

华梵前山校区风速统计分布之探讨^{*}

曹舜评¹, 范晓丹², 李汴军¹

(1. 华梵大学 环境与防灾设计学系, 台湾 台北 223; 2. 屏东科技大学水土保持系, 台湾 屏东 912)

摘要: 研究针对华梵前山气象站 2002 年 5 月至 2006 年 12 月之小时风速资料进行统计分析, 探讨华梵前山风速资料之分布是否合乎韦伯分布(Weibull distribution)及莱利分布(Rayleigh distribution)。研究中韦伯分布使用动差法(Method of moment)及最大概似法(Method of maximum likelihood)求取参数; 莱利分布使用动差法求取参数。进一步以 $K-S$ 检定, 判断各理论累积机率分布与观测值之累积机率是否相符; 再以误差绝对值之总和判断何种分布较为适合。分析结果显示: 韦伯分布与莱利分布均可适用于描述全部数据之机率分布。将资料分为夏半年(4-9月)与冬半年(10月至翌年3月)时, 韦伯分布为最适用之机率分布函数; 且夏半年之参数以最大概似法求得最佳, 而冬半年则以动差法较佳。此结果可作为校区风蚀、防灾、生态、环境风场特性及评估风能潜势研究之参考。

关键词: 韦伯分布; 莱利分布; 动差法; 最大概似法

中图分类号: P425.44

文献标识码: A

文章编号: 1005-3409(2009)06-0248-05

Analysis of Statistical Distribution to Wind Speed Data at Huafan Campus

TSAO Shun-ping¹, FAN Shiao-tan², LEE Beng-chun¹

(1. Department of Environmental & Hazards-Resistant Design, Huafan University, Pingtung, Taiwan 223, China; 2. Department of Soil and Water Conservation, Pingtung University of Science and Technology, Pingtung, Taiwan 912, China)

Abstract: This study investigates the statistic distribution of hourly mean wind speed data recorded at west campus of Huafan University from 2002 to 2006. Weibull distribution and Rayleigh distribution are adopted to fit wind speed data. The parameters of these distribution were estimated by using method of moment and method of maximum likelihood. The results show that the statistical distribution of all hourly mean wind speed data is suitably described by the Weibull distribution and Rayleigh distribution. Furthermore, wind speed data are separated into summer parts (April to September) and winter parts (October to March). The data of two parts are more suitable described by the Weibull distribution. The parameters estimated by method of maximum likelihood are better for summer parts, and the parameters estimated by method of moment are better for winter parts.

Key words: Weibull distribution; Rayleigh distribution; method of moment; method of maximum likelihood

风代表大气中气流运动的物理量, 在量化上主要包含了水平的风向与风速。风可以传送或交换地表的热量和水汽, 不管在短期之天气变化上或是长期之气候特性上, 均是一个重要的影响因子。

风的形成是因为地表上气压分布不均匀所致, 一般可分成广域及区域性两大类, 广域性质的风, 起因于地球转动所形成的气流, 即称为行星风系, 包括赤道无风带、信风带、西风带以及极地东风带等; 区

域性的风, 是由地理环境造成气压不均所形成, 称为地方风, 如海陆风、山谷风、焚风及落山风等。

在风的研究上常将风速视为随机变量, 探讨风速发生之机率, 作为风蚀、防灾、生态、环境特性或风能应用之参考。在风蚀防治时, 风速机率可作为砂粒启动的参考; 进一步评估防风定砂设施或方法的适用性, 可以利用前后风速机率分布的变化, 作为判断的依据。另外台风来临时所带来的强风暴雨, 对

* 收稿日期: 2009-08-16

作者简介: 曹舜评(1959-), 博士, 助理教授, 主要从事坡地水文及土壤物理的研究。E-mail: sptsao@huafan.hfu.edu.tw

于居家环境具有破坏性, 此时极值风速的机率研究, 有助于对建物设施之耐风强度, 提供重要之信息。候鸟栖息、风力发电及建筑物的环境风场等问题, 亦经常需要风速数据作为分析的基本资料。

