

基于改进集成学习算法的海上风电发电功率预测研究

叶 凡¹, 白志阳²

¹上海理工大学管理学院, 上海

²沪东中华造船(集团)有限公司, 上海

收稿日期: 2025年1月24日; 录用日期: 2025年2月17日; 发布日期: 2025年2月27日

摘 要

针对海上风电功率预测的复杂性, 建立以发电功率预测精度为主要优化目标, 以风速、风向、空气密度等多维气象数据输入方式的海上风电功率预测模型, 同时在构建XG-Boost算法的基础上, 提出了一种引入模拟退火算法的参数寻优机制, 以避免算法陷入局部最优解。将该算法应用于企业数据, 并与其他三种算法进行对比, 结果表明: 改进的SA-XG-Boost算法能够有效提高预测精度, 为海上风电企业的稳定运行提供了可靠的支撑。

关键词

海上风电, 发电功率预测, 改进XG-Boost, 皮尔逊相关分析法

Research on Offshore Wind Power Generation Power Prediction Based on Improved Ensemble Learning Algorithm

Fan Ye¹, Zhiyang Bai²

¹School of Management, University of Shanghai for Science and Technology, Shanghai

²Hudong Zhonghua Shipbuilding (Group) Co. Ltd., Shanghai

Received: Jan. 24th, 2025; accepted: Feb. 17th, 2025; published: Feb. 27th, 2025

Abstract

Aiming at the complexity of offshore wind power prediction, an offshore wind power prediction

model is established with the power generation prediction accuracy as the main optimization objective, and multi-dimensional meteorological data input methods such as wind speed, wind direction, air density, etc. At the same time, on the basis of constructing the XG-Boost algorithm, a parameter-seeking optimization mechanism that introduces the simulated annealing algorithm is proposed to avoid the algorithm from falling into the local optimal solution. The algorithm is applied to enterprise data and compared with the other three algorithms, and the results show that the improved SA-XG-Boost algorithm can effectively improve the prediction accuracy and provide reliable support for the stable operation of offshore wind power enterprises.

Keywords

Offshore Wind Power, Generation Power Prediction, Improved XG-Boost, Pearson Correlation Analysis Method

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

中国海上风电资源丰富,在“双碳”背景下,大力发展海上风电有助于加快中国能源转型进程,达成“双碳”目标[1]。与此同时,由于风能的随机性和间歇性,导致风电功率也随之产生波动[2],风电大规模并网将直接影响电力系统稳定运行[3]。对风电功率进行准确预测将提高电力系统对风能的消纳能力,降低弃风率[4],对电网保持平衡具有实践意义。

目前,风电功率预测的方法主要分为物理方法、统计方法、人工智能技术及其组合方法[5]。物理方法主要基于天气预报模型,通过气象和地理信息(包括温度、气压和地形等)对于风电功率进行超短期预测,但需要大量的计算资源和大量的数据来校准[6]。统计方法包括自回归模型和线性预测方法等,当有大量的历史数据可用时,统计方法进行功率预测是可行的。然而,由于历史数据中常蕴含着复杂的非线性关系,传统的统计方法往往难以实现理想的预测效果。基于人工智能技术的预测模型相比物理方法和统计方法更具发展前景,因为其具有潜在的数据挖掘和特征提取能力。近些年受到科学学者的广泛关注。基于支持向量机、人工神经网络、极限学习机和自适应模糊神经网络等人工智能技术的模型通过使误差最小化来处理输入的非线性关系。文献[7]介绍了一种融合了聚类分析与极限学习机的先进预测方法,相较于传统预测模型,该方法显著提升了预测的精确度。文献[8]通过构建基于随机森林、极限学习机及支持向量机的风电功率预测模型,发现随机森林算法在平均预测表现上超越了其他技术。尽管这些预测方法均展现出了良好的预测能力,但它们大多依赖于单一的预测模型,并且模型参数往往采用默认设置或手动调整,这限制了模型达到最佳预测效果的可能性,进而影响了预测性能。针对这一问题,将智能优化算法与预测模型相融合已成为一个值得探索的关键研究领域。文献[9]提出一种结合乌鸦搜索算法和支持向量机的组合模型,与其他智能优化算法相比,该模型在与其他智能优化算法的对比中展现出了更优的预测效果,显著降低了预测误差。同样,文献[10]利用麻雀优化算法来优化极限学习机在风电功率预测中的应用,该方法在鲁棒性和准确性方面均取得了显著提升。通过选择更优的预测模型,并结合智能算法对模型参数进行优化,可以显著提升风电功率预测的精度。

上述研究表明,运用人工智能技术预测海上风电功率能显著提升预测精度,有效应对多变量和非线性特征带来的挑战。然而,过往研究往往止步于直接应用人工智能算法,忽视了模型参数的深度优化,

致使模型性能未达最优。机器学习预测模型因涉及大量计算参数,对数据中的噪声和异常值较为敏感,且模型的验证过程较为复杂。不同计算参数对预测模型的精确程度影响很大[11]。因此,当前亟需开发一种更为简洁、合理且精确的海上风电功率预测方法。

