

考虑机组健康状态的风电机组出力概率预测

杨伟新^{1,2}, 赵洪山², 杨 铎², 张扬帆¹, 林诗雨²

(1. 国网冀北电力有限公司科学研究院, 北京 100045; 2. 华北电力大学 电气与电子工程学院, 河北 保定 071003)

摘要: 风电机组健康状态会影响功率预测精度, 为了提高短期风功率预测的准确度, 促进可再生能源应用, 提出了一种考虑机组健康状态的风电机组出力概率预测方法。首先构建基于自适应最大均值差异算法 (AMMD) 的风电机组健康状态评估模型; 其次对风速等特征向量进行斯皮尔曼相关系数的重要性评估, 将影响权重较大的因素与风电机组健康状态评估结果作为影响因素输入到随机森林概率预测模型中预测风电出力概率; 最后采用某风电场 SCADA 数据进行预测区间准确度实例验证。实验结果表明, 使用的方法相比于传统随机森林方法, 平均绝对误差 (MAE) 降低了 3.521, R^2 提高了 12%; 相比于其他对比算法预测效果也均有不同程度提高。研究成果可为低健康度的风电功率预测工作提供参考。

关键词: 风电概率预测; 最大均值偏差; 决策树; 随机森林

中图分类号: TM614 文献标志码: A 文章编号: 1007-2691 (2026) 02-0099-11

Probabilistic Prediction of Wind Turbine Output Regarding Wind Turbine Health State

YANG Weixin^{1,2}, ZHAO Hongshan², YANG Duo², ZHANG Yangfan¹, LIN Shiyu²

(1. Scientific Research Institute of State Grid Jibei Electric Power Co., Ltd., Beijing 100045, China;

2. School of Electrical and Electronic Engineering, North China Electric Power University, Baoding 071003, China)

Abstract: The health state of the wind turbine affects the accuracy of power prediction. To improve the accuracy of short-term wind power prediction and promote the application of renewable energy, this paper proposes a probabilistic prediction method for wind turbine output that considers the health state of the wind turbine. Firstly, a health state assessment model for wind turbines is constructed based on the adaptive maximum mean difference algorithm (AMMD). Secondly, the importance of Spearman's correlation coefficient is assessed for the wind speed and other eigenvectors, and the factors with larger impact weights and the assessment results of the health state of the wind turbine are inputted as influencing factors into the Random Forest probabilistic prediction model to predict the probability of wind power output; and SCADA data of a wind farm is used for the prediction interval accuracy example. Finally, we use the SCADA data of a wind farm to verify the accuracy of the prediction interval. The experimental results show that, compared to the traditional random forest method, the proposed approach reduces the mean average absolute error (MAE) by 3.521 and improves R^2 by 12%. It also outperforms other comparison algorithms to varying degrees. The research results can provide a reference for wind power prediction in low health conditions.

Key words: probabilistic prediction of wind power; maximum mean deviation; decision tree; random forest

0 引言

风力发电正在成为清洁能源领域的重要实用工具,但目前风力发电存在大面积弃风现象;精确

的风电功率预测不仅可以解决上述问题;还能帮助电网调度部门安排电源调度方案,提高电网运行安全性,因此提高对风电出力预测精度对于促进风电发展,推进新型电力系统构建具有重要意义。

根据预测形式,可以将风电功率预测方法划

分为单点预测和概率预测^[1]。其中,单点预测直接给出预测时刻对应预测值,难以体现风电出力的不确定性和随机性,因此通常用于电力市场竞争方面。而概率预测给出预测时刻的风电功率概率分布,并且将概率分布的期望值作为预测结果,因此应用在风险评估及系统概率优化调度、风险分析和决策方面。

根据预测原理,风机功率预测可以划分为物理方法、统计学方法、数据驱动法。物理方法主要考虑风向、风速、气压以及风电场附近地势等物理信息^[2-5],根据风速功率曲线得到风电场功率,但建模和求解过程复杂,需要丰富精准的气象数据;统计学方法利用数据采集与监控系统(Supervisory Control And Data Acquisition, SCADA)和数值天气预报(Numerical Weather Prediction, NWP)数据寻求历史数据和机组输出功率的映射关系进行预测,包括时间序列法、概率预测法、灰色预测法、小波分析法^[6]等,与物理模型相比更容易建模^[7-10],但预测误差会随预测时间的延长而增加,且对突变情况反应不够灵敏。

数据驱动法通过对数据关系的学习和训练建立非线性模型,更好地适应风功率的时间序列特性,主要包含神经网络法、支持向量机法以及模糊预测法^[11-13]。文献[14]基于互信息理论对风机 SCADA 数据进行筛选,利用筛选特征实现风电功率预测;文献[15]基于极限学习机和粒子群优化算法,提出直接对风电功率进行概率预测的混合智能算法;文献[16]提出基于实数编码小世界优化算法的支持向量机改进模型,使参数快速收敛于全局最优值,取得了较好的风电功率预测效果;文献[17]提出一种改进型果蝇优化算法,将风速子序列重构参数和 LS-SVM 参数作为优化目标建立风速预测模型,通过风速-风功率转化关系可以求得风电场的功率值;文献[18]提出了一种以 CNN-BiLSTM 为基础模型,结合双重注意力机制和贝叶斯优化算法的风电功率超短期预测模型。通过自主挖掘输入特征与风电功率之间的数据关联,构建特征注意力模块。文献[19]提出了一种基于弹球损失优化的两步 WPPF 模型,优化预测分布以生成准确的概率预测。但是以上预测的方法并未考虑风电机组本身健康状态对预测精度的影响,当极端天气情况下风机健康状态较低时,风电功率的预测误差较大。

针对上述问题,本文提出了一种考虑机组健康状态的风机出力概率预测的方法,首先,建立了基于自适应最大均值差分算法的健康状态评估模型,以量化风电机组的运行健康状况;其次,将健康状态评估结果作为关键影响因素,融入到出力概率预测模型中,基于随机森林算法构建了考虑风电机组健康状态的出力概率预测模型;最后,通过利用我国北方某风电场的 SCADA 数据和进行实例验证,证明了所提方法的优越性。

1 基于自适应最大均差的风电机组健康度评估

由于风电机组长期处于正常运行状态,在实际的风电机组采集数据中,为了减少重复性的故障数据挖掘工作,避免对相关先验知识的过度依赖,本文利用风力发电机组的数据建立自适应最大均差算法模型。通过这种方式,从原始 SCADA 数据集中筛选出的高质量健康数据样本是后续研究的基础。在风电机组的实际运行中,由于风速和风向的不确定性以及变速恒频电气控制的约束,风电机组的运行状态通常在不同的运行状态之间随机且频繁地切换,这将产生一系列异常数据^[20]如图 1 所示。

