

风力发电风资源数据分析处理研究

刘文超, 掌旭, 高阳, 吴连龙, 陈志朋, 张煜

沈阳嘉越电力科技有限公司, 辽宁 沈阳

收稿日期: 2022年12月22日; 录用日期: 2023年1月9日; 发布日期: 2023年2月20日

摘要

风资源数据影响着风电系统运行可靠性, 本文针对风资源原始数据存在的问题, 分析研究了风资源所包含的各个特性参数和各特性参数的定义、适用范围以及计算方式。针对基于对测风数据在风资源精细化评估中的实际应用进行了讨论, 提出了四分位法处理数据, 保证了风电场对于风资源信息的掌握更加全面细致, 有利于风电场运行后期的发电产能评估为新能源发展奠定基础。

关键词

风资源, 风力发电, 数据分析, 数据处理

Research on Analysis and Processing of Wind Resource Data of Wind Power Generation

Wenchao Liu, Xu Zhang, Yang Gao, Lianlong Wu, Zhipeng Chen, Yu Zhang

Shenyang Jiayue Electric Power Technology Co., LTD, Shenyang Liaoning

Received: Dec. 22nd, 2022; accepted: Jan. 9th, 2023; published: Feb. 20th, 2023

Abstract

Wind resource data affect the operational reliability of wind power system. Aiming at the problem of original data of wind resource, this paper analyzes the characteristic parameters of wind resource and their definition, application scope and calculation method. The practical application of wind measurement data in the precise evaluation of wind resources is discussed, and the quartile method is proposed to process the data, which ensures that the wind farm has a more comprehensive and detailed grasp of wind resource information, and is conducive to the evaluation of power generation capacity in the later stage of wind farm operation and lays the foundation for the de-

velopment of new energy.

Keywords

Wind Resources, Wind Power, Data Analysis, Data Processing

Copyright © 2023 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

当前风电场的测风塔一般在风向合适位置一次性固定高度安装, 这样容易导致高层数据缺失, 且测风塔一般只能监测固定几个层级的高度, 不能动态实时精确监测多层面数据。大部分测风塔在运行一段时间后, 由于长时间处于缺乏维护的状态, 导致测风塔传感器持续出现偏差且得不到及时校准。测风塔采集实时数据的传输途径噪声大, 从而造成的数据缺失、数据质量差和数据不可靠等问题长期困扰风电行业。因此要对采集的原始测风数据进行处理, 从测风数据的采集分析角度看, 国内外对实时测风数据的处理修正方法非常多样, 利用邻近风电场的气象局提供的有效气象数据进行修正的数字天气预报法, 使用一些经验公式筛选订正采集到的风能数据, 另外可以利用 BP 神经网络模型得到连续函数的人工神经网络法。本文针对原始风资源数据利用四分位法、插补法和相关性处理几种方法进行了综合分析处理[1]。

2. 风能资源特性参数

因为风资源的随机变化和不稳定性, 可以通过风能评估中常用的相应各项指标表示[2], 具体的各项指标如下所示:

1) 风速

平均风速的计算公式可表现为在一个限定的时间段内将采集的所有具体风速数据求和之后除以全部的测量数据个数, 具体公式如式(1)所示:

$$\bar{v} = \frac{\sum_{i=1}^n v_i}{n} \quad (1)$$

其中, \bar{v} 指平均风速;

v_i 指各个测风数据;

n 指测风数据的个数。

在工程中, 常常用幂定律风切变指数作为考量标准, 该系数的经验公式中包含有多个变量, 其中包括了垂直高度、大气稳定度、地表粗糙度和温度。具体的经验公式如式(2)所示:

$$\chi = \frac{\lg\left(\frac{v_2}{v_1}\right)}{\lg\left(\frac{z_2}{z_1}\right)} \quad (2)$$

