

基于随机森林和改进高斯过程的风电机组 功率曲线模型

许灿, 缪书唯

(三峡大学 电气与新能源学院, 湖北 宜昌 443002)

摘要:功率曲线在风电机组状态监测和风电功率预测等领域发挥着重要作用。为提高风电机组功率曲线建模精度,首先应用随机森林算法筛选出影响风能捕捉能力的关键变量。然后,将筛选出的变量输入到改进后的高斯过程模型中,以提高计算效率。最后,应用4种不同的指标评估模型的精度,并应用熵权法处理指标之间的潜在冲突,从而形成一个综合性评估指标,全面衡量功率曲线模型的精度。通过英国某风电场SCADA数据验证所提方法的有效性,结果显示:与现有6类常见方法相比,所提方法在模型精度方面有一定提升。

关键词:风电机组;功率曲线;随机森林;改进高斯过程;熵权法

中图分类号:TK83 **文献标识码:**A **DOI:**10.19457/j.1001-2095.dqcd25930

Wind Turbine Power Curve Model Based on Random Forest and Improved Gaussian Process

XU Can, MIAO Shuwei

(College of Electrical Engineering and New Energy, China Three Gorges University,
Yichang 443002, Hubei, China)

Abstract: Wind turbine condition monitoring and wind power prediction both rely heavily on power curves. Firstly, to increase the modeling accuracy of wind turbine power curves, the random forest technique was used to screen the important variables that influence wind energy capture ability. Then, the screened variables were fed into the improved Gaussian process (GP) model, which improved computational efficiency. Finally, four separate metrics were used to evaluate the model's correctness, and the entropy weight approach was used to resolve any potential conflicts between the metrics, resulting in a comprehensive assessment metric that measured the quality of the power curve model. The suggested approach's effectiveness was validated using supervisory control and data acquisition (SCADA) data from a wind farm in the United Kingdom, and the findings reveal that the proposed method improves model accuracy when compared to the current six types of conventional methods.

Key words: wind turbine; power curve; random forest; improved Gaussian process; entropy weight method

截至2022年年底,全球风能装机容量已达940 GW^[1]。风电机组功率曲线,凭借其能够识别风电机组功率与风速之间的非线性复杂关系,在风电机组功率估算和风机状态监测等领域扮演着至关重要的角色^[2]。

现有文献针对建立风电机组功率曲线模型展开了一系列研究。例如,国际标准比恩法^[3]采用较长时间段内的10 min平均运行数据。这些数据按风速等级划分为若干分区,进而计算每个

分区的平均风速和功率值,以构建风力发电机组的功率曲线模型。文献[4]应用k近邻(k-nearest neighbor, KNN)算法建立风电机组功率曲线模型,结果表明该算法优于多层感知器和其它3类集成学习算法。文献[5]应用四参数和五参数logistic表达式建立风电机组功率曲线模型,并利用遗传算法等先进算法对这些表达式的参数进行求解,得到风电机组功率曲线的最佳参数模型。上述研究中,风速被选定为唯一的输入变

作者简介:许灿(1999—),女,硕士研究生,主要研究方向为风电机组最优运行,Email:JYasc_A@163.com

通讯作者:缪书唯(1989—),男,博士,副教授,主要研究方向为风电场风况建模、风电场并网系统可靠性评估和优化,

Email:jabker@163.com

量,用于构建风电机组功率曲线模型。文献[6]研究风向作为影响风电机组功率的关键因素,而文献[7]探讨风切变和湍流强度对功率的具体影响。这两项研究分别探讨了风向、风切变和湍流强度这几种不同的影响因素,展示了这些因素对风电机组功率的显著影响。然而,它们在考虑其他潜在影响因素方面存在不足,这些因素同样对风电机组的功率表现具有重要影响。随着研究的深入,文献[8]应用径向基神经网络算法,基于4个变量构建了风电机组的功率曲线模型。同样,文献[9]应用高斯过程(Gaussian process, GP)算法,随机选取6个变量进行功率曲线建模。这两项研究通过扩大变量范围,为风电机组功率曲线模型的建立提供了新的视角。然而,该研究在变量选择上显得较为主观,并且主要使用单一的指标评估模型的性能。此方法可能在一定程度上限制了模型的预测精度和准确性。