在风速的机率研究上, 常使用之分布有韦伯分布 (Weibull distribution)、莱利分布 (Rayleigh distribution) 及极端值第一型 (Extreme value Type I distribution)。Lun and Lam^[1] 曾利用韦伯机率密度函数, 分析 1968–1997 年香港京士柏、中环广场和横瀾岛 3 个不同地方之长期风速观测资料。庄月璇^[2] 采用 1961–2000 年台湾地区 25 个主要气象测站之风速和风向资料进行分析, 将气象测站各风向之风速资料分别以极端值第一型分布、极端值第二型分布、韦伯分布及莱利分布 4 种理论机率分布进行分析, 并利用 K-S 适合度检定, 发现韦伯函数最适合用于台湾地区平时风速的机率分布, 极端值第一型最适合极值风速。

张仓荣^[3] 评估台湾地区的风能时, 针对平均风速采用莱利分布与韦伯分布进行统计研究, 研究结果表明: 韦伯分布为最适合用于描述台湾地区风速之机率分布, 而莱利分布次之。但因韦伯分布求解参数较繁复, 而莱利分布之参数只需藉由测站之风速平均值即可求得, 因此研究中最后采用莱利分布描述风场。

丁明等^[4] 指出风力发电厂风速机率分布为风能资源统计特性的重要指标, 并将实测的风速数据, 应用最大似法求解韦伯机率分布之参数, 由此估算得风能特征指标值, 同时与历史风速数据之实测值比较, 结果显示最大似法计算精度高, 且韦伯分布作为风力发电厂风速统计模型能准确拟合风能之实际情况, 具实用价值性。陈永盛^[5] 进行风力发电机组之电力系统可靠度分析时, 考虑韦伯分布具有两个可调整之参数, 故比其它机率分布更能符合风速的分布特性。

台湾因气候和地理因素的复杂性, 形成各地风速分布多变。华梵校区位于台湾东北部之丘陵区, 受地形及海拔高度影响, 风速亦时有所异, 为了解台湾北部丘陵区风速之特性, 故以华梵作为代表, 尝试利用统计方法, 探讨华梵校区风速之特性, 以作为了解区域性风场之前期研究。

本研究利用 2002 年 5 月至 2006 年 12 月, 华梵校区前山气象站观测之风速资料, 将每小时的平均风速, 进行统计分布分析, 研究中使用动差法及最大

似法求得韦伯及莱利机率分布之参数, 进一步利用 K-S 适合度检定, 探讨各机率分布在华梵校区之适合性, 并利用累积机率分布与实测数据计算其绝对误差, 作为最适分布之判定。

1 背景资料

华梵大学校区位于台湾北部, 台北市东南方石碇乡境内之大仑山上, 地形上属台北盆地周边之基隆、新店丘陵。校区共有 34 hm², 一陵线将校区分隔为东西两部分。校区西侧面积约为 20 hm², 高程为 570~430 m, 西侧校区为面向西南之凹型坡地, 校舍建筑多位于此, 因学生活动多位于西侧校区, 故称为前山。校区东侧面积约 14 hm², 高程为 570~350 m, 地形较为陡峭, 为面东之陡坡地, 东界为一谷地, 有常流水通过, 除设有废弃物堆置场外多为未开发区。

华梵前山气象站建于校门口停车场旁调度室之楼顶, 附近宽广无建物屏障, 于 2002 年 4 月设置完成, 5 月开始观测, 量测项目包含风速、风向、降雨、温度及湿度。前山测站海拔 465 m, 地理位置 121°41′29.07″E, 24°58′43.4″N, 气象站之区域位置图如图 1。