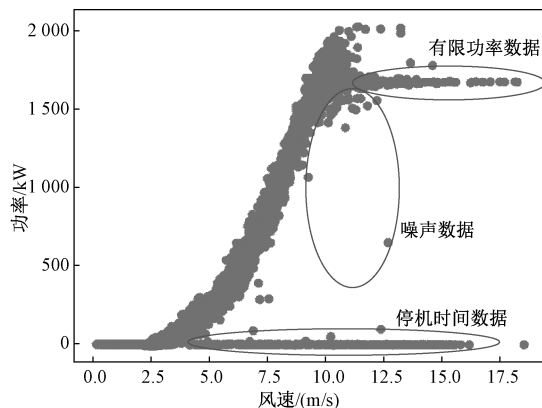


图 1 数据异常机组的“风速-功率”分布图

Fig. 1 “Wind speed-power” distribution diagram of unit with abnormal data

常见异常数据为:

- 1) 停机时间数据: 风电机组的实测风速大于切入风速,输出功率连续为 0,主要原因是人为弃风或通信故障。
- 2) 有限功率数据: 风电机组的输出功率分布在理想功率曲线以下,不随风速的变化而变化(或变化很小)。主要原因是人为控制风电机组限制

出力造成的。

3) 噪声(故障)数据:数据点随机分布在整体数据点之外,主要原因是风电机组故障或噪声干扰。

针对数据异常的情况,导致机组数据出现异常的原因有很多,但在实际运行中主要是由于风电机组中的各种组件发生故障引起的,比如发电机或变桨系统故障,就会导致机组停机或部分功率输出异常的情况。因此当数据出现异常时,为了判断机组的故障程度,我们使用自适应最大均值差异方法对此时风电机组的健康状态进行评估,根据此时机组健康度的值大小来判断此时传感器采集到的数据可靠性并分配该组数据在后续出力概率预测时所占权重,以此来纠正机组异常数据对概率预测造成的影响。由于风电机组的部件很多,每一部分发生故障时对整个机组的影响不尽相同,因需要将每种异常数据的特征与影响权重对应起来,以便出力预测算法进行更完善的学习,从而进一步提高预测的精度。

除此之外,风电机组在进行数据采集时由于各种环境因素或机组健康因素还会出现数据缺失的问题。针对这种情况,本文通过使用牛顿插值法,根据未知函数的已知点,利用已知点函数值及其差商信息构造一个简单函数来逼近原函数,从而完成数据的补齐。

为了分标签处理以上故障数据和正常数据,本文建立自适应最大均差算法模型。形式上,设 $\tilde{X} = \{\omega_1 x_1, \omega_2 x_2, \dots, \omega_m x_m\}^\tau$ 和 $Y = \{y_1, y_2, \dots, y_m\}^\tau$ 是具有分布 p 和 q 的可变集。变量 x_i 和 $y_j (i \in [1, m], j \in [1, n], m, n$ 是传感器的数量)是由不同传感器收集风速、风向等时间序列数据,而 p 和 q 分别表示风电机组处于标准状态和当前状态时的分布。对于每个传感器,时间窗口的长度可以表示为 1,因此, X 和 Y 中的每个元素可以通过 $x_i = (x_i^1, x_i^2, \dots, x_i^l)$ 和 $y_j = (y_j^1, y_j^2, \dots, y_j^l)$ 来详细描述。以这种方式,由最大均值差异(Maximum Mean Discrepancy, MMD)定义的 p 和 q 之间的距离的经验估计为^[21]:

$$MMD(p, q) = \|E_p[f(x)] - E_q[f(y)]\|_H \quad (1)$$

式中: H 为再生核希尔伯特空间(Reproducing Kernel Hilbert Space, RKHS),并且 $f(\cdot)$ 为将变量映射到高维空间的映射函数。通过 $f(\cdot)$ 函数得到风速等特征量到不同健康状态标签的内部映射关

系。然后, MMD 的有偏差的经验估计可以通过用为样本 X 和 Y 计算的经验期望代替总体期望来获得。

$$MMD(X, Y) = \left\| \frac{1}{m} \sum_{i=1}^m f(x_i) - \frac{1}{n} \sum_{j=1}^n f(y_j) \right\| \quad (2)$$

基于传统的 MMD,每个传感器 i 的数据乘以相应的系数 ω_i ,以区分不同传感器对于故障预测的重要性。具体地, MMD 的输入矩阵应重写为 $\tilde{X} = \{\omega_1 x_1, \omega_2 x_2, \dots, \omega_m x_m\}^\tau$ 和 $\tilde{Y} = \{\omega_1 y_1, \omega_2 y_2, \dots, \omega_n y_n\}^\tau$ ^[22]。通过对故障预测结果的召回率进行周期性自动调整,可以为不同的传感器分配不同程度的权重,将其作为传感器权重的系数。自适应最大均值差异(Adaptive Maximum Mean Discrepancy, AMMD)模型可以通过最大化预测的关键故障数量来训练最佳传感器权重。AMMD 的数学表示为

$$AMMD(\tilde{X}, \tilde{Y}) = \left\| \frac{1}{m} \sum_{i=1}^m f(\omega_i x_i) - \frac{1}{n} \sum_{j=1}^n f(\omega_j y_j) \right\| \quad (3)$$

为了计算上述公式,设 H 为函数 $f(\cdot)$ 的完备内积空间: $X \rightarrow H$, 其中 X 为非空紧集。如果对所有 $x \in X$, 线性点评价函数映射 $f \rightarrow f(x)$ 存在且连续。在这种情况下, $f(x)$ 可以通过以下方式表示为内积^[23]:

$$f(x) = \langle f, \phi(x) \rangle_H \quad (4)$$

式中: $X \rightarrow H$ 被称为从 X 到 H 的特征空间映射。本文中使用的无限维 H 对应于高斯核:

$$k(x, x') = e^{\left(-\frac{\|x-x'\|^2}{2\sigma^2}\right)}$$

$$\phi(x) = e^{-\frac{x^2}{2\sigma^2}} \left(1, \sqrt{\frac{1}{1!}} \frac{x}{\sigma}, \sqrt{\frac{1}{2!}} \frac{x^2}{\sigma^2}, \dots, \sqrt{\frac{1}{n!}} \frac{x^n}{\sigma^n}\right) \quad (5)$$