本文在文献[9]的基础上进行改进,提出基于随机森林和改进高斯过程的风电机组功率曲线模型。本文的主要贡献如下:

- 1) 随机森林算法筛选出影响风电机组功率的关键变量;
- 2) 改进高斯过程进行风电机组功率曲线模型的建立;
- 3) 熵权法解决4类指标冲突,得到综合评估指标;
- 4) 应用风电场正常运行过程中的实时数据对该方法进行评估,并与其他常用方法进行比较以验证模型的有效性。

1 基于随机森林的关键变量选择

随机森林是一种集成机器学习方法,通过随机重采样技术和节点随机分裂技术生成多棵决策树,并利用投票机制得出最终分类结果^[10]。相比其他方法,随机森林能够处理复杂的非线性关系和高维数据,并通过集成多棵树有效减少过拟合的风险。

在风电机组功率曲线建模中,通常涉及多个输入变量(如风速、风向、温度等)。这些变量及其复杂的相互作用关系可以被随机森林有效地分析和捕捉。

因此,本文应用随机森林算法进行变量选择,从原始的 n 个变量中筛选出 m 个变量($m < n$)。随机森林算法进行变量选择的实现步骤如下^[11]:

步骤1:从 N 个样本、 n 个变量中随机有放回地抽取 $z(z < N)$ 个样本,并进行 G 次重复采样,以获得 G 个子训练集。其中,未被选中的数据构成袋外数据(out-of-bag, OOB)。

步骤2:针对第 $g(g=1, 2, \dots, G)$ 个子训练集,独立地构建决策树并进行训练。每个决策树计算其在对应的OOB上的投票结果 L_g 。

步骤3:随机将OOB中的变量 $A_i(i=1, 2, \dots, n)$ 施加扰动,重新计算决策树在对应OOB中获得的投票结果 $L_{g,i}$ 。

步骤4:通过计算每个变量的重要性分值 P_i 并进行降序排列,以确定最重要的变量来构建最优的变量组合,其中 P_i 越大表示该变量越重要。其对应公式如下式所示^[11]:

$$P_i = \frac{1}{G} \sum_{g=1}^G (L_g - L_{g,i}) \quad (1)$$

2 基于改进高斯过程的风电机组功率曲线模型

记 $\mathbf{x} = [x_1, x_2, \dots, x_m]$ 为 m 个影响风电机组功率的变量,例如 \mathbf{x} 可包含风速、风向、转子转速、桨距角等, y 为风电机组功率。高斯过程作为一种强大的非参数贝叶斯方法,能够准确地建模和预测这种复杂的关系。因此,本文应用高斯过程模型,将 \mathbf{x} 和 y 之间的函数关系表达如下^[12]:

$$y = f(\mathbf{x}) + \varepsilon \quad (2)$$

式中: ε 为均值为0、方差为 σ_n^2 的正态随机变量; $f(\mathbf{x})$ 为高斯过程,其特性由均值 $m(\mathbf{x})$ 和协方差函数 $k(\mathbf{x}, \mathbf{x}')$ 定义^[13]。

假设已知训练集输入 $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_B\}$,训练集输出 $\mathbf{y} = \{y_1, \dots, y_i, \dots, y_B\}$,此时若已知测试样本 \mathbf{X}_* ,则可得到预测功率 y_* 的后验概率分布为^[12]

$$y_* | \mathbf{X}, \mathbf{y}, \mathbf{X}_* \sim N(\boldsymbol{\mu}_*, \boldsymbol{\Sigma}_*) \quad (3)$$