其中, χ 指风切变指数;

v_1 指高度为 h_1 所测风速;

v_2 指高度为 h_2 所测风速;

h_1 指初始已知高度;

h_2 指风速变化之后对应的所处高度。

若考虑了大气稳定度这一个因素, 公式可以修正为公式(3)所示:

$$\partial = \phi_m \left(\frac{s}{L} \right) / \left[\ln \left(\frac{s}{s_0} \right) - \varphi_m \left(\frac{s}{L} \right) \right] \quad (3)$$

其中, s 指高度;

L 指 M-O 长度;

φ_m 指 M-O 函数系数;

ϕ_m 指 M-O 函数系数。

上式是基于 M-O 相似理论得出的结果, 其中当大气状况为稳定时, $\phi_m = 1 + 5 \left(\frac{s}{L} \right)$; 当大气状况为中

性时, $\phi_m = 1$; 当大气状况为不稳定时, $\phi_m = \left[1 - 16 \left(\frac{s}{L} \right) \right]^{-\frac{1}{4}}$ 。

2) 风向

对于风向数据来说, 也可将其分为瞬时风向和平均风向, 瞬时风向即为每个时间间隔下所测得的实时数据, 而平均风向则可以通过多种求解方法进行判断计算, 得出该时间段的代表性平均风向。具体的求解平均风向较为常用的计算方法如下所示:

a) 算术平均计算方法

$$D_{ave} = \frac{\sum_{i=1}^{60} D_i}{60} \quad (4)$$

其中, D_{ave} 指平均风向;

D_i 指第 i 个风向值。

b) 单位矢量平均计算方法

$$D_{ave} = \arctan \left(\frac{\frac{1}{60} \sum_{i=1}^{60} \sin D_i}{\frac{1}{60} \sum_{i=1}^{60} \cos D_i} \right) \quad (5)$$

其中, D_{ave} 指平均风向;

D_i 指第 i 个风向值。

该方法是建立在算术平均计算方法上改进的, 它比算术平均计算方法在准确度上有了进一步的提升, 能够更接近于真实风向。

c) 矢量平均计算方法

$$D_{ave} = \arctan \left(\frac{\frac{1}{60} \sum_{i=1}^{60} v_i \sin D_i}{\frac{1}{60} \sum_{i=1}^{60} v_i \cos D_i} \right) \quad (6)$$

其中, v_i 指第 i 个风速值;

D_{ave} 指平均风向;

D_i 指第 i 个风向值。

这个方法是建立在单位矢量平均计算方法的基础上改进的, 考虑了各个时刻测得的风向占全部测风数据的比例大小, 使得每个方向的风向带上了权重进行平均计算, 使得结果更加真实。

3) 风功率密度

风功率密度在风能资源评估工作工程项目中往往适用范围更广[3], 风功率密度的求解方法可分为基于威布尔参数法和数学统计法。

a) 利用数学统计法计算的风功率密度[4]的数学表达式如公式(7)所示:

$$W = \frac{\rho v^3}{2} \quad (7)$$

其中, W 指风功率密度;

ρ 指空气密度;

v 指风速。

实际工程中采用年平均风功率密度来形容该地区的风能资源大小, 其中, 年平均风功率密度利用一年之内记录的所有的风功率密度平均值代替, 其数学表达式如公式(8)所示:

$$D_{wp} = \frac{\sum_{i=1}^n \rho v_i^3}{2n} \quad (8)$$

其中, D_{wp} 指年平均风功率密度;

n 指某个规定时间内记录的所有数值个数;

ρ 指当地年平均空气密度;

v_i 指第 i 个记录的瞬时风速值。

如果风电场内安装有测量压强和温度的仪器, 则其计算公式如公式(9)所示:

$$\rho = \frac{P}{RT} \quad (9)$$

其中, P 指年平均压强;

R 指气体常数, 取 $287 \text{ J}/(\text{kg} \cdot \text{K})$;

T 指年均开尔文绝对温度。

若无法实时测量大气压强和温度, 则空气密度改变为海拔高度和温度的函数, 具体数学计算公式如公式(10)所示:

$$\rho = \left(\frac{353.05}{T} \right) \exp^{-0.034 \left(\frac{z}{T} \right)} \quad (10)$$