用 $\mu_p = E_p[\phi(x)]$ 表示特征空间中 (x) 的期望^[24], AMMD 可以重写为

$$AMMD(\tilde{X}, \tilde{Y}) = \|\mu_p - \mu_q\|_H \quad (6)$$

将 AMMD 值在 0 和 1 之间归一化^[25], 则风力发电机健康度可量化为

$$HI = \frac{|x - \mu(x)|}{\mu(x)} \quad (7)$$

式中: $\mu(x)$ 为特征数据的均值, x 为特征数据实际

值。经过多次迭代训练后,AMMD 模型已获得各特征量的稳定权重,并可用于对每组数据所对应的机组健康度进行评估。根据预先设定的标签,该评估结果的归一化值越接近 0 则表明该机组越为健康,越接近 1 则表明该机组存在更高的故障风险。在模型训练过程中,本文采取了针对不同健康度数据分配不同权重的方法,以确保健康数据的特征对模型的影响程度更大。

2 考虑风电机组健康状态的功率概率预测

2.1 决策树原理

决策树是一种二分递归分割技术,它将一个样本集中在除叶节点之外的节点上分割为两个子集。决策树算法以 Gini 指数作为度量,Gini 指数的概率分布表达式为^[26]

$$Gini(p) = \sum_{k=1}^K p_k(1-p_k) = 1 - \sum_{k=1}^K p_k^2 \quad (8)$$

式中: K 为节点中特征样本的种类总数, p_k 为节点中第 K 类特征样本出现的概率。

样本集 D 的 Gini 指数定义为

$$Gini(D) = 1 - \sum_{k=1}^K \left(\frac{|C_k|}{|D|} \right)^2 \quad (9)$$

式中: C_k 为样本集 D 中属于第 k 类的子样本集。

各分区的指数定义为

$$Gini_{split}(D) = \frac{|D_1|}{|D|} Gini(D_1) + \frac{|D_2|}{|D|} Gini(D_2) \quad (10)$$

式中: D_1 、 D_2 为样本集 D 的某个属性点的二值划分数据集。

2.2 随机森林原理

随机森林算法由许多基分类器决策树 $\{h(X, \theta_k), k=1, 2, \dots, n_{tree}\}$ 通过一定规则聚合而成。设原始数据集为 $D\{X, Y\}$, 其中 X 为风速等原始数据的特征属性, Y 为 X 对应的类别属性也即风机出力。其原理如图 2 所示^[27]。

随机森林通过 Bootstrap 重采样生成随机训练集 δ_i 和相应的决策树, 单个决策树的预测值 Y_i 为

$$Y_i(\delta_i) = \sum_{i=1}^N \lambda_i y_i \quad (11)$$

式中: N 为单个决策树的叶节点数量, λ_i 为叶节点 i 处的预测值权重, y_i 为叶节点 i 处的预测值。

对所有决策树的预测结果进行求均值处理, 得到随机森林的最终预测结果 Y 为

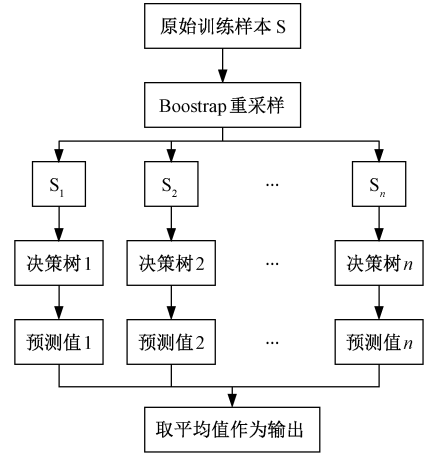


图 2 随机森林原理图

Fig. 2 Random forest schematic

$$Y = \frac{1}{N_{tree}} \sum_{i=1}^{N_{tree}} Y_i(\delta_i) \quad (12)$$

式中: N_{tree} 为决策树数量。

2.3 考虑健康状态的随机森林模型建立

考虑健康状态的随机森林模型建立基本思路如下, 首先, 从风电场出力相关数据中选取样本集 D 。将风速、风向等输入向量作为影响出力的特征向量;

其次, 筛选特征变量重要性。将通过 AMMD 算法得到的健康状态评估结果与原有的特征矩阵进行合并, 得到一个新的特征矩阵。由于除了健康状态以外的特征向量对机组出力的影响权重并不相同, 需要筛选出权重较大的变量, 作为预测出力的主要影响因素。为了判断哪些特征变量对于机组出力的贡献更大, 本文使用斯皮尔曼相关系数进行分析。相关系数越大, 表明该特征变量 M 的重要性越高^[28]。

再次, 生成随机向量序列 $\theta_i (i=1, 2, \dots, k)$, 利用 Bootstrap 重抽样方法从原始数据集 D 中抽取 k 个 Bootstrap 样本集 分别建立风电场出力 CART 回归决策树模型 $\{T(X, \theta), i=1, 2, \dots, k\}$ ^[29], 其中 X 是从风电场出力相关数据中筛选出来的用于预测风电场出力的自变量, 参数集 $\{\theta_k\}$ 是独立同分布的随机向量。

最后, 通过 k 次训练, 得到分类预测模型序列 $\{T_1(X), T_2(X), \dots, T_k(X)\}$, 再用他们构成一个分类回归预测模型系统^[30], 该系统就是随机森林预测模型。

假设对所用数据进行有放回的随机抽样建立

k 棵分类回归树 $\{T_1(X), T_2(X), \dots, T_k(X)\}$ 从而建立随机森林,其中 $X = \{x_1, x_2, \dots, x_m\}$ 为 m 维的输入向量。而每棵树对应一个输出结果 Y_i , 则共有 k 个输出值,其分别是 $Y_1 = T_1(X), Y_2 = T_2(X), \dots, Y_k = T_k(X)$ [31,32]。其中输入向量包括风速、风向等影响风电场出力的特征变量,输出向量为风电场出力值。则流程图如图 3 所示。