其中

$$\bar{y}_* = \mathbf{K}(\mathbf{X}_*, \mathbf{X}) \mathbf{K}^{-1}(\mathbf{X}, \mathbf{X}) \mathbf{y} \quad (4)$$

$$\boldsymbol{\Sigma}_* = \mathbf{K}(\mathbf{X}_*, \mathbf{X}_*) - \mathbf{K}(\mathbf{X}_*, \mathbf{X}) \times [\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}]^{-1} \mathbf{K}(\mathbf{X}, \mathbf{X}_*) \quad (5)$$

式中: $\mathbf{K}(\mathbf{X}, \mathbf{X})$ 为协方差函数 k 在训练集 \mathbf{X} 上的所有成对组合 (x_i, x_j) 之间的评估结果所形成的 $B \times B$ 维协方差矩阵; $\mathbf{K}(\mathbf{X}, \mathbf{X}_*)$ 为测试数据 \mathbf{X} 与训练数据 \mathbf{X}_* 之间的 $B \times B_*$ 阶协方差矩阵; $\mathbf{K}(\mathbf{X}, \mathbf{X}_*) = \mathbf{K}(\mathbf{X}_*, \mathbf{X})^\top$; \mathbf{I} 为 $B \times B$ 维单位矩阵; $\mathbf{K}(\mathbf{X}_*, \mathbf{X}_*)$ 为测试点 \mathbf{X}_* 自身的 $B_* \times B_*$ 维协方差矩阵; $\bar{y}_*, \boldsymbol{\Sigma}_*$ 分别为测试样本 \mathbf{X}_* 对应预测功率 y_* 的预测均值和预测

方差。

在改进的高斯过程回归模型中,本文通过LDL分解方法将协方差矩阵分解为下三角矩阵和对角矩阵的乘积,从而简化计算过程。通过LDL分解,矩阵K的逆可以计算为

$$K^{-1} = (L^T)^{-1} D^{-1} L^{-1} \quad (6)$$

则预测均值 \bar{y}_* 为

$$\bar{y}_* = K(X_*, X) (L^T)^{-1} D^{-1} L^{-1} y \quad (7)$$

3 基于熵权法的综合评价指标在模型性能评估中的应用

为评估风电机组功率曲线模型的效果,应用4种典型指标:平均绝对误差(mean absolute error, MAE)^[14]、均方根误差(root mean square error, RMSE)^[14]、确定系数(the coefficient of determination, R²)^[15]和平均绝对值误差百分比(mean absolute percentage error, MAPE)^[14],为保证这些指标数值方向的一致性,本文应用1-R²代替原有R²。

熵权法是一种在多指标评价中常用的赋权方法。该方法依赖于数据本身的信息量来分配权重,从而避免了人为因素的影响。此外,熵权法通过信息熵反映数据的离散程度和信息量,能够更好地揭示数据的内在结构和规律,从而分配更合理的权重。考虑到不同指标之间的差异,本文应用熵权法^[16]为这4个指标分配不同权重,从而得到综合评价指标。熵权法的具体步骤如下:

步骤1:收集M个样本在4种预测指标下的取值,构成决策矩阵:

$$C = \begin{bmatrix} c_{1,MAE} & c_{1,RMSE} & c_{1,1-R^2} & c_{1,MAPE} \\ c_{2,MAE} & c_{2,RMSE} & c_{2,1-R^2} & c_{2,MAPE} \\ \vdots & \vdots & \vdots & \vdots \\ c_{M,MAE} & c_{M,RMSE} & c_{M,1-R^2} & c_{M,MAPE} \end{bmatrix} \quad (8)$$

