

风电机组异常风速的识别和修正方法研究

郝玉珠 石岚* 贾晓红

(内蒙古自治区气象服务中心, 呼和浩特 010051)

摘要 有效的数据清洗手段可提高风电机组测风资料的质量, 而风机数据质量对风资源评估、风功率发电有重要意义。本文提出了风功率区间识别风速、功率的异常值, 再基于风机高相关片区修正风速的方法。资料选用 2020—2022 年内蒙古乌兰察布市北部某风电场的风机测风数据进行分析。结果表明: 风功率区间修正后的风机数据完整率提高至 90% 以上, 利用风机高相关片区, 修正了部分异常风速。该方法提高了风电场风电机组测风风速的数据质量, 实现了风速和功率互相校准, 为风电场发电量预测、调控提供基础性支撑数据。

关键词 风电机组资料; 识别修正; 风机片区; 风功率区间

中图分类号: P49 **DOI:** 10.19517/j.1671-6345.20230264 **文献标识码:** A

引言

风能作为目前最具开发潜力的可再生能源, 风力发电是有效利用风能的主要途径。大规模风电并网对电力及相关部门平稳运营提出了更高的要求, 风功率预测的准确性变得至关重要。可靠完整的风机测风资料获取困难^[1], 需通过技术手段清洗风机运行数据, 以获得有效、完整的风速、功率等数据^[2-3]。

风电场都装有 SCADA (Supervisory Control and Data Acquisition) 系统实时监控风机状态和采集风机运行数据, 为及时、准确地了解风机运行状态提供了可靠的数据来源。由于人为干涉、通信故障等原因, 采集到的数据会出现不合理及丢失的情况, 停机检修、异常天气等情况引发测量数据与实际相悖。异常数据会给后续处理和分析带来较大误差, 通过数据清洗降低误差带来的影响^[4-5]。常见的数据清洗方法有 3 种^[6-8]: 根据风机自身参数等指标判别并剔除异常值; 根据附近风机、测风塔的观测数据进行研判; 风功率曲线识别判断^[9-10]。若对风电场风机数据进行精细化的识别研判, 势必会增加计算

复杂度和计算时间^[11]。目前风电场等效建模已经成为一项广泛应用的技术^[12]。风功率等值曲线建模方法一般分为离散法^[13-14]、参数法^[15]、非参数法^[16]、随机法^[17], 通过对风电场风速—功率曲线分析为风机布局、性能评判、估算发电量等提高重要参考。离散法、随机法由于精度不高, 使用较少; 参数方法和非参数法较为广泛, 且非参数方法更能体现风机实际运行状态。杨茂等^[18]论述了多种风功率等效建模的方法及效果, 非参数方法无论是从模型精度还是可实现操作上有显著效果。Bremne^[19]通过拟合风功率曲线, 局部线性分位数回归方法提出了基于天气预报和风电历史值概率模型。阎洁^[20]在风功率不确定分析中使用了分位数回归。分位数回归方法能准确地还原实际风功率曲线的运行状态。基于异常、缺失数据采用不同填补方式, 单点异常采用填补; 大量连续缺失考虑相关性高的风机; 风功率等值曲线适用于风速或功率单个要素异常、缺失的填补^[21-23]。

本文分析了风机异常风速产生的原因, 提出了两种清洗方法: 一是根据风机自身特性和风功率概率分布区间结合的方法; 二是根据风机间风速高相

<http://www.qxkj.net.cn> 气象科技

内蒙古自治区自然科学基金(2022MS04019)资助

作者简介: 郝玉珠, 女, 1989 年生, 硕士, 工程师, 主要从事专业气象预报技术研究, Email: haoyuzhu.cool@163.com

收稿日期: 2023 年 8 月 23 日; 定稿日期: 2024 年 7 月 22 日

* 通信作者, Email: lan_shi@sina.com

©《气象科技》编辑部 (CC BY-NC-ND 4.0)

关特征的方法^[24]。根据风功率概率分布方法确定合理区间,对异常数据的识别率更精确,修正后进行二次检查,再利用风机间高相关特征进行修正、填补,两种方法结合使用可有效提高风电场资料的精度。

1 资料与方法

1.1 资料

内蒙古自治区乌兰察布市北部风电场 X,地势开阔,风电场东西跨度 4.5 km,南北跨度 3.3 km,占地约 34 km²,平均海拔约 1500 m,有 33 台同型号(F01~F33)轮毂高度 65 m 的风电机组,切入、切出、额定风速分别为 3 m/s、25 m/s、12 m/s,总装机容量 49.5 MW。资料时间为 2020—2022 年 33 台风机的逐 15 min 测风风速、有功功率。风电场西侧 0.37 km 处设有逐 15 min 代表风况的梯度(10 m、65 m、80 m)测风塔,西侧 13 km 处为国家区域自动气象站(包含逐时 10 m 风速、风向等要素),资料统一处理为小时平均数据。分别计算测风塔、自动气象站与风机间风速的相关系数,结果为 0.74~0.86 和 0.81~0.91(通过 0.01 显著性检验),自动气象站观测风速、测风塔 80 m 风速与风机显著相关,自动站气象站风向与测风塔 80 m 风速可作为风电场参证资料^[25]。