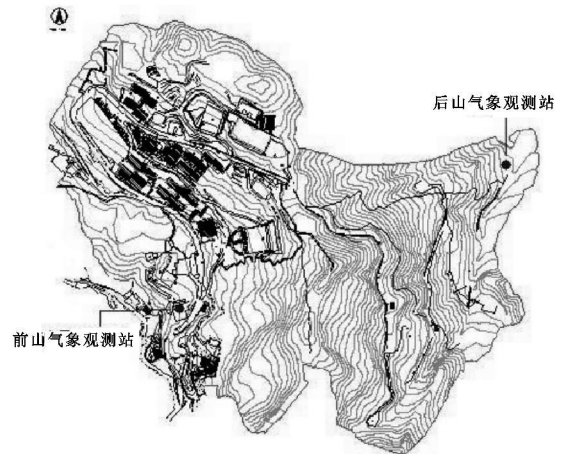


图 1 前山气象站之位置图

本研究使用华梵校区前山气象观测站 2002 年 5 月至 2006 年 12 月之每小时风速数据, 此小时数据依世界气象组织 (WMO) 所订, 为每一正时前 10 min 内之风速平均, 记为此一小时之平均风速。为避免高估风速, 故将蒲福风级表 7 级以上之疾风风速予以删除, 再进行分析。另由于风速受季风影响而有不同之变化, 故将一年分为两部分, 将 4–9 月定义为夏半年, 10 月至翌年 3 月定义为冬半年, 以利比较差异。因气象站建于 2002 年 4 月, 故资料起始于 2002 年 5 月; 分析期间至 2006 年 12 月 (如表 1)。

表 1 前山气象站资料 2002- 2006 年夏半年冬半年之时间范围

年份	前山气象站	
	夏半年	冬半年
2002	5- 9 月	10- 12 月及翌年 1- 3 月
2003- 2005	4- 9 月	10- 12 月及翌年 1- 3 月
2006	4- 9 月	10- 12 月

2 研究理论

2.1 统计参数

在统计分析中, 常未知母体 (population) 的特性, 往往只能针对样本进行统计分析, 为了解母体的特性, 及预测未来母体的发生机率, 故藉由样本之相关参数进行分析, 其参数如下:

平均值:

$$x = \frac{1}{n} \sum_{i=1}^n x_i \tag{1}$$

样本变异数:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \tag{2}$$

式中: x_i ——观测值; n ——样本数; \bar{x} ——平均值。

2.2 韦伯累积机率分布

针对平时风速进行研究时, 在机率密度函数中, 以韦伯分布与莱利分布的使用最为普遍^[1-3]。

韦伯分布为二参数机率分布, 其累积机率函数为:

$$F(x) = 1 - \exp[-(\frac{x}{b})^a] \tag{3}$$

式中: a ——形状参数 (Shape parameter); b ——尺度参数 (Scale parameter); x ——风速 (m/s)。其参数可利用动差法 (Method of moment) 或使用最大概似法 (Method of maximum likelihood) 参数迭代求得。

动差法是利用风速的平均值、标准差推求机率函数之参数, 其优点是简单快速, 其缺点是风速的样本数不够多时, 计算之平均值和标准差不具有代表性。韦伯分布参数 a 、 b 与变量平均值、标准差之关系为:

$$x = b \cdot \Gamma(1 + \frac{1}{a}) \tag{4}$$

$$s^2 = b^2 [\Gamma(1 + \frac{2}{a}) - \Gamma^2(1 + \frac{1}{a})] \tag{5}$$

式中: x ——平均值; s ——标准差; Γ ——伽玛函数 (Gamma function), 其定义为:

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx \tag{6}$$

若平均值与标准差皆为已知, 则形状参数之近似值为:

$$a = (\frac{s}{x})^{-1.086} \tag{7}$$

最大概似法是假设风速 x 符合某机率密度函数 $f(x, a, b)$, 其中 a 和 b 为模式参数。为了推求最可能之模式参数, 假设在随机取样的条件下, 发生某特定风速值 x_1 的机率为 $f(x, a, b)$, 而得到 n 个相互独立的风速值 x_1, x_2, \dots, x_n 的可能性为:

$$L(x_1, x_2, \dots, x_n; a, b) = f(x_1; a, b) \cdot f(x_2; a, b) \dots f(x_n; a, b) \tag{8}$$

式中: $L(x_1, x_2, \dots, x_n; a, b)$ ——风速 x_1, x_2, \dots, x_n 的概似函数, 将概似函数对参数 a 和 b 微分, 使其为零即可求得概似值, 而此概似函数中, 为最大值之模式参数则称最大概似值。