步骤2:对决策矩阵C进行标准化处理,这里仅列举指标MAE的标准化方式,其它指标类似。

$$Q_{i,MAE} = \frac{c_{i,MAE}}{\sum_{i=1}^M c_{i,MAE}} \quad (9)$$

步骤3:计算评价指标MAE的信息熵:

$$E_{MAE} = \frac{\sum_{i=1}^M Q_{i,MAE} \ln Q_{i,MAE}}{\ln M} \quad (10)$$

当 $Q_{i,MAE}=0$ 时,令 $Q_{i,MAE} \ln Q_{i,MAE}=0$ 。

步骤4:计算指标MAE的权重:

$$w_{MAE} = \frac{1 - E_{MAE}}{(1 - E_{MAE}) + (1 - E_{RMSE}) + (1 - E_{1-R^2}) + (1 - E_{MAPE})} \quad (11)$$

w_{MAE} 越大,则对应指标MAE在综合评价中权重越高。 w_{RMSE} , w_{1-R^2} 和 w_{MAPE} 的计算方法类似。

最后,得到综合指标T的表达式如下所示:

$$T = w_{MAE} S_{MAE} + w_{RMSE} S_{RMSE} + w_{1-R^2} S_{1-R^2} + w_{MAPE} S_{MAPE} \quad (12)$$

式中: S_{MAE} , S_{RMSE} , S_{1-R^2} , S_{MAPE} 分别为风电机组功率曲线模型MAE, RMSE, $1-R^2$ 和MAPE的取值。

4 算例分析

4.1 英国Penmanshiel风电场的SCADA数据

本文收集英国Penmanshiel风电场2021-01-01至2021-02-16的每10 min SCADA运行数据,用于实验分析^[17]。原始数据集中包含6 768条记录,涉及的风电机组型号为Senvion MM82,用以进行实验分析。

在数据预处理阶段,首先移除功率值为零或负的记录。接着,应用四分位法^[18]对剩余数据进行过滤,从而获得5 538条有效数据。这些数据包含15个不同的变量,包括实测功率、风速和转子转速等。随后,通过随机森林算法对这些变量的重要性进行评估,结果按照重要性分值进行降序排列,并在表1中展示。

表1 15个变量的重要性分值

Tab.1 Importance scores for 15 variables

变量	重要性分值/%	变量	重要性分值/%
密度调整风速	4.65	定子温度	1.24×10^{-2}
风速	4.48	机舱温度	8.50×10^{-3}
转子转速	9.98×10^{-1}	环境温度	6.84×10^{-3}
湍流强度	5.40×10^{-1}	叶片位置	4.72×10^{-3}
桨距角	2.06×10^{-1}	齿轮箱油温	3.32×10^{-3}
叶尖速比	1.19×10^{-1}	转子轴承温度	3.25×10^{-3}
变压器温度	2.82×10^{-2}	偏航误差	7.97×10^{-4}
风向	1.94×10^{-2}	—	—

为简化输入变量,同时尽可能多地保留变量信息,仅选择重要性高的部分变量。本文选择以下6个变量作为模型输入:密度调整风速、转子转速、湍流强度、桨距角、叶尖速比和变压器温度。

4.2 验证本文功率曲线模型

将上述选择的变量输入到改进的高斯过程模型中,得到的验证结果如图1所示。

如图1b所示,大多数归一化残差集中在-0.1~0.1之间,归一化残差可由下式计算得到^[19]:

$$H_i = \frac{S_i - \hat{S}_i}{\max(\hat{S}_i)} \quad i = 1, 2, \dots, 1000 \quad (13)$$

式中: S_i 为功率的实际值; \hat{S}_i 为功率预测值。

将实际功率与预测功率的归一化值转换回原始的功率度量范围,得到如图2所示的风电机组实际功率与本文模型预测功率的散点图。

为验证改进后的高斯过程的有效性,本文给出QR分解、SVD分解、Cholesky分解和LDL分解法计算高斯过程逆矩阵的时间对比,如图3所示。该程序用Matlab编写,计算机处理器为Intel(R)