1.2 方法

1.2.1 分段法

分段法是风速功率曲线建模的一种参数化方法^[26-28],将一个复杂的问题分解成多个简单的子问题,逐层求解,最终得到整个问题的解决方案。计算步骤:(1)将风速或功率进行分段,风速每 0.1 m/s 为一段,功率每 10 W 为一段;(2)各段采用分位数回归法,统计每段风速或功率的概率分布;(3)依据分布集中程度确定风速或功率的上限和下限,并设定阈值;(4)调整区间阈值进行验证,最终确定合理风速或功率区间。

1.2.2 分位数回归

分位数回归(quantile regression)方法^[29-30]不仅可以度量风功率的分布中心,而且能够捕捉分布的尾部特征,尤其是极值的情况,可更加全面的刻画风功率区间分布特征。

假设随机变量 Y ,受到 k 个因素 X_1, X_2, \dots, X_k 的影响,分位数回归模型为:

$$Y = [1, X_1, X_2, \dots, X_k] \boldsymbol{\beta}(\tau) \quad (1)$$

$$\boldsymbol{\beta}(\tau) = [\beta_0(\tau), \beta_1(\tau), \beta_2(\tau), \dots, \beta_k(\tau)]' \quad (2)$$

式(1)展开式为:

$$Y = \beta_0(\tau) + \beta_1(\tau)X_1 + \beta_2(\tau)X_2 + \dots + \beta_k(\tau)X_k \quad (3)$$

式中,随机变量 Y 在解释变量 $\mathbf{X} = [X_1, X_2, \dots, X_k]'$ 给定下的条件 τ 分位数; $\tau \in (0, 1)$ 为分位点; $\beta_0(\tau), \beta_1(\tau), \beta_2(\tau), \dots, \beta_k(\tau)$ 为回归系数,随着分位点 τ 变动,这一点与均值回归分析存在显著差异,在均值回归分析中,回归系数向量为常数向量。

式(2)中参数向量 $\boldsymbol{\beta}(\tau)$ 的估计,可以转化求解以下优化问题:

$$\min_{\boldsymbol{\beta}} \sum_{i=1}^k \rho_{\tau}(Y - X'_i \boldsymbol{\beta}) = \min_{\boldsymbol{\beta}} \sum_{i | Y_i \geq X'_i \boldsymbol{\beta}} \tau |Y - X'_i \boldsymbol{\beta}| + \sum_{i | Y_i < X'_i \boldsymbol{\beta}} (1 - \tau) |Y - X'_i \boldsymbol{\beta}| \quad (4)$$

式中, $X'_i (i = 1, 2, 3, \dots, k)$ 为在分位数条件下的解释变量; $i | Y_i \geq X'_i \boldsymbol{\beta}$ 表示第 i 个相应变量实际值大于等于线性回归估计值;分位数 $\tau \in (0, 1)$; $\rho_{\tau}(u) = u[\tau - I(u)]$ 为分位回归的检验函数, u 是指示关系式,当分位数为 0.5 时,就是最小一乘回归,即中位数回归, $I(\ast)$ 为示性函数。

$$I(u) = \begin{cases} 1, & u < 0 \\ 0, & u \geq 0 \end{cases} \quad (5)$$

由式(4)得出,在不同的分位数 τ 下得到不同的参数估计 $\boldsymbol{\beta}(\tau)$,就能计算出在不同分位数下解释变量对响应变量的条件分位数的影响,分位数 τ 在 $(0, 1)$ 内连续取值时,得到响应变量的条件分布,然后得到条件密度,最终得出条件密度预测。

1.2.3 线性回归法

风机风速间一般具有显著线性关系,选取至少 3 个月风速时间序列,采用线性回归法构建风机风速的回归方程,假设缺测风机风速(输出变量) P 与 n 台风机风速 W_1, W_2, \dots, W_n 有显著线性关系,满足回归方程,公式如下:

$$P = \omega_0 + \omega_1 W_1 + \omega_2 W_2 + \dots + \omega_n W_n + \varepsilon \quad (6)$$

其中, ω_0 表示截距,系数 $\omega_1, \omega_2, \dots, \omega_n$ 为回归系数, ε 为误差项。

1.2.4 权重法

利用风机风速的相关系数,选取高相关风机进行权重分配,计算缺测风机的风速,假设选取选取 n 台风机风速 W_1, W_2, \dots, W_n 和缺测风机风速(输出