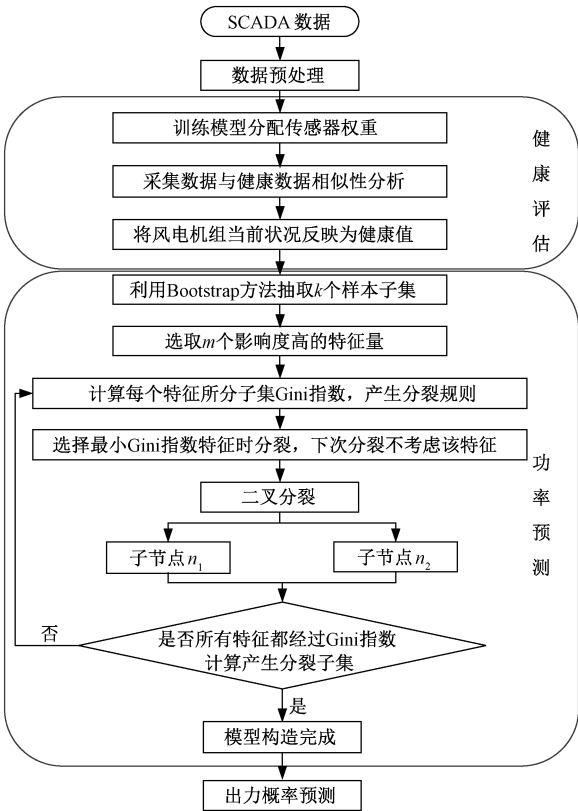


图 3 考虑风电机组健康状态的风电机组出力概率预测流程图

Fig.3 Flowchart of WTG output probability prediction considering WTG health state

3 算例分析

3.1 数据来源

为验证所提出的功率预测模型的有效性,本文选取某风电场在 2016-2017 年期间的 SCADA 数据进行验证。数据采样间隔为 10 min,预测步长为 30。从数据点 350 开始,我们选择了因 FPM 通信故障、操作超时制动、风扇叶片存在故障等因素导致机组健康度较低的运行数据,以与健康数据进行对比分析。这些数据点被用于评估所提方法的预测性能。

3.2 机组健康状态评估结果

本文采用 AMMD 算法评估机组健康度的量化

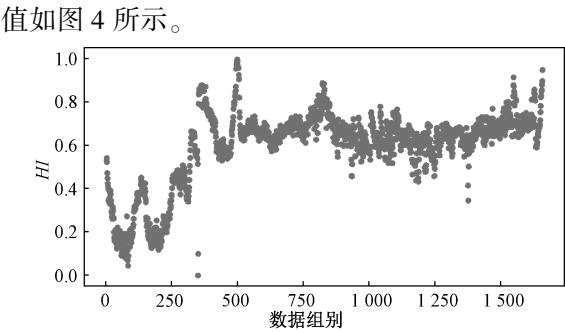


图 4 AMMD 量化机组健康度
Fig.4 AMMD recognizes unit health states

在 AMMD 模型的训练过程中,参数矩阵中各个特征量的权重将通过自动匹配的方式获得最适合的数值,以确保对机组健康状态的准确识别。一旦训练完成,该模型将利用采集到的标准数据与风电机组处于健康状态时的数据进行比较,即比较它们之间的数据分布相似性,从而将风电机组的当前状态转化为健康度值。这样的健康度值能够反映机组当前的状况。

由图 4 可知,应用自适应最大均差算法对机组的健康度进行量化评估具备清晰直观的特点,能够有效判断机组的健康状态。这种评估方法有助于筛选出所需的机组数据,并提供了对机组健康状态进行准确评估的手段。

3.3 预测特征筛选

为筛选影响风电机组出力的特征,本文对 18 维风电 SCADA 数据进行斯皮尔曼相关性分析,各特征变量之间的斯皮尔曼系数分布如图 5 所示。

由图 5 可知,风速、桨距角、变压器温度、转子轴承温度与风电出力的相关系数分别为 0.84、0.84、0.80、0.83。表明以上因素与风电出力具有强相关特性,而其余因素与机组出力的斯皮尔曼相关系数均不超过 0.55,相关性不明显。因此本文选取风速、桨距角、变压器温度以及转子轴承温度作为主要特征量构建概率预测模型。

3.4 风电出力概率预测结果

本文对比了传统随机森林算法与考虑机组健康度的随机森林算法分别在 80%置信度和 90%置信度下风电出力概率预测情况,结果如图 6~7 所示。

图 6 为未经过健康状态评估,只使用随机森林算法进行风电机组出力预测的曲线,可以看到尽管功率曲线也包含在预测区间内,但预测区间范围较大,不能很好的为机组调度运行提供可靠