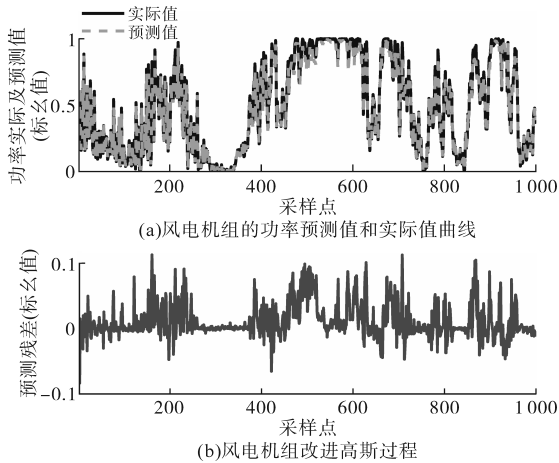


图1 高斯过程功率曲线模型验证结果

Fig.1 Gaussian process power curve model validation results

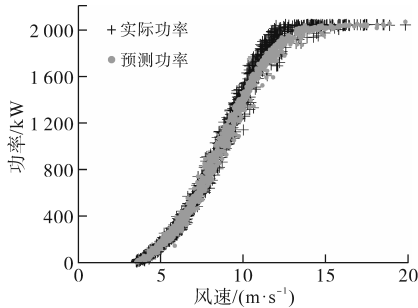


图2 风电机组实际功率与本文模型预功率的散点图

Fig.2 Scatter plot of the actual power of wind turbines against the predicted power of the model in this paper

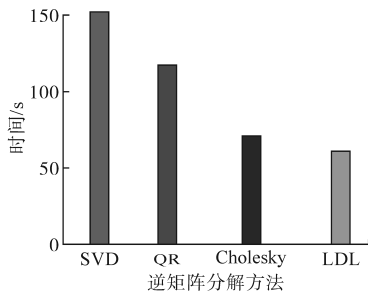


图3 逆矩阵分解方法的时间对比柱状图

Fig.3 Time-contrast histogram of inverse matrix decomposition methods

Pentium(R) Gold G6405 CPU @ 4.10 GHz。

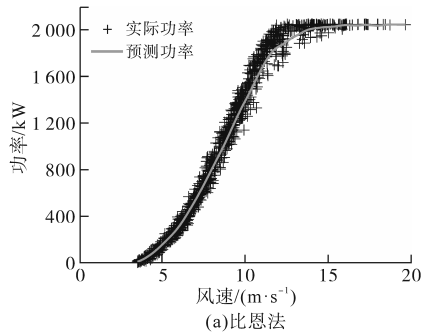
从图3中可以看到,在求逆矩阵的过程中,LDL分解所需的时间明显短于其他3种方法。这一结果有效地展示了LDL分解在效率方面的潜在优势。特别是在处理大规模数据集时,此发现为选择合适的矩阵分解技术提供了重要依据,有助于优化高斯过程模型的计算效率。

4.3 与现有风电机组功率曲线模型方法的对比

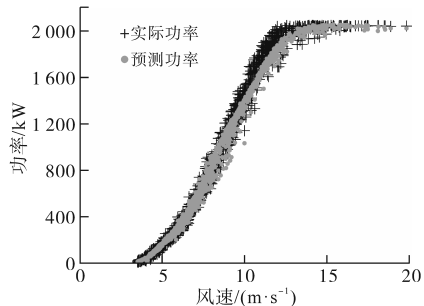
为评估本文提出的功率曲线模型的有效性,本文对比6种不同的方法:k近邻算法^[20]、比恩法(bins method, BM)^[21]、人工神经网络(artificial neural network, ANN)算法^[22]、XGBoost算法^[23]、15个变量的高斯过程模型(简称GP-15)以及原始高斯过程模型^[9](简称GP-0)。值得注意的是,上述模型除比恩法只应用风速作为单一输入变量以及原始高斯过程模型应用主观选择的多变量输入外,其它模型均应用与本文模型相同的多变量输入策略。

由于篇幅限制,图4仅展示了比恩法和人工神经网络算法所绘制的功率曲线。可以看到,两种传统方法在从切入风速至切出风速的区间内,其预测效果均低于本文所提模型(见图2)。