其中, Z 指空气密度。

b) 基于威布尔参数法计算的年平均风能密度 ω 的数学计算公式如公式(11)所示:

$$\omega = \frac{1}{2} \rho c^3 \Gamma \left(\frac{3}{k} + 1 \right) \quad (11)$$

其中, ρ 指年平均空气密度, 标准状态取 $\rho = 1.225 \text{ kg}/\text{m}^3$

k 、 C 指威布尔分布参数。

4) 平均有效风能密度

$$\bar{w} = \frac{\frac{1}{2} \rho \int_{v_1}^{v_2} v^3 f(v) dv}{\exp\left[-\left(\frac{v_1}{c}\right)^k\right] - \exp\left[-\left(\frac{v_2}{c}\right)^k\right]} \quad (12)$$

5) 风能可利用小时数

风能可利用小时数指将风速数值处于预先定义的某段有效风速区间相加的小时数总和[5]。在风电领域实际工程中常常用年风能可利用小时数代替作为相应指标参数。该参数越大，说明该测风区域所掌握的风资源情况越理想。具体的数学计算公式如公式(13)所示：

$$t = n \left\{ \exp\left[-\left(\frac{v_1}{c}\right)^k\right] - \exp\left[-\left(\frac{v_2}{c}\right)^k\right] \right\} \quad (13)$$

其中， v 指测量的风速；

n 指测风设备根据测风间隔统计的平均风速段总小时数；

v_1 、 v_2 分别指机组的切入风速和切出风速。

3. 风能资源数据集的建立

3.1. 风能资源数据的收集

本文数据的采集设备是位于安徽省某风电场相距 15 千米左右的两期声雷达，采集 3~6 月份数据后输入到后台的测风数据结果表中同时记录了特定高度层的环境温度、湿度和压强大小。取其中 2020 年 3 月 11 日的 10 米和 30 米高度层的部分数据为例，收集后的部分数据见表 1：

Table 1. Real-time data recording table of acoustic radar

表 1. 声雷达实时数据记录表

时间	10 米风速	10 米风向	30 米风速	30 米风向	10 米温度	10 米湿度	10 米压强
03:00:00	1.50	101.00	0.90	81.00	8.88	76.90	1050.400024
03:01:00	1.48	105.00	0.76	94.00	9.00	76.10	1050.400024
03:02:00	1.29	107.00	0.63	100.00	9.00	76.10	1050.400024
03:03:00	1.30	110.00	0.69	105.00	8.76	75.80	1050.500000

而后将一期和二期声雷达分别在 3~5 月份各层风速平均值与最大风速值进行统计得出如下结论：

- 1) 在平均风速方面，3、4 月份一期声雷达布置地点风速明显较二期声雷达布置地点风速大，5 月份不及二期声雷达布置地点风速数值。
- 2) 在最大风速方面，一期声雷达所测得当地最高风速明显高于二期声雷达所测得当地最高风速。
- 3) 一期声雷达所布置地点风速变化较大，陡度较高。

3.2. 风能资源数据的处理

3.2.1. 数据处理的基本原理

本文基于四分位法的数据清洗办法，四分位法即是将整个数据库的数据按大小有序排列，而后通过三个数据点将该序列进行等部分分组，分成等量的四组数据，设数据库中的三个数据点分别为 P_1 、 P_2 、 P_3 ，该数据序列为 $X = \{x_1, x_2, \dots, x_n\}$ 。具体三个数据点的计算公式如(14)~(21)所示：

1) 首先计算数据序列的中点 P_2

$$P_2 = \begin{cases} \frac{x_{n+1}}{2} & n = 2k + 1, k = 0, 1, \dots \\ \frac{x_k + x_{k+1}}{2} & n = 2k, k = 1, 2, \dots \end{cases} \quad (14)$$