为此,本文引入模拟退火算法来优化极端梯度提升树(XG-Boost),构建了一种基于模拟退火算法下优化极端梯度的风电功率预测模型(SA-XG-Boost),通过模拟退火算法对 XG-Boost 进行参数自适应优化,建立针对海上风电功率的 SA-XG-Boost 预测模型,并采用了真实风电数据进行模型性能的验证。实验结果显示,经过改进的 SA-XG-Boost 模型在预测精度上表现出色,进一步揭示了海上风电功率预测的内在规律。

2. 算法原理

2.1. XG-Boost 算法

极端梯度提升算法被证实为一种高性能的决策树算法(Ma *et al.*, 2023)。决策树是一种常见的机器学习方法,常用于分类和回归问题[12],由于其计算复杂度较低、输出结果直观易懂、不易受到缺失数据的影响,并且能够有效处理与目标变量无关的特征数据等优势,被广泛应用于数据分析和数据挖掘等领域[13]。XG-Boost 算法通过构建目标函数并加入正则化项,有效减少了模型过拟合的问题,为工程仿真提供了既快速又可靠的解决方案。XG-Boost 算法的基本原理如下:针对数据集 $DATA = \{(X_i, Y_i)\}$ 中,算法中每一棵新树都旨在修正前一棵树在预测样本 i (特征向量为 X_i , 真实值为 Y_i) 时产生的误差,从而不断迭代提升预测准确性。所以,对样本 i 在前 t 棵树预测结果如式(1):

$$\hat{y}_i^t = \sum_{i=1}^n f_i(x_i) = \hat{y}_i^{t-1} + f_i(x_i) \quad (1)$$

在公式(1)中: x_i 代表输入数据; $f_i(x_i)$ 则指的是第 t 棵决策对样本 x_i 所给出的预测结果,而 \hat{y}_i^{t-1} 是基于之前 $t-1$ 棵树对样本 i 的预测输出进行修正后得到的。XG-Boost 算法的目标函数,是包括模型误差和正则项,其表达式如式(2)

$$O_{bj}^{(t)} = \sum_{i=1}^n L(y_i, \hat{y}_i) + \sum_{i=1}^l \Omega(f_i) \quad (2)$$

$$\Omega(f_k) = \gamma T + \frac{\lambda}{2} \sum_{i=1}^T w_j^2 \quad (3)$$

在该表达式中, $\Omega(f_k)$ 代表了模型的复杂程度; $L(y_i, \hat{y}_i)$ 则表示模型的损失函数,用于衡量预测值与真实值之间的差异; T 是决策树中叶子节点的数量; γ 作为一个系数,用于调控叶子节点的复杂性; w_j 指的是每个叶子节点的权重;而 λ 则是正则化项的参数。把式(1)和式(3)代入到式(2)中后,目标函数会相应地转化为一个新的形式,如式(4):

$$O_{bj}^{(t)} = \sum_{i=1}^n L[y_i, \hat{y}_i^{t-1} + f_i(x_i)] + \gamma T + \frac{\lambda}{2} \sum_{i=1}^T w_j^2 + C \quad (4)$$

表达式(4)中: C 表示常量,是 $t+1$ 棵树的模型复杂度之和。

对表达式(5)应用泰勒展开的方法时,第 t 步时 $L(y_i, \hat{y}_i)$ 常数,常数对优化过程不产生影响,因此可以忽略这些常数项,为了简化表达,引入以下定义: $f_i(x) = w_{q(x)}$, $g_i = \partial_{\hat{y}_i^{t-1}} L(y_i, \hat{y}_i^{t-1})$, $h_i = \partial_{\hat{y}_i^{t-1}}^2 L(y_i, \hat{y}_i^{t-1})$, w_q 为决策树的预测值, g_i 表示损失函数 L 关于 \hat{y}_i^{t-1} 的一阶偏导数,它反映了损失函数对预测值变化的敏感度; h_i 则为损失函数 L 对 \hat{y}_i^{t-1} 的二阶偏导数,它提供了损失函数曲率的信息。基于这些定义,将目标函

数转换为式(5):

$$O_{bj}^{(t)} = \sum_{i=1}^n \left[g_i w_q(x_i) + \frac{1}{2} h_i w_q^2(x_i) \right] + \gamma T + \frac{\lambda}{2} \sum_{j=1}^T w_j^2 \quad (5)$$

令 $G_j = \sum_{i \in I_j} (g_i)$, $H_j = \sum_{i \in I_j} (h_i)$, 其中 I_j 为第 j 个叶子节点中包含的所有样本, 即 $I_j = \left\{ \frac{i}{q(x_i)} = j \right\}$, 关于 $x_i \in I_j$ 有 $w_q(x_i) = w_j$, 目标函数可以转化为式(6):

$$O_{bj}^{(t)} = \sum_{j=1}^T \left[G_j w_j + \frac{1}{2} (H_j + \lambda) w_j^2 \right] + \gamma T \quad (6)$$