变量) P , 公式表示如下:

$$P = k_1 W_1 + k_2 W_2 + \dots + k_n W_n \quad (7)$$

其中, 系数 k_1, k_2, \dots, k_n 为权重系数, k_i 计算方法:

$$k_i = \frac{r_i}{\sum_{i=1}^n r_i} \quad (8)$$

r_i 为第 i 个风机与缺测风机风速的相关系数。

2 异常数据识别和填补

2.1 风机实测资料的识别

假设忽略风机内部影响, 只考虑风速与功率的关系。依据《QX/T 645—2022 风电机组测风资料质量审核与订正》^[25] 气象行业标准的方法, 参照风电机组测风设备的量程参数(表 1), 对风电机组风速、功率进行合理性范围检查^[31-32], 当风速、功率长时间保持不变、波动异常时, 均视为不合理风速。经合理性检验 33 台风电机组有效数据完整率为 72.5%~97.3%。

表 1 风电机组风速、功率合理范围指标

Table 1 Reasonable range indicators for wind speed and power of wind turbines

指标	合理范围
有功功率/W	$[-0.1 \times \text{额定功率}, \text{额定功率} \times 1.1)$
1 h 风速连续无变化持续时间/h	< 6
1 h 风速差/($\text{m} \cdot \text{s}^{-1}$)	≤ 6
15 min 风速连续无变化持续时间/h	< 2
15 min 风速差/($\text{m} \cdot \text{s}^{-1}$)	≤ 20

2.2 风功率区间识别

图 1 为 X 风电场 F05 风机实测风速与功率散点图, 其分布基本符合风功率理论曲线, 部分数据点受检修、停机等人或天气原因表现为异常。图中在 500 W 功率处有一条较集中区域, 随风速增加, 发电功率始终不变, 可能与人工控制输出电量有关。在 0 kW 分布少量数据点, 与停机有关。依据风机测风数据合理范围统计, 对不符合风功率理论曲线集中度 70% 的点判断为数据异常点^[33]。

风功率预测通常是功率曲线的延伸预测或者功率公式参数的改进^[34]。实际风功率曲线会偏离其理论功率曲线, 具有离散性分布特征。本文基于分位数回归法, 分段统计风速、功率概率分布, 确定风速功率关系区间的合理范围, 并运用到风速识别和

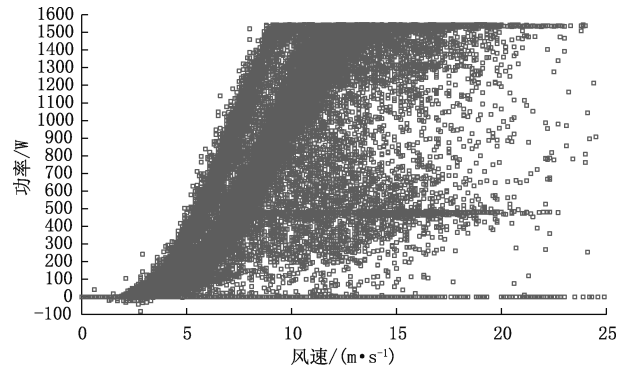


图 1 2020 年 1 月至 2022 年 12 月 X 风电场 F05 风机风速-功率散点分布

Fig. 1 Wind speed and power scatter distribution of turbine F05 in the wind farm X from January 2020 to December 2022

修正中。

具体做法参考易春燕等^[33] 风速功率关系带的算法: 将大于等于切入风速且小于切出风速的风电机组风速进行分段, 保证每段样本量充足, 计算、检验风速段及相应风功率标准偏差分布, 总体满足 70% 集中度为分段风速与实际风功率的合理波动范围, 如果该段过于集中, 提高集中度, 如果不够集中, 70% 概率范围偏大, 影响区间准确率, 因此逐步剔除 1%~10% 的边缘区, 直至满足 70% 的集中度再取值。为获取的风功率曲线更准确, 风速分段后, 再以功率分段, 两者叠加取重叠部分。

以 X 风电场风机 F05 为例, 实测风功率曲线是带状集中分布, 按照风速和功率误差波动控制在 10% 以内为分层依据^[34], 当功率以 10 W 分段, 1005 W 对应风速范围 9.1~9.9 m/s(图 2a); 当风速以 0.1 m/s 分段时, 10 m/s 风速对应功率波动范围 1021~1170 W(图 2b)。风电场每 0.1 m/s 风速间隔的功率区间, 功率按 10 W 分层, 图 2a~d 依次为风速区间、功率区间、叠加区间、风功率区间, 风功率区间则是各风速层与功率层取交集后上限和下限交点的连线。图 2d 中自 7.5 m/s 随风速增加功率区间变宽, 至 11~12 m/s 达到额定风速, 功率区间变窄, 达到额定功率后趋于平稳。例如叠加后的风功率区间, 10 m/s 风速层对应的功率变化为 930~1145 W。利用实测风功率曲线区间可以对风电机组风速、功率进行初次和订正后的审核, 经实测风功率曲线区间识别法检查后, 剔除风速、功率分布不一致的数值, X 风电场 2020—2022 年有效数据从