为更客观分析本文模型的有效性,本文应用上述所提到的MAPE, MAE, RMSE和1-R²4种指标计算不同功率曲线模型方法的预测精度,结果如表2所示。



(a)比恩法



(b)神经网络算法

图4 不同功率曲线模型的散点图

Fig.4 Scatter plot of different power curve models

表2 不同功率曲线模型的四类预测指标

模型	MAPE/%	MAE/kW	RMSE/kW	1-R ² /%
KNN	7.70	44.92	74.73	1.18
ANN	6.09	44.02	69.55	1.01
XGBoost	8.76	39.48	58.38	0.71
GP-15	7.34	44.29	64.59	0.87
GP-O	4.17	40.48	68.36	0.98
BM	8.96	55.50	74.71	1.17
本文模型	3.86	36.46	58.48	0.71

表2的数据显示,本文提出的模型在3个关键指标MAPE,MAE和1-R²上均优于对比模型,表明该模型在预测准确性和适用性方面具有一定优势。然而,在指标RMSE上,本文模型略逊于XGBoost算法,这种现象可能源于XGBoost在处理某些特定类型的数据结构时更高效。此外,尽管BM模型在RMSE和1-R²要低于KNN模型,但是MAPE和MAE却高于KNN模型。这一结果揭示了依赖单一性能指标可能导致评价局限性。为解决这一问题,本文应用熵权法综合考量各性能指标,计算得到各模型的综合得分,如图5所示。图5显示,本文模型的综合得分显著低于其他模型,进一步验证其在多个方面的性能优势。这种综合评估方法不仅提供了更科学的模型选择依据,也促进了评价方法的客观性和全面性的提升。

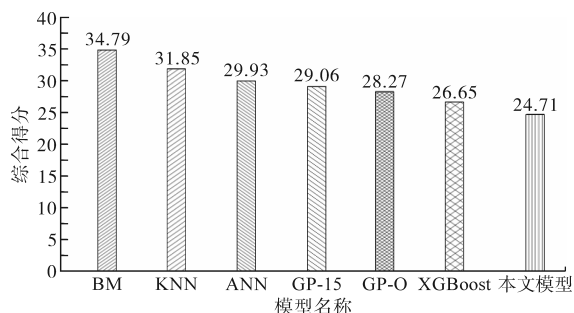


图5 不同模型综合得分对比结果

Fig.5 Comparison results of composite scores of different models

4.4 不同风电机组的功率预测残差分布

本文通过引入Penmanshiel风电场同型号的另外两台风电机组的数据,进一步验证模型的准确性,并绘制相应的功率预测残差分布图(图6和图7)。两台风电机组分别记为WT₁和WT₂。

风电机组的功率预测残差分布是评估预测方法准确性的关键指标。理想情况下,一个高精度的功率预测模型会显示一个尖峰的残差分布,表明大部分预测值与实际值之间的误差很小。相反,一个低精度模型的残差分布则较为分散^[24]。

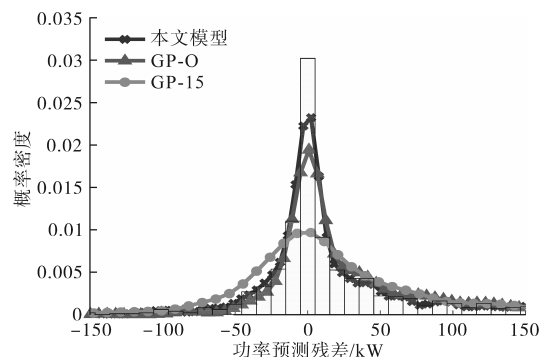


图6 WT₁的3种功率曲线模型的预测残差分布

Fig.6 Predicted residual distributions for three power curve models of WT₁

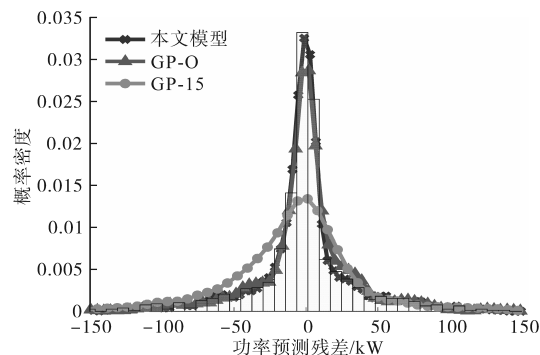


图7 WT₂的3种功率曲线模型的预测残差分布

Fig.7 Predicted residual distributions for three power curve models of WT₂

从图6和图7的观察结果可见,未经变量选择的高斯过程模型与包含15个变量的高斯过程模型在功率预测残差分布上表现出较大的分散性。相比之下,本研究提出的模型显示出更尖锐的功率预测残差分布,特别是在0附近。