2) 接着计算数据序列分点的另外两个 P_1 、 P_3

当 $n = 2k + 1, k = 0, 1, \dots$ 时, 则有

$$\begin{cases} P_1 = \frac{x_{k+1}}{4} & k = 1, 3, \dots \\ P_3 = \frac{x_{3(k+1)}}{4} & k = 1, 3, \dots \end{cases} \quad (15)$$

$$\begin{cases} P_1 = \frac{\frac{x_k + x_{k+2}}{2}}{2} & k = 0, 2, \dots \\ P_3 = \frac{\frac{x_{3k+2} + x_{3k+4}}{2}}{2} & k = 0, 2, \dots \end{cases} \quad (16)$$

当 $n = 2k, k = 1, 2, \dots$ 时, 则有

$$\begin{cases} P_1 = \frac{x_{k+1}}{2} & k = 1, 3, \dots \\ P_3 = \frac{x_{3(k+1)}}{2} & k = 1, 3, \dots \end{cases} \quad (17)$$

$$\begin{cases} P_1 = \frac{\frac{x_k + x_{k+2}}{2}}{2} & k = 2, 4, \dots \\ P_3 = \frac{\frac{x_{3k} + x_{3k+2}}{2}}{2} & k = 2, 4, \dots \end{cases} \quad (18)$$

3) 定义四分位距 O_{PR} 的计算公式如下所示:

$$O_{PR} = P_3 - P_1 \quad (19)$$

4) 而后可通过四分位距和数据分点计算上下限, 若是超过如下范围的数据即可认为是错误数据, 设上下限分别为 W_1 、 W_2 , 由此可计算出合理数据的范围, 具体如下所示:

$$[W_1, W_2] = [P_1 - 1.5O_{PR}, P_3 + 1.5O_{PR}] \quad (20)$$

通过四分位法对无效数据进行清除后, 利用插值法对空余时间点对应的数据进行补充。具体计算公式如下所示:

$$x_n = \frac{x_{n-1} + x_{n+1}}{2} \quad (21)$$

依据以上计算过程拟对数据库采集的数据进行修正操作, 实际步骤如下:

1) 将数据根据实际功率从小到大有序排列, 整体功率范围处于 0~150 兆瓦内, 将其按照每 10 兆瓦进行等距分割, 重新分为依次的 15 个同等大小范围的小集合, 分别记为 $X = \{x_1, x_2, \dots, x_n\}$ 。

2) 对每个小集合 $X = \{x_1, x_2, \dots, x_n\}$ 再依据各层高度的风速大小进行重新排序, 得到新的集合 $X' = \{x'_1, x'_2, \dots, x'_n\}$, 而后利用以上的计算公式进行数据剔除和插值工作, 其中通过四分位法清除的数据

占总数据集的比例见表 2:

Table 2. Data cleaning proportion table
表 2. 数据清洗占比表

各高度层的数据清洗所占比例						
时间	10 米	30 米	70 米	80 米	85 米	95 米
3 月	1.66%	1.34%	1.85%	1.78%	1.85%	2.10%
4 月	8.43%	7.98%	7.58%	8.47%	9.36%	9.05%
5 月	0	0.13%	0.09%	0.13%	0.36%	0.36%
6 月	1.42%	1.26%	0.47%	0.31%	0	0

其中数据清洗所占的比例较低, 不会影响数据本身的真实性和合理性, 由此建立起的数据集相较之前更加的合理化, 以便后续的工作。

3.2.2. 问题数据的处理办法

在数据验证合理性的过程中, 采用国标 GB/T18710-2002 风电场风能资源评估方法[6] [7], 该标准内规定了适用范围、引用标准等。

在现场操作中会出现采集到的初始数据存在部分缺失或少数明显超出合理范围的错误, 其具体的解决办法如下所示:

1) 对于数据的缺失: 可利用多个高度测得的风速数据, 根据经验公式对其缺失的数据进行估计; 可采用测风区域条件类似的声雷达对应的风速数据, 并对其进行相关性修正作为缺失后弥补的数据; 也可通过气象数据的近似分析进行修补。