72.5%~97.3%降至 68.7%~92.5%。

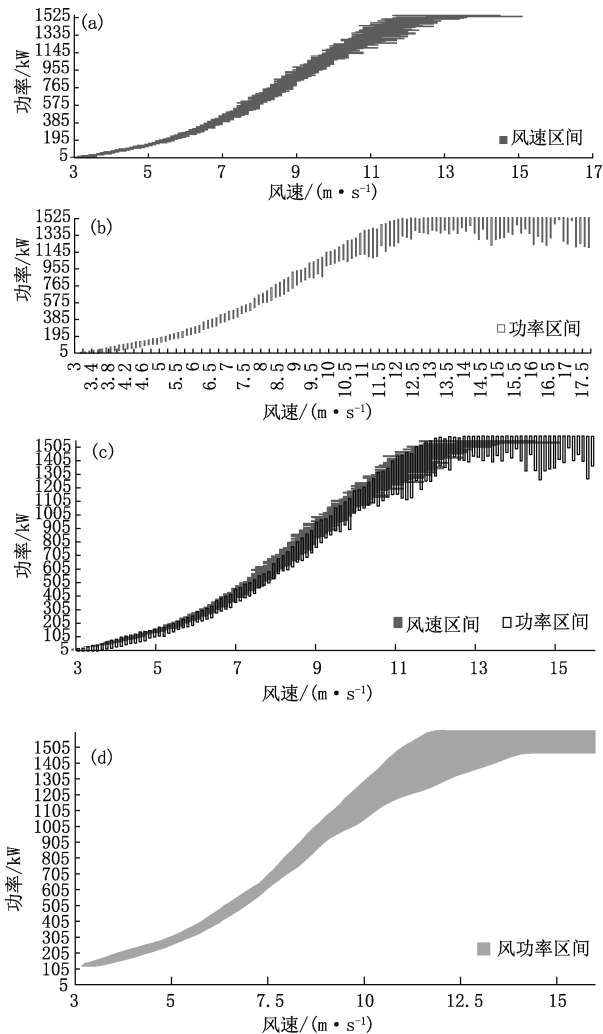


图 2 2020 年 1 月至 2022 年 12 月 X 风电场 F05 风机分段区间:(a)基于功率分段的风速区间,(b)基于风速分段的功率区间,(c)风速区间与功率区间叠加,(d)风速区间与功率区间重叠区域风功率区间

Fig. 2 F05 section interval diagram of wind farm X from January 2020 to December 2022: (a) wind speed interval based on power segmentation, (b) power interval based on wind speed segmentation, (c) wind speed interval and power interval overlay, (d) wind power interval in the overlapping area of wind speed interval and power interval

2.3 风速高相关分区和填补

通过研究风机间风速变化特征,已形成了成熟的风机风速高相关片区技术方法^[24],具体划分方法如下:将风向划分为 NW,SW,NE,SE 四个方位,即

西北、西南、东北、东南。分析风速在 1—12 月风机间的相关性,相关性范围为 0.54~0.98,将相关系数较高的值作为风速显著相关的标准,本文将 0.88 (通过 0.05 显著性水平)作为划分片区指标,当风机与风机相关系数大于 0.88,划入同一片区。将同风向不同月的相似片区进行合并(表 2),片区类型由 4 风向×12 个月,共计 48 类,合并后减少至 15 类,各片区类型包含 1~6 个风机片区,每个片区内包含若干台风速特征类型高度相同的风机。风机片区受环境影响基本一致,通过风机片区可准确推算风机异常或缺测时的实际风速和功率。

利用 X 风电场 2020 年 1 月至 2022 年 12 月风机轮毂高度逐时观测风速,采用线性相关法和权重分配法,通过风机高相关片区对缺测、异常风速进行填补、修正。分别选取风电场西北、西南、东北、东南四个方位的风机(F33,F05,F13,F24),填补和修正缺测、异常风速。模型构建期为 2020 年 1 月至 2021 年 12 月,验证期为 2022 年 1—12 月,表 3 是基于风速高相关分区的线性^[35]和权重法修正后的风速,填补相同时刻风机风速。结果表明:线性回归方法更接近实测风速,权重法在低风速段(<3 m/s)填补更接近实测风速。从构建期和验证期的平均绝对误差来看,线性法平均绝对误差低于 0.74 m/s,优于权重法 1.53 m/s 的平均绝对误差。图 3 为 2022 年 2 月 9 日 F05 风机实测风速与线性相关、权重法填补逐时风速,线性相关和权重法填补后的风速与风机实测风速变化一致性较好,线性法填补后平均绝对误差为 0.53 m/s,权重法填补平均绝对误差为 0.58 m/s,权重法填补风速偏高,但风速变化与实测风速一致性较好,线性法误差更小,修正的风速更接近实际值。