5 结论

本文提出一种基于随机森林和改进高斯过程的风电机组功率曲线模型,并通过英国Penmanshiel风电场2.05 MW实验机组的运行数据进行实证分析,验证模型的有效性。相较于现有模型,本文原创性工作如下:

1)本文应用随机森林算法识别出影响功率的关键变量,有效提升了Penmanshiel风电场2.05 MW实验机组功率曲线模型的预测精度。

2)引入LDL分解以计算高斯过程中的逆矩阵,有效缩短了求逆矩阵的计算时间。

3)应用熵权法生成综合指标,有效解决了4类指标在预测精度方面的冲突。

4)对Penmanshiel风电场另外两台风电机组的预测残差分布进行分析,进一步证实了模型的

有效性。

参考文献

- [1] Global Wind Energy Council. Global wind report 2023[EB/OL]. (2023-04-13)[2024-05-13]. <https://gwec.net/>.
- [2] WEI D, WANG J, LI Z, et al. Wind power curve modeling with hybrid copula and grey wolf optimization[J]. IEEE Transactions on Sustainable Energy, 2021, 13(1):265-276.
- [3] IEC. Power performance measurements of electricity producing wind turbines: IEC 61400-12-1: 2005[S]. British Standard: IEC, 2005.
- [4] KUSIAK A, ZHENG H, SONG Z. On-line monitoring of power curves[J]. Renewable Energy, 2009, 34(6):1487-1493.
- [5] LYDIA M, SELVAKUMAR A I, KUMAR S S, et al. Advanced algorithms for wind turbine power curve modeling[J]. IEEE Transactions on Sustainable Energy, 2013, 4(3):827-835.
- [6] YOU M, LIU B, BYON E, et al. Direction-dependent power curve modeling for multiple interacting wind turbines[J]. IEEE Transactions on Power Systems, 2017, 33(2):1725-1733.
- [7] PANDIT R, INFIELD D, SANTOS M. Accounting for environmental conditions in data-driven wind turbine power models[J]. IEEE Transactions on Sustainable Energy, 2022, 14(1):168-177.
- [8] KARAMICHAILIDOU D, KALOUTSA V, ALEXANDRIDIS A. Wind turbine power curve modeling using radial basis function neural networks and tabu search[J]. Renewable Energy, 2021, 163:2137-2152.
- [9] GUO P, INFIELD D. Wind turbine power curve modeling and monitoring with Gaussian process and SPRT[J]. IEEE Transactions on Sustainable Energy, 2018, 11(1):107-115.
- [10] 姚登举, 杨静, 詹晓娟. 基于随机森林的特征选择算法[J]. 吉林大学学报(工学版), 2014, 44(1):137-141.
YAO Dengju, YANG Jing, ZHAN Xiaojuan. Feature selection algorithm based on random forest[J]. Journal of Jilin University (Engineering and Technology Edition), 2014, 44(1):137-141.
- [11] 杨晓梅, 郭林明, 肖先勇, 等. 基于可调品质因子小波变换和随机森林特征选择算法的电能质量复合扰动分类[J]. 电网技术, 2020, 44(8):3014-3020.
YANG Xiaomei, GUO Linming, XIAO Xianyong, et al. Classification of multiple power quality disturbances based on TQWT and random forest feature selection algorithm[J]. Power System Technology, 2020, 44(8):3014-3020.
- [12] 孙斌, 姚海涛, 刘婷. 基于高斯过程回归的短期风速预测[J]. 中国电机工程学报, 2012, 32(29):104-109, 5.
SUN Bin, YAO Haitao, LIU Ting. Short-term wind speed forecasting based on Gaussian process regression model[J]. Proceedings of the CSEE, 2012, 32(29):104-109, 5.