表达式(6)中, G_j 表示汇聚了第 j 个叶子结点内所有样本的一阶偏导数之和; H_j 则代表这些样本的二阶偏导数之和。

对表达式(6)中的 w_j 进行偏导数求解, 令 $\frac{\partial O_{bj}^{(t)}}{\partial w_j} = 0$, 解得 $w_j = -\frac{G_j}{H_j + \lambda}$, 随后, 将这一最优解代入到表达式(3)中, 最终得到的目标函数如式(7):

$$O_{bj}^{(t)} = -\frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \gamma T \quad (7)$$

2.2. 模拟退火(SA)算法

模拟退火算法(SA)灵感来源于固体退火过程中原子结晶行为, 是一种基于概率突跳的智能优化算法。在固体燃烧过程中, 原子或分子因高温而动能增加, 随机无序运动并可能通过碰撞反应生成新化合物[14]。随时间推移, 温度降低, 原子或分子运动逐渐减缓, 无序减少, 最终达到有序状态, 内能最小化, 即化学反应平衡。将这一物理过程类比到算法中, 原子或分子的运动就代表搜索解空间的过程, 而化学反应则相当于接受新解。Metropolis 等人在 1953 年首次提出了这一概念, 然而, 直到 1983 年, Kirkpatrick 等人才成功将其引入组合优化的应用领域中, 作为一种寻优方法, 并命名为“模拟退火算法”。该算法本质上是一种基于概率的随机优化方法, 通过模仿物理退火过程实现智能优化。

算法在数学层面的原理为: 模拟退火采用 Metropolis 抽样规则, Metropolis 抽样规则是一种用于确定是否接受新解的策略, 通常用于在接受较差解的情况下跳出局部最优解, 能够提高发现更优全局解的可能性。基本思想是根据新解相对于当前解的能量差(ΔE)以及当前的温度来确定是否接受新解。该规则基于统计物理学中的 Metropolis 抽样法, 被引入到模拟退火算法中以控制接受劣质解的概率。Metropolis 准则公式为下式(8):

$$p = \exp\left(-\frac{E(x_{\text{new}}) - E(x_{\text{old}})}{kT}\right) \quad (8)$$

式中, k 为玻尔兹曼常数; T 为当前温度值($^{\circ}\text{C}$)。

目标函数 $E(x)$ 为当前微观状态下固体的能量状态, 当系统从一个能量状态 $E(x_{\text{old}})$ 变化到新的能量状态 $E(x_{\text{new}})$ 时, 以概率 $P(x_{\text{old}} \rightarrow x_{\text{new}})$ 接受新解, 概率 P 可用式(9)计算:

$$P(x_{\text{old}} \rightarrow x_{\text{new}}) \begin{cases} 1 & E(x_{\text{max}}) \leq E(x_{\text{old}}) \\ \exp\left(-\frac{|\Delta E|}{kT}\right) > \text{rand}[0,1] & E(x_{\text{max}}) > E(x_{\text{old}}) \end{cases} \quad (9)$$

这个公式表明, 在整个算法求解的迭代过程中, 对现有旧解 x_{old} 生成的(0,1)之间的随机数, 产生一个新解 x_{new} 。若 $E(x_{new}) \leq E(x_{old})$, 那么将采用新解 x_{new} 替换旧解 x_{old} ; 若 $E(x_{new}) > E(x_{old})$, 则接受这一较差解 x_{new} 的概率 p 较低。由公式可以看出, 模拟退火算法迭代过程中逐渐降低温度值 T , 由公式可知, 正是由于当温度 T 较高时, 此算法设计以较高概率接纳非最优解, 从而具备了逃离局部最优解的能力, 当新解与当前解的能量差距微小(意味着新解相对更优)或系统“温度”参数较高时, 算法采纳新解的可能性会相应增大, 而当温度趋近于零时, 算法倾向于接受更优的解。在不同温度下, 算法的搜索方式也有所不同, 在温度较高时, 算法进行全局搜索, 而在温度较低时则转为局部搜索。通过这种概率计算, 高温下接受较差解的概率较大, 有助于跳出局部最优解; 而低温时, 接受劣质解的概率较小, 有助于算法收敛到全局最优解。在迭代过程中, 温度逐渐降低, 从而提高解的精度。这也是快速合理并正确找到最优解的方法。

2.3. SA-XG-Boost 算法构建

在运用 XG-Boost 算法进行非线性回归预测时, 参数的精准选择对于确保预测结果的准确性至关重要。XG-Boost 的预测性能主要受 `n_estimator`、`learning_rate`、`max_depth` 以及 `gamma` 这几个关键参数的影响, 具体见表 1。通常在计算中通过人为调参来获得更优的拟合效果, 但这种方法难以避免人为因素带来的主观偏见, 并且容易使结果陷入局部最优的困境。

Table 1. Computational parameters of XG-Boost model
表 1. XG-Boost 模型计算参数

参数	参数说明
<code>n_estimator</code>	该参数决定了模型的预测能力, 其值越大则模型的学习及数据分析能力越强, 预测、结果越精确, 但参数过大将浪费计算资源。
<code>learning_rate</code>	该参数决定了迭代的步长, 参数过大将影响运行的准确率, 过小则影响运行的速度。
<code>max_depth</code>	该参数定义了树的最大深度, 用于调节算法以防止过度拟合。当此参数设定得更大时, 模型将能够更加细致地捕捉数据特征, 进行更深入的学习。
<code>gamma</code>	该参数设定了决策树节点分裂时所需达到的最小损失函数减少量, 其值越大, 算法在决策树构建过程中会表现得更为谨慎, 即更加保守。