利用风功率区间识别法对线性相关或权重法填补后的风速进行审核,若其风速、功率通过风功率区间一致性检验,则插补有效,否则重新进行填补修正。经数据完整性检验,X 风电机组 33 台风机有效数据完整率从 68.7%~92.5%提高至 90.1%~92.7%(表 4)。基于风机风速高相关片区方法的线性相关和权重法有效的提高了数据完整性,为后期风速预报、功率预测提供了高质量的支撑数据。

表 2 不同风向各月的风机片区划分

Table 2 The wind turbine regional division of months with characteristics in different wind directions

风向	片区			
西北	NW1(1,7,12月)	NW2(4—6,9—11月)	NW3(2—3,8月)	
西南	SW1(1—2月)	SW2(3—5,9月)	SW3(7,10—12月)	SW4(6—8月)
东北	NE1(2—3,6,10月)	NE2(4—5月)	NE3(7—8月)	NE4(1,9,11—12月)
东南	SE1(1—2,4—6,12月)	SE2(9—11月)	SE3(3月)	SE4(7—8,10月)

表 3 基于风速高相关分区的线性填补和权重法填补风速对比

Table 3 Comparison of linear filling and weighted filling wind speeds based on high wind speed correlation zones (m/s)

	F33 风速			F05 风速			F13 风速			F24 风速		
	实际	线性法	权重法	实际	线性法	权重法	实际	线性法	权重法	实际	线性法	权重法
2022-12-31 19:00	12.9	13.9	14.9	12.6	13.7	15.0	14.8	13.6	12.6	13.3	14.3	15.0
2022-11-25 00:00	16.2	16.7	16.6	14.6	16.2	19.1	12.8	15.8	21.8	17.8	17.6	17.6
2022-11-23 03:00	6.7	5.8	6.2	5.3	5.6	6.2	6.3	6.1	4.9	4.9	5.8	6.1
2022-07-06 15:00	3.8	3.5	2.9	1.5	1.7	2.1	1.5	2.1	1.4	1.1	2.0	1.9
2022-11-14 03:00	9.4	9.5	10.2	8.3	10.0	11.7	8.1	9.3	8.8	10.3	10.8	11.3
构建期平均绝对误差		0.49	0.61		0.73	0.87		0.49	1.24		0.43	0.44
验证期平均绝对误差		0.56	1.10		0.70	1.20		0.58	1.53		0.74	0.89

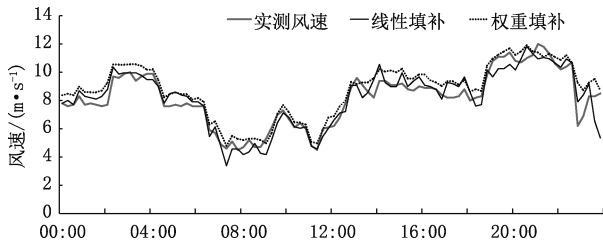


图 3 2022 年 2 月 9 日风机 F05 线性相关和权重法风速填补与实测风速对比

Fig. 3 Comparison between linear correlation and weighted wind speed filling and observed wind speed on 9 February 2022

3 结论与讨论

与常规有效性、合理性检验相比,风功率区间识别方法更科学地检查出看似合理的异常数据,再对其进行修正、复查,直至异常数据点符合风功率区间分布。基于风速高相关片区的线性法和权重法有效地提高了风机风速的完整性、准确性。两种方法的结合,将风机风速完整率从 68.7%~92.5% 提升至 90.1%~92.7%,大大提高了风机数据的质量,降低地形、尾流等环境因素的干扰。

表 4 X 风电场风机风速审核和填补修正前后完整率对比

Table 4 Comparison of data completeness rate in wind farm X after the wind speed of wind power interval verified and correction filled %

风机号	数据完整率		风机号	数据完整率	
	风功率区间审核后	填补订正后		风功率区间审核后	填补订正后
F01	72.8	90.7	F18	82.5	91.0
F02	92.5	92.7	F19	77.3	90.4
F03	71.2	90.8	F20	74.3	91.4
F04	79.6	90.6	F21	79.4	90.1
F05	82.1	92.6	F22	81.2	91.3
F06	91.8	91.9	F23	72.0	90.9
F07	91.1	92.5	F24	88.5	91.8
F08	81.9	90.9	F25	71.2	91.9
F09	89.7	90.5	F26	71.6	92.4
F10	90.7	92.7	F27	82.3	92.3
F11	69.7	92.0	F28	86.1	91.5
F12	72.1	91.7	F29	77.0	92.4
F13	81.7	92.1	F30	80.6	91.7
F14	73.8	92.2	F31	76.4	91.3
F15	69.8	91.2	F32	74.9	90.1
F16	78.1	90.1	F33	79.1	90.9
F17	81.3	90.4			