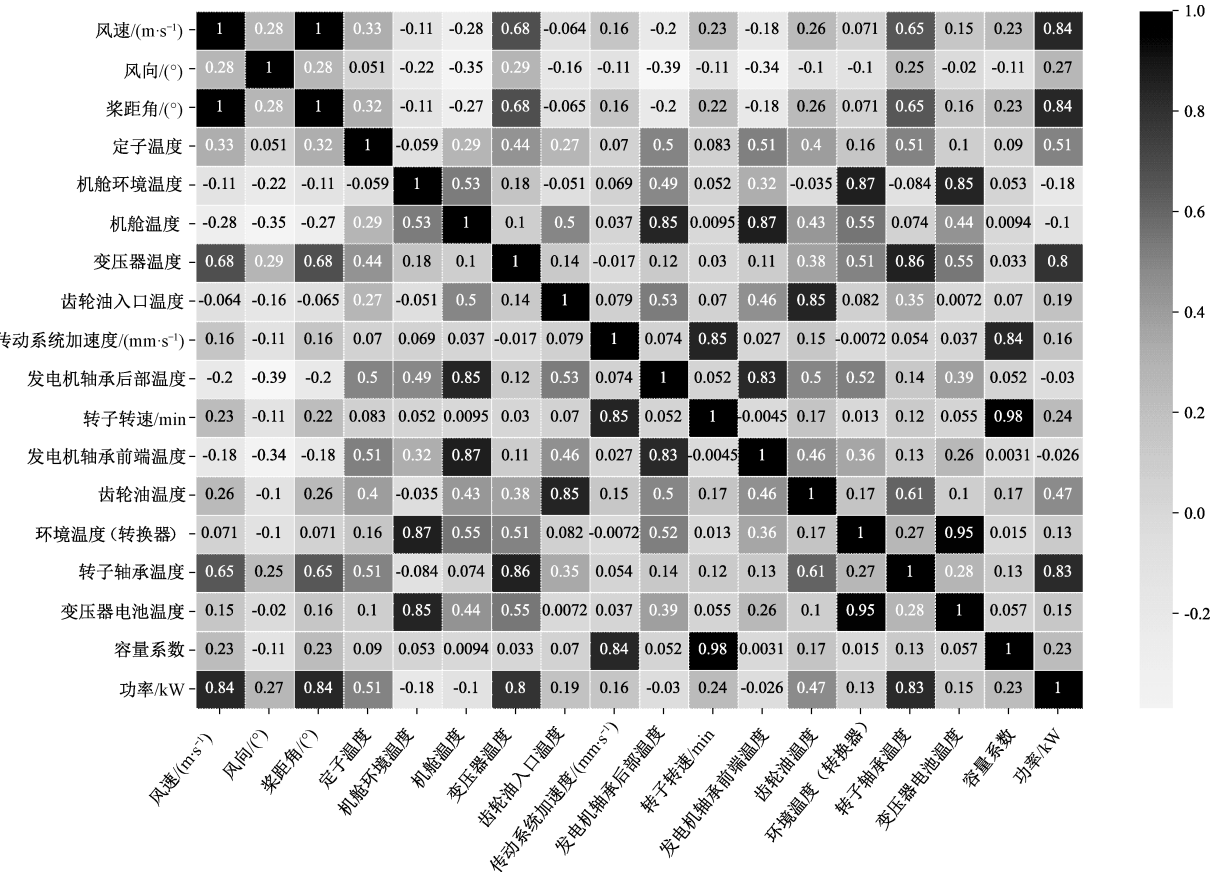


图5 斯皮尔曼相关系数热点图

Fig.5 Spielman related coefficient hotspots

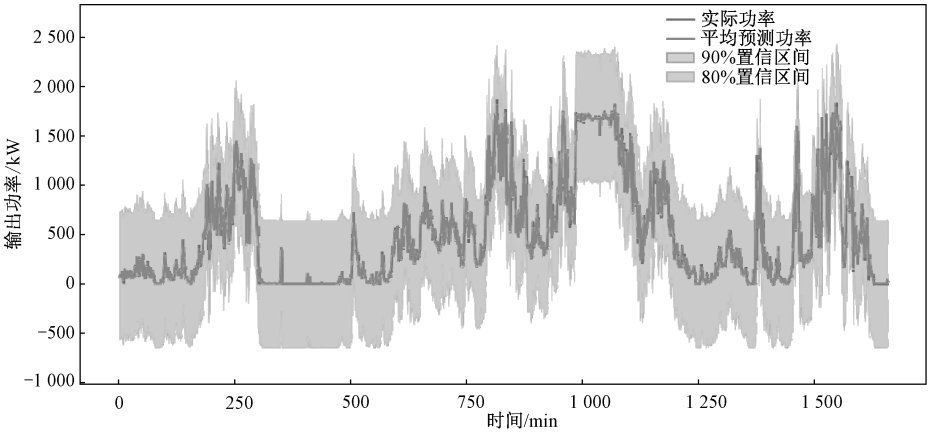


图6 传统随机森林算法的90%与80%置信度下机组出力概率预测结果

Fig.6 Probability prediction results of unit output under 90% and 80% confidence of traditional random forest algorithm

的输出功率。这是因为只使用随机森林算法的功率预测方法没有考虑机组健康状态出现问题时产生的异常数据,对这些错误的数 据按正常情况下进行分析,导致得到的结果误差很大。

由图 7 可见,考虑机组健康度的随机森林算法在 80%置信度下机组出力概率预测区间可以包

含大部分机组实际的出力曲线,然而在某些情况下,例如受到外界影响导致运行状态突变时,出力曲线将在短时间内发生较大转折,此时 80%置信区间不能完全包含这部分曲线;但 90%置信度下预测区间可以完整包括机组的实际出力曲线。这是因为本文方法通过 AMMD 评估了风电机组健康

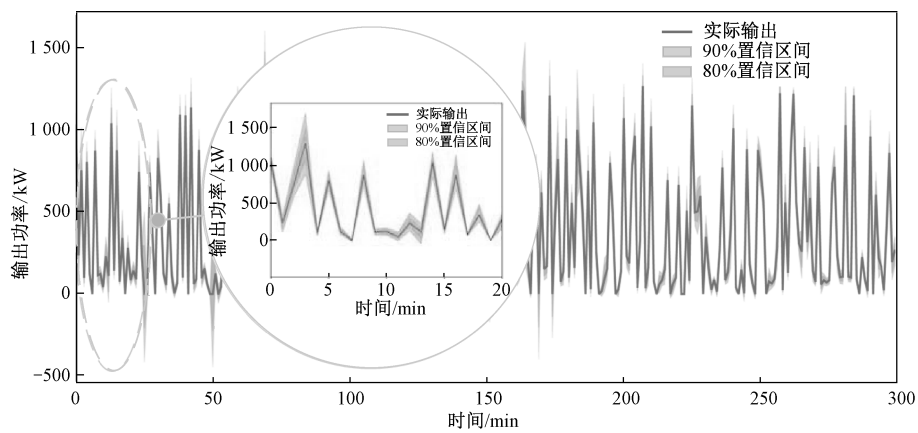


图 7 本文方法 90%与 80%置信度下机组出力概率预测

Fig. 7 Probability prediction results of unit output with 90% and 80% confidence obtained by this method

度,对低健康状态时的机组特征值降低权重,从而保证预测区间的精确性和可靠性。

为了验证考虑机组健康度的必要性,本文使用平均绝对误差(MAE)以及决定系数 R^2 来评估预测结果。MAE以及 R^2 的计算公式如下^[33]:

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - P_i|$$

(13)

式中: Y_i 和 P_i 分别为风电功率的实际值和预测

值, n 为预测样本总数。

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - P_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

(14)