为了获得更加精确的预测结果, 将模拟退火(SA)算法的优化机制与 XG-Boost 算法相结合, 对 XG-Boost 的参数配置进行寻优调整。图 1 详细展示了 SA-XG-Boost 预测算法的执行流程, 具体步骤如下:

Step 1: 导入数据集后, 首先进行数据的归一化处理, 接着将数据集分割为测试集与训练集两部分。

Step 2: 模拟退火算法进行初始化设置, 包括确定算法的起始温度以及模型的基础参数。

Step 3: 明确 XG-Boost 模型中待优化参数的范围, 并基于此范围生成模拟退火算法的初始参数集合。

Step 4: 根据每次得到的能量状态, 依据 Metropolis 准则来评估每次迭代产生的能量状态。

Step 5: 通过多次迭代, 将预测模型的均方误差(MSE)作为衡量优化效果的关键评价指标, 迭代达到预设的停止条件时, 算法将终止, 并输出在最优能量状态下所对应的参数。

Step 6: 将迭代后的最佳参数应用到 XG-Boost 模型, 获得训练完成的最佳预测模型, 接着评估模型性能, 从而实现海上风电功率的预测并输出结果。

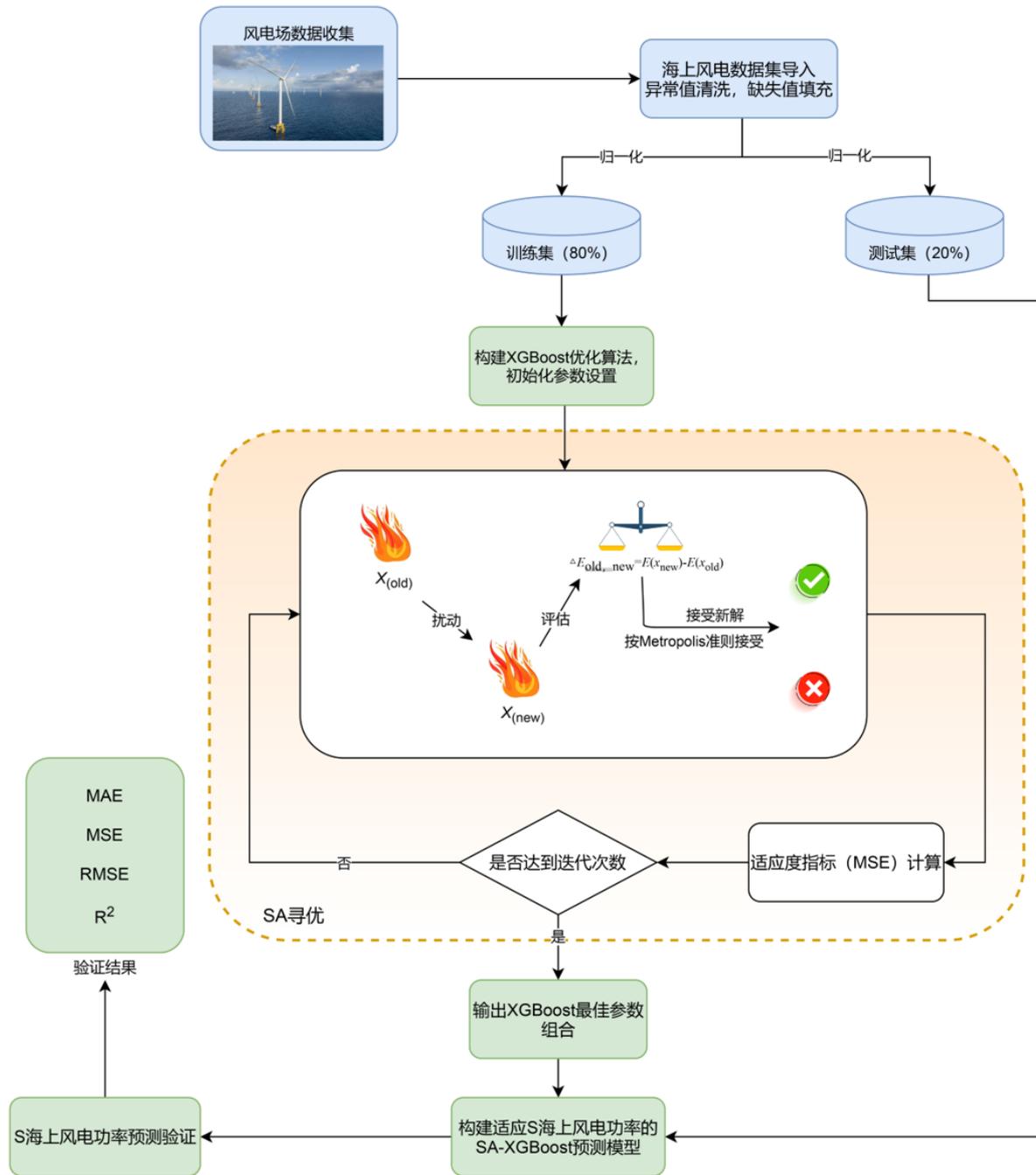


Figure 1. Flowchart of SA-XG-Boost algorithm
图 1. SA-XG-Boost 算法流程图

3. 基于 SA-XG-Boost 的海上风电功率预测模型构建

3.1. 数据集采集

为了评估改进后的 SA-XGBoost 对海上风电功率预测精度问题, 本文以江苏某海上风电场为例, 数据集使用的是 2022 年 1 月 10 日 0 时至 2022 年 12 月 20 日 23 点 50 分的风电发电数据集, 提供了将近一年时间的发电数据, 这些数据将作为模型的输入来训练预测模型。选择这样一段完整的时间周期进行观