- [13] 赵洋, 王瀚墨, 康丽, 等. 基于时间卷积网络的短期电力负荷预测[J]. 电工技术学报, 2022, 37(5):1242-1251.
ZHAO Yang, WANG Hanmo, KANG Li, et al. Temporal convolution network-based short-term[J]. Transactions of China Electrotechnical Society, 2022, 37(5):1242-1251.
- [14] 张天瑞, 刘玉亭, 王译可. 基于改进BiLSTM的多工序产品质量预测研究[J]. 系统仿真学报, 2023, 35(11):2321-2332.
ZHANG Tianrui, LIU Yuting, WANG Yike. Research on multi-process product quality prediction based on improved BiLSTM[J]. Journal of System Simulation, 2023, 35(11):2321-2332.
- [15] 谢国民, 陈天香. 改进黑猩猩算法的光伏发电功率短期预测[J]. 电力系统及其自动化学报, 2024, 36(2):135-143.
XIE G M, CHEN T X. Improved short-term prediction of photovoltaic power generation power by chimpanzee algorithm[J]. Proceedings of The CSU-EPSA, 2024, 36(2):135-143.
- [16] 胡杰, 鲍帆. 基于组合赋权-TOPSIS模型的机场运行效率评价[J]. 系统仿真学报, 2023, 35(12):2570-2581.
HU Jie, BAO Fan. Airport operational efficiency evaluation based on combined weighting-TOPSIS model[J]. Journal of System Simulation, 2023, 35(12):2570-2581.
- [17] PLUMLEY C. Penmanshiel wind farm data[DB/OL]. Zenodo, [2023-05-13]. <https://zenodo.org/record/5946808#.YgpAmvso-V5>.
- [18] 吴永斌, 张建忠, 袁正舫, 等. 风电场功率异常数据识别与清洗研究综述[J]. 电网技术, 2023, 47(6):2367-2380.
WU Yongbin, ZHANG Jianzhong, YUAN Zhengxi, et al. Review on identification and cleaning of abnormal wind power data for wind farms[J]. Power System Technology, 2023, 47(6):2367-2380.
- [19] 刘琳, 郭鹏. 基于改进决策树的多变量功率曲线建模方法[J]. 动力工程学报, 2019, 39(8):647-653.
LIU Lin, GUO Peng. Wind turbine power curve modelling based on improved decision tree considering multiple input variables[J]. Journal of Chinese Society of Power Engineering, 2019, 39(8):647-653.
- [20] JANSSENS O, NOPPE N, DEVRIENDT C, et al. Data-driven multivariate power curve modeling of offshore wind turbines[J]. Engineering Applications of Artificial Intelligence, 2016, 55:331-338.
- [21] 杨茂, 代博祉. 基于比恩法的风电场风速-功率曲线建模误差分析[J]. 电力自动化设备, 2020, 40(12):81-89.
YANG Mao, DAI Bozhi. Modeling error analysis of wind speed-wind power curve for wind farm based on Bins method[J]. Electric Power Automation Equipment, 2020, 40(12):81-89.
- [22] PELLETIER F, MASSON C, TAHAN A. Wind turbine power curve modeling using artificial neural network[J]. Renewable Energy, 2016, 89(4):207-214.
- [23] TORRES-BARRANARRÁN A, ALONSO Á, DORRONSORO J R. Regression tree ensembles for wind energy and solar radiation prediction[J]. Neurocomputing, 2019, 326(31):151-160.
- [24] NASERY P, EZZAT A A. Yaw-adjusted wind power curve modeling: a local regression approach[J]. Renewable Energy, 2023, 202(1):1368-1376.

收稿日期:2024-05-13

修改稿日期:2024-07-03