本文提出了提高风电场风速资料完整性和准确性的方法,在修正风速上基于风机高相关片区选取了线性法等常规统计算法,取得了较好的效果,下一步拟尝试选取随机森林、神经网络等人工智能方法,根据算法的适用性,进一步提高数据质量,提升风速、功率的精准度。为电力等相关部分提供更精细化的气象服务产品提供指导和依据。

参考文献

- [1] Lydia M, Kumar S S, Selvakumar A I, et al. A comprehensive review on wind turbine power curve modeling techniques [J]. *Renewable & Sustainable Energy Reviews*, 2014, 30(2): 452-460.
- [2] 杨茂, 马秀达, 温道扬, 等. 风电功率预测研究综述[J]. *电测与仪表*, 2013, 50(7): 7-10+89. Yang M, Ma X D, Wen D Y, et al. Review of wind power prediction [J]. *Electrical of Measurement & Instrumentation (in Chinese)*, 2013, 50(7): 7-10+89.
- [3] 白雁翔, 王德林, 马宁宁, 等. 大型风电场的动态等值方法研究[J]. *电工技术*, 2018(13): 46-49+52. Bai Y X, Wang D L, Ma N N, et al. Study on dynamic equivalence method of large scale wind farm [J]. *Electric Engineering (in Chinese)*, 2018(13): 46-49+52.
- [4] 高祝宇, 何好斐, 杨明. 杭州飞机气象观测资料处理及质量分析[J]. *气象科技*, 2023, 51(6): 794-804. Gao Z Y, He Y F, Yang M. Data processing and quality assessment of aircraft meteorological observation around Hangzhou [J]. *Meteorological Science and Technology (in Chinese)*, 2023, 51(6): 794-804.
- [5] 陈中钰, 杜冰, 刘康. 风廓线雷达水平风数据质量控制[J]. *气象科技*, 2020, 48(6): 801-807. Chen Z Y, Du B, Liu K. Quality control of wind profiler radar horizontal wind data [J]. *Meteorological Science and Technology (in Chinese)*, 2020, 48(6): 801-807.
- [6] 刘静, 罗勇, 杨凤志, 等. 一种测风数据处理与分析方法[J]. *船舶工程*, 2020, 42(S2): 243-247. Liu J, Luo Y, Yang F Z, et al. A method of wind data processing and analysis [J]. *Ship Engineering (in Chinese)*, 2020, 42(S2): 243-247.
- [7] 遇茜, 钱政, 聂志鹏. 风机异常及缺失数据的填补方法研究[J]. *电测与仪表*, 2020, 57(23): 1-8. Yu Q, Qian Z, Nie Z P. Research on filling method of abnormal and missing data of wind turbines [J]. *Electrical of Measurement & Instrumentation (in Chinese)*, 2020, 57(23): 1-8.
- [8] 杨茂, 白玉莹. 测风塔异常风速数据识别与补齐方法研究[J]. *可再生能源*, 2021, 39(6): 811-817. Yang M, Bai Y Y. Study on identification and completion method of abnormal wind speed data of wind tower [J]. *Renewable Energy Resources (in Chinese)*, 2021, 39(6): 811-817.
- [9] 李莉, 梁袁, 林娜, 等. 考虑时空相关性的风电机组风速清洗方法[J]. *太阳能学报*, 2023, 1-9. Li L, Liang Y, Lin N, et al. Data cleaning method considering temporal and spatial correlation for measured wind speed of wind turbines [J]. *Acta Energetica Solaris Sinica (in Chinese)*, 2023, 1-9.
- [10] 王一妹, 刘辉, 宋鹏, 等. 基于多阶段递进识别的风电机组异常运行数据清洗方法[J]. *可再生能源*, 2020, 38(11): 1470-1476. Wang Y S, Liu H, Song P, et al. An approach for the cleaning of abnormal wind turbine operation data based on multi-phase progressive recognition [J]. *Renewable Energy Resources (in Chinese)*, 2020, 38(11): 1470-1476.
- [11] 刘红柳, 杨茂. 基于小波变换和支持向量机的风电功率爬坡事件识别与预测[J]. *东北电力大学学报*, 2016, 36(6): 30-35. Liu H L, Yang M. An approach for wind power climbing event identification and prediction based on wavelet and SVM [J]. *Journal of Northeast Electric Power University (in Chinese)*, 2016, 36(6): 30-35.
- [12] 杨茂, 吕天峰, 季本明. 混沌理论在电力系统负荷预测中应用综述[J]. *东北电力大学学报*, 2015, 35(3): 18-21. Yang M, Lv T F, Ji B M. A review of chaos theory in power system load prediction [J]. *Journal of Northeast Electric Power University (in Chinese)*, 2015, 35(3): 18-21.
- [13] Carmona M, Sanz-bobi M A. Normal power generation area of wind turbines for the detection of abnormal performance [C]// *IEEE International Conference on Renewable Energy Research and Applications*. IEEE, 2016: 335-340.
- [14] 欧阳庭辉, 查晓明, 秦亮, 等. 含核函数切换的风电功率短期预测新方法[J]. *电力自动化设备*, 2016, 36(9): 80-86. Ouyang T H, Zha X M, Qin L, et al. Short-term wind power prediction based on kernel function switching [J]. *Electric Power Automation Equipment (in Chinese)*, 2016, 36(9): 80-86.
- [15] 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.
- [16] 江岳文, 温步瀛. 结合风电功率超短期预测值偏差的实时市场调度[J]. *电力自动化设备*, 2015, 35(3): 12-17. Jiang Y W, Wen B Y. Real-time market dispatch based on ultra-short-term forecast error of wind power [J]. *Electric Power Automation Equipment (in Chinese)*, 2015, 35(3): 12-17.
- [17] 赵亮, 韩华玲, 陈宁, 等. 基于模糊滑模控制的风电机组最大风能追踪策略[J]. *电力自动化设备*, 2012, 32(12): 74-79. Zhao L, Han H L, Chen N, et al. Maximum wind energy tracking based on fuzzy sliding mode control for wind power generation system [J]. *Electric Power Automation Equipment (in Chinese)*, 2012, 32(12): 74-79.
- [18] 杨茂, 杨琼琼. 风电机组风速-功率特性曲线建模研究综述[J]. *电力自动化设备*, 2018, 38(2): 34-43. Yang M, Yang Q Q. Calculation and analysis of total supply capability of distribution network considering connection mode at low voltage