察有助于捕捉风电功率的季节性变化和日常波动, 从而提高预测结果的准确性和可靠性。

3.2. 风电特征相关性分析

在清洗采集到的风电场数据时, 首要步骤是评估数据质量, 并识别出与风电功率关联性最强的气象要素。本研究运用了皮尔逊相关分析法来评估江苏某海上风电场历史数据的质量, 通过计算两个变量 x 和 y 之间的皮尔逊相关系数(其定义如式(10))来衡量它们之间的相似程度:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (10)$$

表达式中 \bar{x} 、 \bar{y} 分别为 x 、 y 中元素的平均值。

显然, Pearson 相关系数 r_{xy} 的取值范围在 $[-1, 1]$ 间的实数。当 $r_{xy} > 0$ 时, 表示两个变量之间存在正相关关系, 相反, 则表示负相关关系。 $|r_{xy}|$ 数值越高, 则表明 x 和 y 之间的相关性越强。

分析结果见表 2。其中列出了湍流强度(TI)、风速(WS)、风向(WD)、温度(AT)、大气密度(AD)、风电功率(WP)、湿度(RH)气压(PA)等 8 组时间序列之间的相关系数。从表 2 可以看出, 该风电场的发电功率与风速相关性最强, 为 0.997; 与风向、湿度、大气密度、湍流强度相关性次之, 相关系数为 0.868、0.818、0.814 和 0.781; 此外, 风电功率与温度及气压相关性较低, 相关系数分别为 0.244、0.082。

Table 2. Correlation between wind power and meteorological factors

表 2. 风电功率与气象因素的相关性

变量	TI	WS	WD	AD	RH	AT	PA	WP
TI	1	0.812	0.640	0.800	0.818	0.455	0.095	0.781
WS	0.812	1	0.849	0.819	0.826	0.262	0.095	0.997
WD	0.640	0.849	1	0.726	0.827	-0.049	-0.275	0.868
AD	0.800	0.819	0.726	1	0.706	0.275	0.024	0.814
RH	0.818	0.826	0.827	0.706	1	0.060	-0.176	0.818
AT	0.455	-0.262	-0.049	0.275	0.060	1	0.860	0.244
PA	0.095	0.095	-0.275	0.024	-0.176	0.860	1	0.082
WP	0.781	0.997	0.868	0.814	0.818	0.244	0.082	1

在得出相关系数后, 为了更清晰地展示各特征间的相关性绘制了这段时间内风电影响因素与风电功率之间的相关性热力图见图 2 所示。从图中可以直观地观察到, 风速、风向、湿度、空气密度、湍流强度因素与风电功率之间存在显著的相关性, 这些相关性的强度在热力图中以深红色区域为主, 在图 2 中, 风速与风电功率的相关性最为突出。

根据皮尔逊相关分析法, 选取与发电功率相关性最强的五个因素作为特征输入量。这些因素包括风速、风向、空气密度、湿度、湍流强度和风电功率, 构成了对海上风电场发电情况的全面描述。在数据集, 未记录或缺失的数据点在本次实验中已被剔除, 以确保数据的完整性和分析的有效性。总共包含了超过万个样本, 每个样本对应一个 10 分钟时间窗口内的统计信息。由于数据量庞大, 该数据集的部分数据集展示如表 3 所示。



Figure 2. Correlation heat map
图 2. 相关性热力图

Table 3. Historical dataset for the S offshore wind project
表 3. S 海上风电项目历史数据集

日期	风速	风向	空气密度	湿度	湍流强度	功率
2022/4/19 15:00	5.678	210.531	1.237	42.120	0.035	5.523
2022/4/19 15:10	5.778	213.120	1.238	42.901	0.0345	6.163
2022/4/19 15:20	5.977	209.697	1.237	43.857	0.033	8.153
2022/4/19 15:30	6.076	211.328	1.237	43.491	0.049	8.843
					
2022/4/19 16:00	5.976	212.288	1.236	43.068	0.033	7.686
2022/4/19 16:10	5.776	215.734	1.236	42.598	0.034	6.327
2022/4/19 16:20	5.877	218.556	1.237	42.750	0.034	7.310

3.3. 算法结果分析和对比

Table 4. Comparison of experimental evaluation of each algorithm
表 4. 各算法实验评价对比

算法	MAE	RMSE	MSE	R ²
SA-XG-Boost	0.9405	1.8335	3.4073	0.988
XG-Boost	1.9473	3.0709	9.4606	0.961
Random Forest	4.9246	6.6790	44.6882	0.937
Bayesian Ridge	6.0247	8.5888	73.8473	0.895