式中: $\bar{Y}$ 为测试数据的平均值, R^2 越接近 1,表示预测效果越好。考虑机组健康度的随机森林算法与传统随机森林算法对比如图 8 所示。

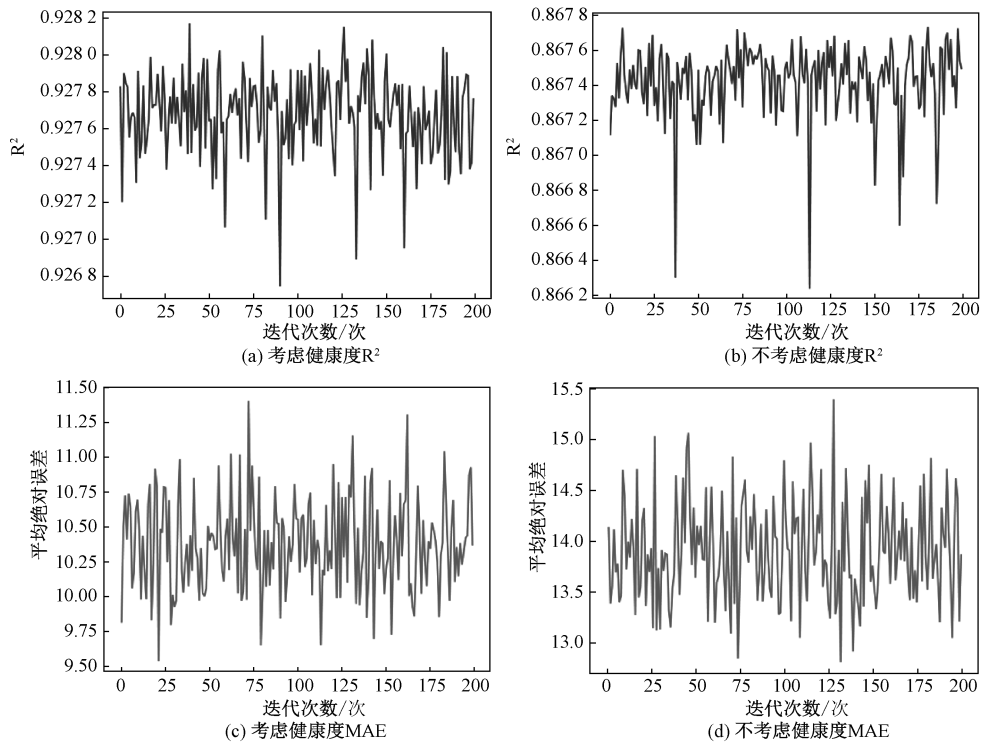


图 8 考虑机组健康度前后模型预测误差

Fig. 8 Model prediction error before and after considering unit health

由图 8 可见,考虑机组健康度的随机森林算法的 MAE 均小于单一随机森林算法, R^2 更接近

1,验证了考虑机组健康状态的必要性。

3.5 模型对比

为进一步验证预测方法的准确性,本文还与梯度提升树 (eXtreme Gradient Boosting, XGBoost)、支持向量机 (Support Vector Machine, SVM)、长短期记忆神经网络 (Long Short-Term Memory, LSTM) 这些常用的风电出力预测方法,以及他们分别与 AMMD 相结合的方法进行对比分析,对比情况如表 1 所示。

由表 1 可知,本文方法 MAE 和 R^2 均优于对比方法,进一步验证了本文方法的有效性。

3.6 泛化性能分析

为了验证本文所提方法的泛化性能,本文选取了同一风电场在 2023 年 3 月-6 月的 SCADA 数据进行算法验证,得到的风电出力概率预测结果如图 9 所示。

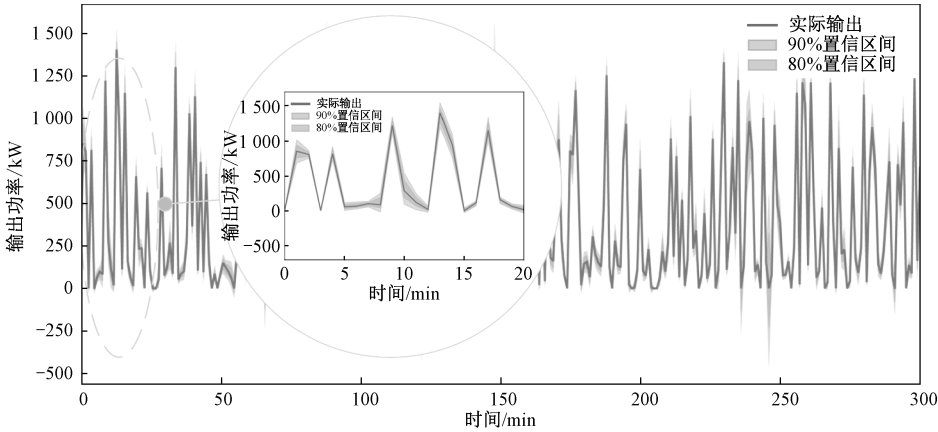


图 9 本文方法 90%与 80%置信度下机组出力概率预测泛化性能对比图

Fig. 9 Comparison of generalization performance of this paper's method for predicting probability of unit output at 90% and 80% confidence levels

4 结 论

为解决实际风机传感器采集数据有效值少且分布稀疏,以及未考虑风电机组健康状态对数据影响的缺点,本文提出了考虑机组健康状态的风电机组出力概率预测算法,通过自适应最大均差方法评估机组健康度以及随机森林方法预测机组出力算例分析,得出了以下结论:

- 1)本文提出了一种基于自适应均差随机森林的风电机组概率预测模型,通过考虑风电机组健康度,提高了风电机组概率预测精度;
- 2)所提考虑机组健康度的概率预测模型相比于传统为考虑机组健康度的风电出力概率预测模

表 1 不同出力预测方法误差对比

Tab. 1 Error comparison of different output prediction methods

使用方法	MAE	R^2
本文方法	10.389	0.987 7
AMMD-XGBoost	13.865	0.912 2
AMMD-SVM	21.396	0.886 3
AMMD-LSTM	16.128	0.905 3
XGBoost	15.910	0.867 7
SVM	25.729	0.878 2
LSTM	17.527	0.864 2

由图 9 可见,采用该风电场的近期数据进行风电出力的概率预测,所得到的预测区间依然可以完整包含实际风电出力曲线,并且区间范围仍然保持在较小范围内,为机组调度运行提供可靠的输出功率的同时又保证了预测精度,符合上一节中得到的结论,验证了本文方法的泛化性能。

型,MAE 降低了 3.521, R^2 提高了 12%,验证了考虑机组健康度的必要性;

3)经过健康状态评估后的预测区间相较于未经过健康状态评估更加精确,可以为机组调度运行提供可靠的输出功率。

4)本文方法还存在对噪声数据处理不够仔细问题,下一步的工作是研究如何进一步减小受数据噪声影响,进一步提高预测精度。

参考文献:

[1] 李姣,郭鹏. 基于 ARGAN 表面阴影预处理与迁移学习风电机组叶片故障识别 [J]. 华北电力大学学报 (自然科学版), 2021, 48 (2): 73-79.
LI Jiao, GUO Peng. ARGAN-based surface shadow