2) 对于数据的错误: 可去除个别不符合整个风速变化规律的数值, 预估一个近似范围的数值区间, 尽量使得风速变化曲线连贯、圆滑; 也可利用同处于邻近平坦地形或同处于同一水平高度复杂地形下的另一个声雷达所测得数据加以修正代替。

3.3. 风能数据的精细化评估应用

1) 风向

通过相应工具对测得数据进行统计修正更精确地评估风资源概况, 得到相应风资源概况图, 有利于直观地看出风速风向等参数的变化情况。本文基于安徽省装设的两期声雷达, 将一期与二期的声雷达 2020 年 3~5 月份的多层测风数据提取出来, 通过预处理得出每个月每间隔 1 分钟所测得的风向所处的 16 个风向方位区间的概率大小, 分别得到了一期和二期 3~5 月份的 16 方位风向图, 一期声雷达 3 月风向见图 1。

安徽省 3~5 月份风向数据在大体上的风向仍是比较相似; 同一个声雷达在同一时刻测得的风向在不同高度层会出现不同方向的可能性, 这是由于当处于近地面时, 风能受到地形、地面存在的大型障碍物和地表粗糙度等影响; 风能的变化具有极大的随机性, 可以通过声雷达的多层测风, 多时间频率测风, 有利于时刻准确地测量风速、风向等风资源数据。

2) 风速

对声雷达的风速数据在实际应用中的特点进行分析发现, 其可以通过风速数据图得出地形图范围因素对风电场风能资源评估准确性的影响。一期声雷达月测风风速数据见图 2, 可知中风存在着不确定性显而易见, 无规律可寻。

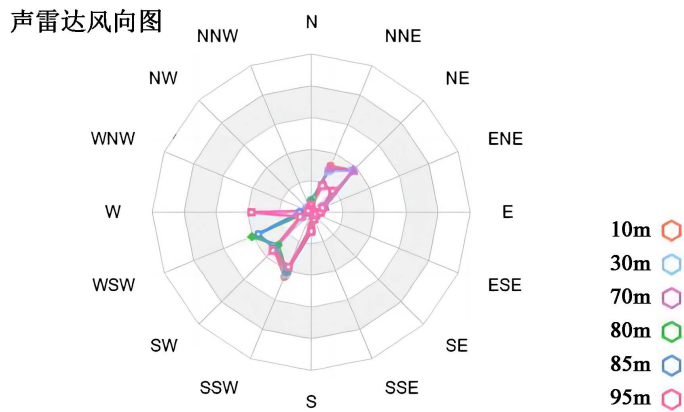


Figure 1. Wind direction chart of 1st stage acoustic radar in March
图 1. 一期声雷达 3 月风向图

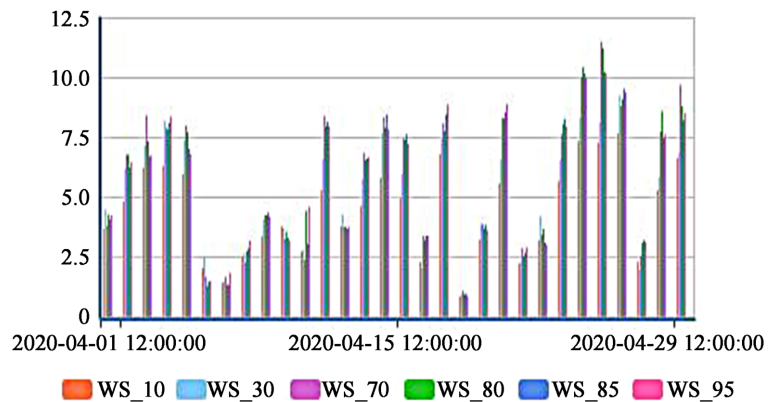


Figure 2. Monthly wind speed data of acoustic radar
图 2. 一期声雷达月测风风速数据图

同时分析复杂山地地形图范围因素对风电场风能资源评估准确性的影响，特扩大测量时间范围，将 3~6 月份每间隔 1 分钟所测得的风速数据进行统计分析，一期声雷达 3~6 月测风风速数据见图 3。