在本次研究中, 选用了常用的评估标准来评估模型的预测性能, 分别是一元多项式回归方程拟合度 R^2 、平均绝对误差(Mean Absolute Error, MAE)、均方根误差(Root Mean Square Error, RMSE)和均方误差(Mean-square error, MSE), 结果如下表 4 所示。

观察表 4 可得, SA-XG-Boost 模型的 R^2 值达到了 0.988, 比其他几种模型的拟合系数更优, 但单一的拟合系数可能会由于出现过拟合而产生较高的值, 因此, 为了全面评估预测模型的性能, 同时运用均方根误差(RMSE)和平均绝对误差(MAE)这两个指标来评估模型的预测效果。根据表格本文所提模型的均方根误差 RMSE 分别低于第二、三种模型的 6.6790 和 8.5888, 并比改进前的误差值 3.0709 有明显的降低, 对于 MAE 这个评价指标, 对比表中数据同样可以发现, SA-XG-Boost 预测模型比起另外三种预测结构具有更小的 MAE, 其值分别降低了 1.0068、3.9841 和 5.0842。说明其预测数据相对于真实数据而言离散程度更小, 即预测模型的稳定性更好, 综上所述, 所提出的 SA-XG-Boost 模型能够显著提升预测精度。证明了改进算法在海上风电功率预测中的可行性与优越性。

根据实验结果数据, 利用误差图作为统计分析工具得出各个算法误差分析, 如图 3~图 6 所示。从图中可以得出 SA-XG-Boost 算法在数据离散程度上相较于其他三种算法更小, 数据更为集中, 说明 SA-XG-Boost 的改进操作可以有效避免算法陷入局部最优值, 增强其对参数全局寻优能力, 从而提高算法精度。