- preprocessing and transfer learning fault identification of wind turbine blade [J]. Journal of North China Electric Power University (Natural Science Edition), 2021, 48 (2): 73-79 (in Chinese).
- [2] 史光宇, 徐健, 杨强. 基于卷积神经网络的风电机组轴承机械故障智能诊断方法 [J]. 华北电力大学学报 (自然科学版), 2020, 47 (4): 71-79.
- SHI Guangyu, XU Jian, YANG Qiang. Intelligent fault diagnosis on wind turbine bearing based on convolutional neural network [J]. Journal of North China Electric Power University (Natural Science Edition), 2020, 47 (4): 71-79 (in Chinese).
- [3] 吴城, 米增强, 杨玉新, 等. 大型风机海上风电场集电系统拓扑优化 [J]. 华北电力大学学报 (自然科学版), 2023, 50 (6): 31-39.
- WU Jian, MI Zengqiang, YANG Yuxin, et al. Topology Optimization of offshore wind farm power collection system with large wind turbine. [J]. Journal of North China Electric Power University (Natural Science Edition), 2023, 50 (6): 31-39 (in Chinese).
- [4] 王丽婕, 冬雷, 高爽. 基于多位置 NWP 与主成分分析的风电功率短期预测 [J]. 电工技术学报, 2015, 30 (5): 79-84.
- WANG Lijie, DONG Lei, GAO Shuang. Wind power short-term prediction based on principal component analysis of NWP of multiple locations [J]. Transactions of China Electrotechnical Society, 2015, 30 (5): 79-84 (in Chinese).
- [5] 李莉, 刘永前, 杨勇平, 等. 基于 CFD 流场预计算的短期风速预测方法 [J]. 中国电机工程学报, 2013, 33 (7): 27-32+22.
- LI Li, LIU Yongqian, YANG Yongping, et al. Short-term wind speed forecasting based on CFD pre-calculated flow fields [J]. Proceedings of the CSEE, 2013, 33 (7): 27-32+22 (in Chinese).
- [6] LIU H, DUAN Z, LI Y, et al. A novel ensemble model of different mother wavelets for wind speed multi-step forecasting [J]. Applied Energy, 2018, 228: 1783-1800.
- [7] 闫家铭, 毕天姝, 胥国毅, 等. 海上风电经 VSC-HVDC 并网改进频率控制策略 [J]. 华北电力大学学报 (自然科学版), 2021, 48 (2): 11-19.
- YAN Jiaming, BI Tianshu, XU Guoyi, et al. An improved frequency control strategy for offshore wind farm connected by VSC-HVDC [J]. Journal of North China Electric Power University (Natural Science Edition), 2021, 48 (2): 11-19 (in Chinese).
- [8] 王贺, 胡志坚, 仇梦林. 基于模糊信息粒化和最小二乘支持向量机的风电功率波动范围组合预测模型 [J]. 电工技术学报, 2014, 29 (12): 218-224.
- WANG He, HU Zhijian, JI Menglin. A combined forecasting model for range of wind power fluctuation based on fuzzy information granulation and least squares support vector machine [J]. Transactions of China Electrotechnical Society, 2014, 29 (12): 218-224 (in Chinese).
- [9] KUSIAK A, LI W. Short-term prediction of wind power with a clustering approach [J]. Renewable Energy, 2010, 35 (10): 2362-2369.
- [10] 张国珍, 王其乐, 叶天泽, 等. 基于随机森林和 LSTM-自编码算法的风机高温降容状态智能检测方法 [J]. 华北电力大学学报 (自然科学版), 2021, 48 (3): 81-88.
- ZHANG Guozhen, WANG Qile, YE Tianze, et al. Intelligent detection method of high-temperature capacity reduction of wind turbine based on random forest and LSTM-Autoencoder [J]. Journal of North China Electric Power University (Natural Science Edition), 2021, 48 (3): 81-88 (in Chinese).
- [11] AZAD H, MEKHILEF S, GANAPATHY V. Long-term wind speed forecasting and general pattern recognition using neural networks [J]. IEEE Transactions on Sustainable Energy, 2014, 5 (2): 546-553.
- [12] LEE D, BALDICK R. Short-term wind power ensemble prediction based on gaussian processes and neural networks [J]. IEEE Transactions on Smart Grid, 2014, 5 (1): 501-510.
- [13] 孔波利, 崔丽艳, 丁钊, 等. 基于风光混合模型的短期功率预测方法研究 [J]. 电力系统保护与控制, 2015, 43 (18): 62-66.
- KONG Boli, CUI Liyan, DING Zhao, et al. Short-term power prediction based on hybrid wind-PV forecasting model [J]. Power System Protection and Control, 2015, 43 (18): 62-66 (in Chinese).
- [14] DU M, YI J, MAZIDI P, et al. A Parameter selection method for wind turbine health management through SCADA data [J]. Energies, 2017, 10 (2): 253.
- [15] WAN C, XU Z, PINSON P, et al. Probabilistic forecasting of wind power generation using extreme learning machine [J]. IEEE Transactions on Power Systems, 2014, 29 (3): 1033-1044.
- [16] 王爽心, 李涛, 孙东旭, 等. 基于小世界优化的支持向量机风电功率预测 [J]. 太阳能学报, 2015, 36 (3): 720-726.
- WANG Shuangxin, LI Tao, SUN Dongxu, et al. Wind power forecasting based on small-world optimization and