- side of substation [J]. *Electric Power Automation Equipment* (in Chinese), 2018, 38(2): 34-43.
- [19] Bremnes B J. Probabilistic wind power forecasts using local quantile regression [J]. *Wind Energy*, 2010, 7(1): 47-54.
- [20] 阎洁, 刘永前, 韩爽, 等. 分位数回归在风电功率预测不确定性分析中的应用[J]. *太阳能学报*, 2013, 34(12): 2101-2107. Yan J, Liu Y Q, Han S, et al. Quantile regression in uncertainty analysis of wind power forecasting [J]. *Acta Energiæ Solaris Sinica* (in Chinese), 2013, 34(12): 2101-2107.
- [21] 李亚玲, 李涛. 改进 K-means 算法在风电异常数据的识别研究[J]. *计算机时代*, 2020(2): 6-8. Li Y L, Li T. Application of improved K-means algorithm in recognition of wind power abnormal data [J]. *Computer Era* (in Chinese), 2020(2): 6-8.
- [22] 杨茂, 王金鑫, 都键. 基于极限学习机与格兰杰检验的风电功率缺失数据补齐[J]. *东北电力大学学报*, 2019, 39(5): 9-16. Yang M, Wang J X, Du J. The complement of the missing data based on the extreme learning machine and granger test in wind power [J]. *Journal of Northeast Electric Power University* (in Chinese), 2019, 39(5): 9-16.
- [23] 胡阳, 乔依林. 基于置信等效边界模型的风功率数据清洗方法[J]. *电力系统自动化*, 2018, 42(15): 18-23+149. Hu Y, Qiao Y L. Wind power data cleaning method based on confidence equivalent boundary model [J]. *Automation of Electric Power Systems* (in Chinese), 2018, 42(15): 18-23+149.
- [24] 石岚, 徐丽娜, 郝玉珠. 基于风速高相关分区的风电场风速预报订正[J]. *应用气象学报*, 2016, 27(4): 506-512. Shi L, Xu L N, Hao Y Z. The correction of forecast wind speed in a wind farm based on partitioning of the high correlation of wind speed [J]. *Journal of Applied Meteorological Science* (in Chinese), 2016, 27(4): 506-512.
- [25] 中国气象局. 风电机组测风资料质量审核与订正: QX/T 645-2022 [S]. 北京: 气象出版社, 2022. China Meteorological Administration. Quality check and correction of wind observation data from wind turbine: QX/T 645-2022 (in Chinese) [S]. Beijing: China Meteorological Press, 2022.
- [26] 崔杨, 陈正洪, 何英杰, 等. 湖北省近地层风切变特征[J]. *气象科技*, 2021, 49(5): 785-793. Cui Y, Chen Z H, He Y J, et al. Characteristics of surface layer wind shear in Hubei province [J]. *Meteorological Science and Technology* (in Chinese), 2021, 49(5): 785-793.
- [27] Thapar V, Agnihotri G, Sethi V K. Critical analysis of methods for mathematical modeling of wind turbines [J]. *Renewable Energy*, 2011, 36(1): 3166-3177.
- [28] 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.
- [29] 顾光芹, 李宏宇, 郑雄伟, 等. 电网架空导线载流量计算气象参数组合特征及风险分析[J]. *气象科技*, 2021, 49(3): 483-490. Gu G Q, Li H Y, Zheng X W, et al. Combination characteristics and risk analysis of meteorological parameters for carrying capacity calculation of overhead line in a power grid [J]. *Meteorological Science and Technology* (in Chinese), 2021, 49(3): 483-490.
- [30] 钱磊, 邱学兴, 郑淋淋. 基于概率密度匹配方法的 WRF 模式阵风风速误差订正[J]. *气象科技*, 2019, 47(6): 916-926. Qian L, Qiu X X, Zheng L L. Error correction of WRF model gust speed based on probability density function matching method [J]. *Meteorological Science and Technology* (in Chinese), 2019, 47(6): 916-926.
- [31] 李艳萍, 黎锦雷, 陶伟, 等. 临桂站极大风速观测值可靠性分析[J]. *气象科技*, 2021, 49(3): 504-508. Li Y P, Li J L, Tao W, et al. Reliability analysis of observed maximum wind speed at Lingui station [J]. *Meteorological Science and Technology* (in Chinese), 2021, 49(3): 504-508.
- [32] 汪学渊, 丘祖京, 徐竹泉. 面向同化的探空温度数据质控参数优化及误差特性分析[J]. *气象科技*, 2023, 51(5): 629-638. Wang X Y, Qiu Z J, Xu Z Q. Analysis of parameter optimization of quality control and error characteristics of radiosonde temperature for assimilation [J]. *Meteorological Science and Technology* (in Chinese), 2023, 51(5): 629-638.
- [33] 易春燕, 李俊芳, 张步涵. 基于风速功率关系带的时间序列法预测研究[J]. *湖北工业大学学报*, 2012, 27(1): 9-13. Yi C Y, Li J F, Zhang B H. Wind power prediction based on time series [J]. *Journal of Hubei University of Technology* (in Chinese), 2012, 27(1): 9-13.
- [34] 胡庆林, 付盼, 胡忠林, 等. 风力发电场并网电力调度稳定性控制研究[J]. *计算机仿真*, 2017, 34(6): 93-97+240. Hu Q L, Fu P, Hu Z L, et al. Research on the stability control methods in power dispatch of wind power generation integrating into power system [J]. *Computer Simulation* (in Chinese), 2017, 34(6): 93-97+240.
- [35] 廖代强, 吴遥, 柴闯闯. 大气污染物非线性回归模型构建[J]. *气象科技*, 2020, 48(6): 871-876. Liao D Q, Wu Y, Cai C C. Construction of nonlinear regression model based on relationship between air pollutants and meteorological elements [J]. *Meteorological Science and Technology* (in Chinese), 2020, 48(6): 871-876.