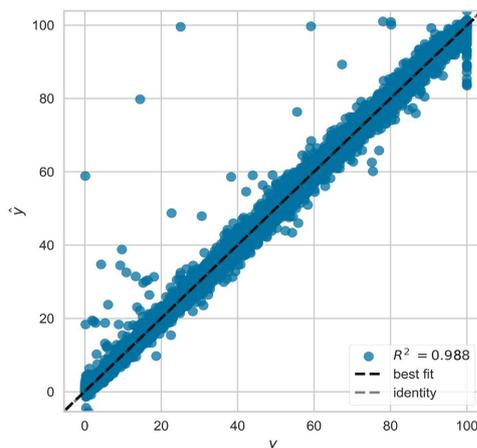


Figure 3. SA-XG-Boost algorithm error plot
图 3. SA-XG-Boost 算法误差图

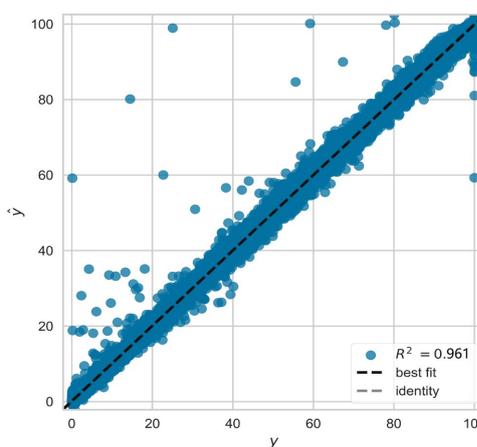


Figure 4. XG-Boost algorithm error plot
图 4. XG-Boost 算法误差图

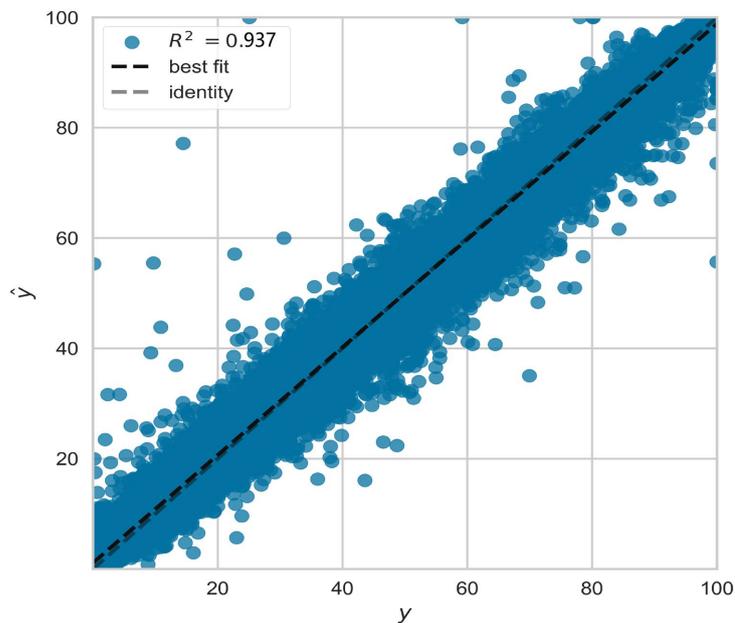


Figure 5. Random forest algorithm error plot
图 5. 随机森林算法误差图

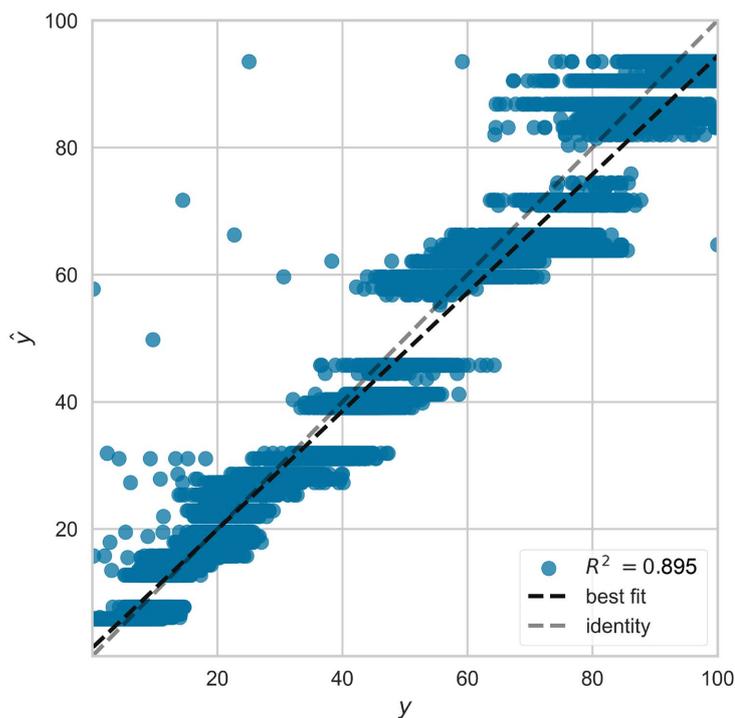


Figure 6. Bayesian ridge regression algorithm error plot
图 6. 贝叶斯岭回归算法误差图

为更直观比较不同结构的预测性能, 图 7~图 10 给出了预测结果的残差图。可以看出, 由随机森林和贝叶斯岭回归构成的经预测模型的预测结果较差, XG-Boost 模型预测稍好, 本文所提出的 SA-XG-Boost 模型结构在残差评价指标均为最优。