- support vector machine [J]. *Acta Energiæ Solaris Sinica*, 2015, 36 (3): 720-726 (in Chinese).
- [17] 程启明, 陈路, 程尹曼, 等. 基于 EEMD 和 LS-SVM 模型的风电功率短期预测方法 [J]. *电力自动化设备*, 2018, 38 (5): 27-35.
- CHENG Qiming, CHEN Lu, CHEN Yiman, et al. Short-term wind power forecasting method based on EEMD and LS-SVM model [J]. *Electric Power Automation Equipment*, 2018, 38 (5): 27-35 (in Chinese).
- [18] JIA X, JIN C, BUZZA M, et al. A deviation based assessment methodology for multiple machine health patterns classification and fault detection [J]. *Mechanical Systems & Signal Processing*, 2017, 99: 244-261.
- [19] 郭鹏, 姜漫利, 李航涛. 基于运行数据和高斯过程回归的风电机组发电性能分析与监测 [J]. *电力自动化设备*, 2016, 36 (8): 10-15+25.
- GUO Peng, JIANG Manli, LI Hangtao. Performance analysis and monitoring based on SCADA data and gaussian process regression for wind turbine power generation [J]. *Electric Power Automation Equipment*, 2016, 36 (8): 10-15+25 (in Chinese).
- [20] WANG X, LI P, YANG J. Short-term wind power forecasting based on two-stage attention mechanism [J]. *IET Renewable Power Generation*, 2020, 14 (2): 297-304.
- [21] SUN M, FENG C, CHARTAN E, et al. A two-step short-term probabilistic wind forecasting methodology based on predictive distribution optimization [J]. *Applied Energy*, 2019, 238: 1497-1505.
- [22] EL-FOULY T, EL-SAADANY E, SALAMA M. Grey predictor for wind energy conversion systems output power prediction [J]. *IEEE Transactions on Power Systems*, 2006, 21 (3): 1450-1452.
- [23] YAN K, WANG X, KIM J, et al. A new aggregation of DNN sparse and dense labeling for saliency detection [J]. *IEEE Transactions on Cybernetics*, 2021, 51 (12): 5907-5920.
- [24] 王焱, 汪震, 黄民翔, 等. 基于 OS-ELM 和 Bootstrap 方法的超短期风电功率预测 [J]. *电力系统自动化*, 2014, 38 (6): 14-19+122.
- WANG Yan, WANG Zhen, HUANG Minxiang, et al. Ultra-short-term wind power prediction based on OS-ELM and Bootstrap method [J]. *Automation of Electric Power Systems*, 2014, 38 (6): 14-19+122 (in Chinese).
- [25] 陈道君, 龚庆武, 金朝意, 等. 基于自适应扰动量子粒子群算法参数优化的支持向量回归机短期风电功率预测 [J]. *电网技术*, 2013, 37 (4): 974-980.
- CHEN Daojun, GONG Qingwu, JIN Chaoyi, et al. Short-term wind power prediction based on support vector regression machine optimized by adaptive disturbance quantum-behaved particle swarm optimization [J]. *Power System Technology*, 2013, 37 (4): 974-980 (in Chinese).
- [26] 刘蕾. 基于谱分析的大规模风电场功率超短期概率区间预测研究 [D]. 吉林: 东北电力大学, 2019.
- LIU Lei. Study on ultra-short-term probability interval prediction of wind power based on spectral analysis [D]. Jilin: Northeast Electric Power University, 2019 (in Chinese).
- [27] SIMAO H, POWELL W, ARCHER C, et al. The challenge of integrating offshore wind power in the US electric grid. Part II: Simulation of electricity market operations [J]. *Renewable Energy*, 2017, 103: 418-431.
- [28] 于群, 霍筱东, 何剑, 等. 基于斯皮尔曼相关系数和系统惯量的中国电网停电事故趋势预测 [J]. *中国电机工程学报*, 2023, 43 (14): 5372-5381.
- YU Qun, HUO Xiaodong, HE Jian, et al. Trend prediction of power blackout accidents in Chinese power grid based on Spearman's correlation coefficient and system inertia [J]. *Proceedings of the CSEE*, 2023, 43 (14): 5372-5381 (in Chinese).
- [29] 张鹏飞. 基于混合神经网络技术的风功率预测研究 [D]. 沈阳: 沈阳工业大学, 2015.
- ZHANG Pengfei. Wind power prediction method based on hybrid neural network [D]. Shenyang: Shenyang Institute of Technology, 2015 (in Chinese).
- [30] 张仕良. 基于深度神经网络的语音识别模型研究 [D]. 合肥: 中国科学技术大学, 2017.
- ZHANG Shiliang. Research on deep neural networks based models for speech recognition [D]. Hefei: China University of Science and Technology, 2017 (in Chinese).
- [31] 董雷, 周文萍, 张沛, 等. 基于动态贝叶斯网络的光伏发电短期概率预测 [J]. *中国电机工程学报*, 2013, 33 (S1): 38-45.
- DONG Lei, ZHOU Wenping, ZHANG Pei, et al. Short-term photovoltaic output forecast based on dynamic Bayesian network theory [J]. *Proceedings of the CSEE*, 2013, 33 (S1): 38-45 (in Chinese).
- [32] 杨静, 张楠男, 李建, 等. 决策树算法的研究与应用 [J]. *计算机技术与发展*, 2010, 20 (2): 114-116+120.
- YANG Jing, ZHANG Nannan, LI Jian, et al. Research and application of decision tree algorithm [J]. *Com-*

puter Technology and Development, 2010, 20 (2): 114-116+120 (in Chinese).

[33] MENDEZ G, LOHR S. Estimating residual variance in random forest regression [J]. Computational Statistics & Data Analysis, 2011, 55 (11): 2937-2950.

(上接第 98 页)

LI Guanghui, WANG Weisheng, LIU Chun, et al. Mechanism analysis and suppression method of wide-band oscillation of PMSG wind farms connected to weak grid (part I): analysis of wideband impedance characteristics and oscillation mechanism [J]. Proceedings of the CSEE, 2019, 39 (22): 6547-6561 (in Chinese).

[29] 李景一, 毕天姝, 于钊, 等. 直驱风机变流控制系统对次同步频率分量的响应机理研究 [J]. 电网技术, 2017, 41 (6): 1734-1740.

LI Jingyi, BI Tianshu, YU Zhao, et al. Study on response characteristics of grid converter control system of permanent magnet synchronous generators (PMSG) to subsynchronous frequency component [J]. Power System Technology, 2017, 41 (6): 1734-1740 (in Chinese).

作者简介:

杨伟新 (1987), 男, 博士研究生, 研究方向为风电机组功率预测, 健康状态评估, E-mail: 18510511136@163.com;

林诗雨 (1999), 男, 博士研究生, 研究方向为风电机组故障诊断技术, E-mail: lin17778884717@163.com。

作者简介:

高本锋 (1981), 男, 副教授, 研究方向为高压直流输电和电力系统次同步振荡, E-mail: gaobenfeng@126.com;

郑展翔 (2000), 男, 硕士研究生, 通信作者, 研究方向为电力系统分析与控制, E-mail: zhanxiang_zheng@126.com;

邓晓洋 (1989), 男, 研究方向为新能源并网技术, E-mail: 14117389@bjtu.edu.cn;

王潇 (1987), 男, 高级工程师, 研究方向为新能源并网技术, E-mail: wangxiao2017@foxmail.com;

吴林林 (1986), 男, 高级工程师, 研究方向为新能源并网技术, E-mail: wulin226@163.com;

刘培鑫 (2000), 男, 硕士研究生, 研究方向为电力系统次同步振荡, E-mail: liupeixin2022@126.com;

陈宗良 (2000), 男, 硕士研究生, 研究方向为电力系统分析与控制, E-mail: 3196599550@qq.com。