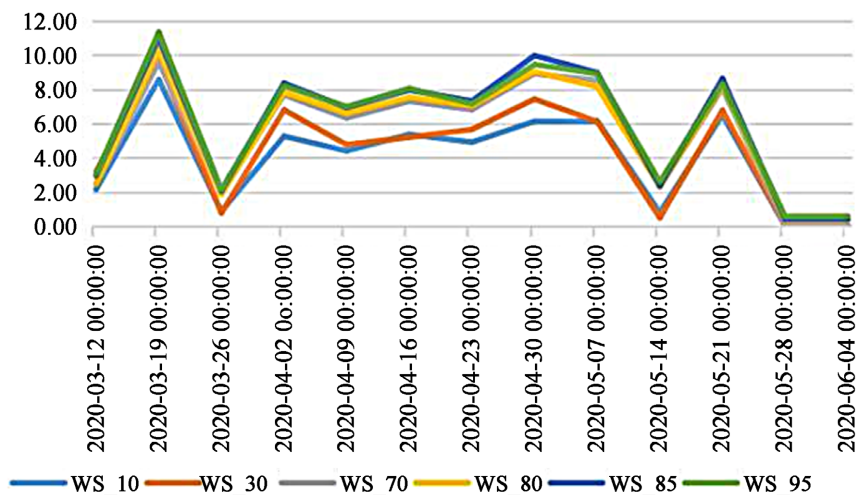


Figure 3. Wind speed data of acoustic radar from March to June
图 3. 一期声雷达 3~6 月测风风速数据图

通过比较数据图, 一期二期声雷达大体上风速变化呈现相似趋势, 说明某一地区的某一局部范围风速变化情况大体一致, 这有利于风资源精细化测量、场站进一步精细化选址。从整个 3~6 月风速数据来看, 声雷达在各层级的测风数据依旧符合随着测量高度的升高, 风速数据普遍上升的规律。

4. 结论

本文针对原始测风数据进行了研究并以安徽省某风电场的风能数据为例进行了分析并得到以下结论。对于数据的缺失可利用经验公式对其缺失的数据进行估计, 可采用测风区域条件类似的声雷达对应的风速数据, 并对其相关性修正作为缺失后弥补的数据, 也可通过气象数据的近似分析进行修补; 风能的变化具有极大的随机性, 可以通过声雷达的多层测风, 多时间频率测风, 有利于时刻准确地测量风速、风向等风资源数据; 风速变化呈现相似趋势, 说明某一地区的某一局部范围风速变化情况大体一致, 这有利于风资源精细化测量、场站进一步精细化选址。

参考文献

- [1] 李小倩. 基于风资源评估的风电场选址研究[J]. 中国战略新兴产业, 2018(40): 16.
- [2] 王欢. 48 MW 风电场选址及其实现研究[D]: [硕士学位论文]. 北京: 华北电力大学, 2018.
- [3] 殷冬琴. 风电场风资源评估及利用情况分析[J]. 轻工科技, 2018, 34(11): 65-70.
- [4] 单然. 复杂地形下的风能资源分析与评估[D]: [硕士学位论文]. 北京: 华北电力大学, 2018.
- [5] 李天成, 吴风波, 李艺等. 基于 WAsP 软件对某风电场的风资源评估研究[J]. 四川建筑, 2019, 39(2): 350-352.
- [6] 李海波. 结合计算流体力学方法的山地风电场风能资源评估研究[D]: [硕士学位论文]. 北京: 华北电力大学, 2016.
- [7] 刘玮. 风电场风能资源精细化评估系统及应用研究[D]: [硕士学位论文]. 天津: 天津大学, 2016.