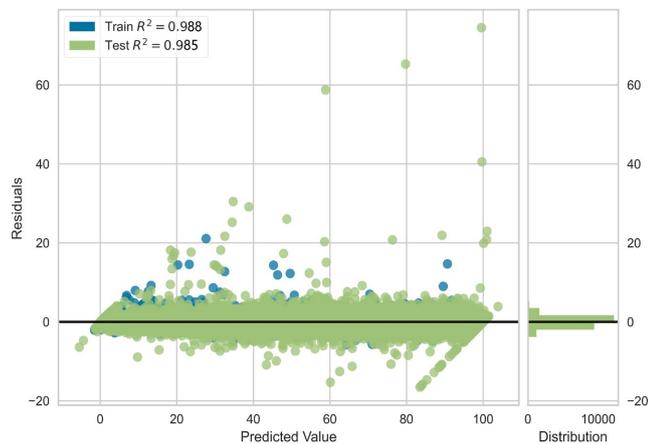


Figure 7. SA-XG-Boost algorithm residual plot
图 7. SA-XG-Boost 算法残差图

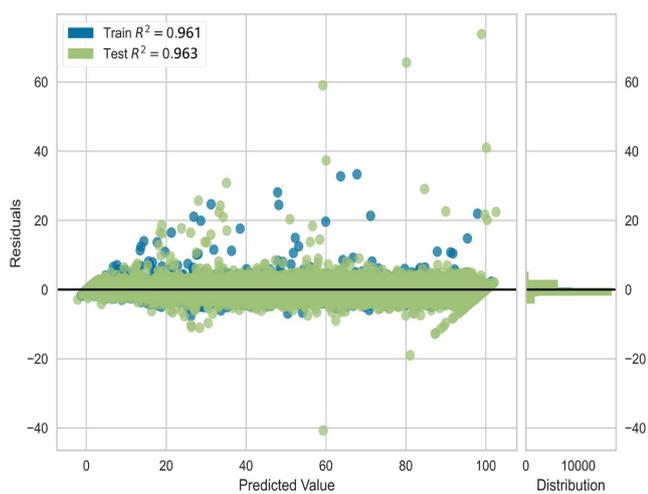


Figure 8. XG-Boost algorithm residual plot
图 8. XG-Boost 算法残差图

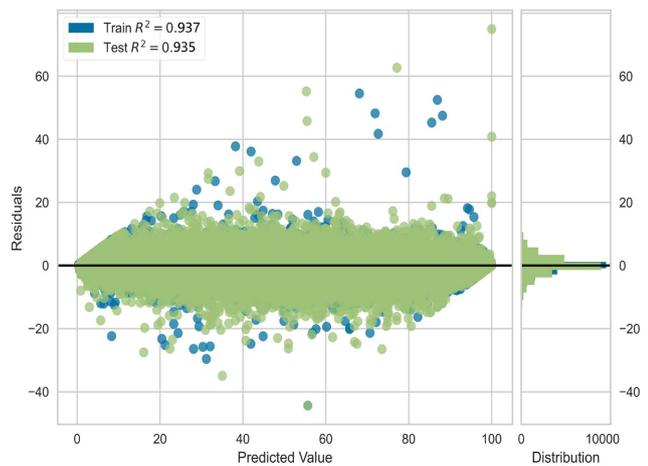


Figure 9. Random forest algorithm residual plot
图 9. 随机森林算法残差图

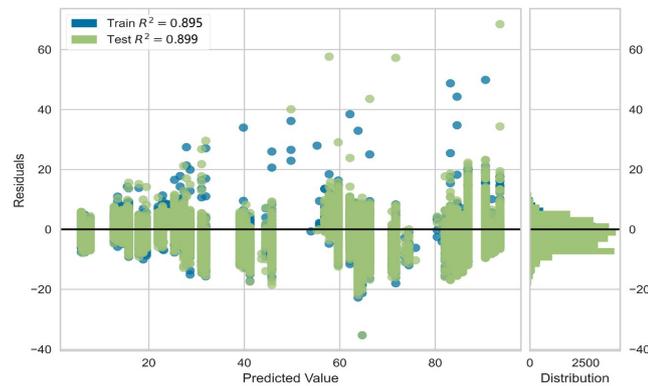


Figure 10. Bayesian ridge regression algorithm residual plot
图 10. 贝叶斯岭回归算法残差图

4. 结论

本文对经海上风电场风电预测问题进行贴近实际的拓展研究, 为改善海上风功率预测精度, 本文提出一种基于改进 XG-Boost 算法的海上风电功率预测模型, 得出以下主要结论

(1) 为了确保数据集的精准度和完整性。进一步地, 利用皮尔逊相关分析法, 科学筛选出了与风电功率最具相关性的气象因素, 为模型特征选择提供了有力支撑。

(2) 在模型构建阶段, 本研究引入了 SA 算法对 XG-Boost 算法的参数进行寻优, 并与其他经典预测算法进行了全面对比分析。实验结果显示, SA-XG-Boost 模型在拟合度、均方根误差、平均绝对误差等多个关键评价指标上均表现出了优异的性能, 从而大幅提高了海上风电功率预测的准确性。

(3) 本研究的成果不仅验证了 SA-XG-Boost 算法在海上风电功率预测中的有效性, 同时也在对风电功率预测方面的实用价值, 为未来海上风电场的运行优化、能源管理以及电网调度提供了重要参考。未来研究可进一步拓展至其他影响因素, 如海洋环境、设备状态等, 以持续优化预测模型, 提升预测性能。

参考文献

- [1] 张龙, 孙丹, 马志豪, 等. 海上风电柔直送出系统改进频率协同控制策略[J]. 太阳能学报, 2024, 45(10): 478-486.
- [2] 赵连娟. 基于混沌神经网络的风电预测系统建模研究[J]. 自动化技术与应用, 2018, 37(10): 105-108+135.
- [3] 张健, 田海. 基于 DE-XG-Boost 的短期风电功率预测[J]. 信息技术, 2024(7): 136-142.
- [4] 刘雅婷. 考虑转折性天气的日前风电功率预测及爬坡事件检测方法研究[D]: [硕士学位论文]. 济南: 山东大学, 2024.
- [5] 王颀, 刘兴杰, 梁英, 等. 一种基于 MGWO-Informer 的超短期风电功率预测方法[J]. 太阳能学报, 2024, 45(11): 477-485.
- [6] 罗潇远, 刘杰, 杨斌, 等. 基于改进鱼鹰优化算法与 VMD-LSTM 的超短期风电功率预测[J]. 太阳能学报, 2024: 1-9.
- [7] 陈天阳, 钱政, 荆博, 等. 基于 K-means++与 ELM 的短期风电功率预测模型研究[J]. 电测与仪表, 2024, 61(6): 45-50.
- [8] 陈金车, 王金艳, 苏士翔, 等. 3 种机器学习方法预报风电功率的对比[J]. 兰州大学学报(自然科学版), 2022, 58(1): 124-129+136.
- [9] 李国全, 高建宇, 白天宇, 等. 基于 SVM 与改进型乌鸦搜索算法的风电功率预测方法[J]. 国外电子测量技术, 2022, 41(2): 40-45.
- [10] 刘栋, 魏霞, 王维庆, 等. 基于 SSA-ELM 的短期风电功率预测[J]. 智慧电力, 2021, 49(6): 53-59+123.
- [11] 李振阳, 张宝岗, 熊信, 等. 基于 PSO-XGBoost 的露天矿山 PPV 预测模型研究[J]. 黄金科学技术, 2024, 32(4): 620-630.

- [12] 周捷盼. 基于大数据的跨境电商平台供应商信用评估研究[J]. 现代商业, 2024(22): 99-102.
- [13] 张铭良. 基于机器学习的钢柱底板连接破坏模式识别研究[D]: [硕士学位论文]. 济南: 山东建筑大学, 2024.
- [14] 李书益. 基于遗传模拟退火粒子群算法的微电网优化运行研究[D]: [硕士学位论文]. 南京: 南京邮电大学, 2021.