Research on Identification and Correction Methods for Abnormal Wind Speed in Wind Turbines

HAO Yuzhu SHI Lan JIA Xiaohong

(Inner Mongolia Meteorological Service Center, Hohhot 010051)

Abstract: Effective data cleaning methods can improve the quality of wind turbine measurement data. The quality of wind turbine data plays a very important role in wind resource assessment, wind power accurate prediction, and performance diagnosis of wind turbines. There are many uncertainties in the data collection and monitoring systems of different wind turbines for fault diagnosis, which result in uneven quality of wind measurement data for wind turbines. This paper proposes a new method for identifying the probability interval of wind power. This method uses the characteristic changes between wind speed and power to clean and correct the effective data of wind turbine measurement data. It can effectively improve the utilisation rate of wind turbine data. This paper selects wind turbine data from a wind farm in the northern part of Ulanqab, Inner Mongolia Autonomous Region from 2020 to 2022. By sequentially subjecting the data to rationality and validity tests, wind power interval checks, and finally, cleaning and correcting abnormal data, which are carried out by utilising the correlation of the turbine. The final results indicate that: by using the wind power interval method, it is difficult to distinguish abnormal wind speeds. This method can improve data quality and enhance the accuracy of wind speed and power. According to statistics, the data integrity has been significantly improved from 68.7%—92.5% to 90.1%—92.7%. Above all, the data integrity has been significantly improved. This method achieves mutual calibration between wind speed and power through the wind power probability interval recognition method. It provides fundamental support data for predicting and regulating the power generation of wind farms. It provides guidance and a basis for more refined meteorological service products for power and other related sectors.

Keywords: wind turbine data; identification and correction; wind turbine area; wind power range

Received: 2023-8-23 Accepted: 2024-7-22

© Editorial Office of *Meteorological Science and Technology* (CC BY-NC-ND 4.0)