其韦伯分布参数公式如下:

$$a = \left[\frac{\sum_{i=1}^n x_i^a \ln(x_i)}{\sum_{i=1}^n x_i^a} - \frac{\sum_{i=1}^n \ln(x_i)}{n} \right]^{-1} \tag{9}$$

$$b = (\frac{1}{n} \sum_{i=1}^n x_i^a)^{\frac{1}{a}} \tag{10}$$

(9) 式的 a 值是利用试误法迭代推求, 使其左右两式相等得到 a 值, 而后再将 a 值代入式 (10) 即可求得 b 值。

2.3 莱利累积机率分布

莱利分布为单一参数之机率分布, 其累积机率函数为:

$$F(x) = 1 - \exp(-\frac{x^2}{2b^2}) \tag{11}$$

式中: b ——尺度参数; x ——风速 (m/s)。尺度参数 b 的推求可利用动差法求得, 与平均值 x 之关系为:

$$x = b \cdot \sqrt{\frac{\pi}{2}} \tag{12}$$

2.4 K-S 适合度检定

决定机率函数之参数后, 必须检验假设之理论机率函数和变量观测之发生机率是否相符, 此种检验称为适合度检定 (Goodness-of-Fit Test)。K-S 检定为适合度检定中的其中一种, 其仅适用于连续型机率函数, 因 K-S 检定不需将数据分组, 可避免误差, 且此检定使用累积序列数据 (Cumulative ranked data) 为基础, 故可用来检验变量的观测累积机率与理论累积机率函数的密合程度。K-S 适合度检定之计算流程如下:

- (1) 计算观测的累积机率曲线 $S_n(x_k)$ 。
- (2) 计算理论累积机率函数 $F(x)$ 与观测之累积机率 $S_n(x_k)$ 之间最大差值: $D_n = \max |F(x) - S_n(x)|$ 。
- (3) 在某一显著水平 α 下, 比较最大差值 D_n 与临界值 D_n^α : $P(D_n \leq D_n^\alpha) = 1 - \alpha$ 。临界值可由查表得知。

(4)若最大差值 D_n 小于临界值 D_n^α , 则表示在显著水平 α 下, 可接受所假设的机率分布函数。

3 分析与讨论

3.1 全部风速资料之分析

利用 2002 年 5 月至 2006 年 12 月华梵校区前山气象站观测之风速资料, 将每小时的平均风速, 采用韦伯及莱利机率分布进行分析, 利用动差法及最大概似法来求得参数, 所得结果如表 2。从表中可看出, 韦伯分布中参数 a 和 b 之值, 两种方法所得之参数值相当接近。再将求得之参数代回机率分布函数, 并绘制观测与理论之累积机率分布图(图 2-4)。

表 2 前山气象站韦伯分布与莱利分布之参数

2002-2006 年	韦伯分布		莱利分布
	动差法	最大概似法	动差法
a	1.54	1.61	-
b	1.58	1.61	1.14

注: a 为形状参数, b 为尺度参数。

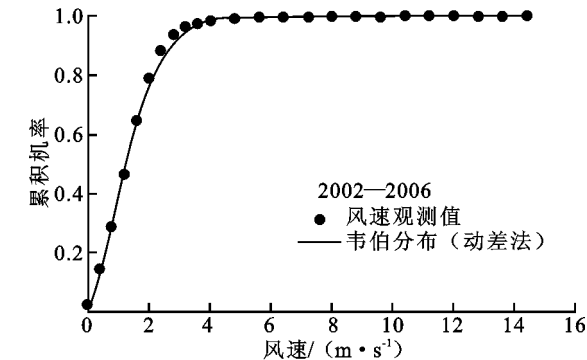


图 2 观测值与以动差法求得参数之韦伯累积机率分布

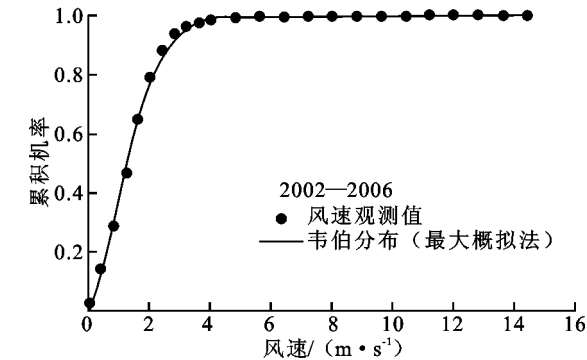


图 3 观测值与以最大概似法求得参数之韦伯累积机率分布

为进一步了解假设之理论累积机率函数与观测累积机率函数是否符合, 故利用 $K-S$ 适合度检定来判别, 如表 3, 表中之临界值为显著水平在 5% 之数值, 经计算结果发现, 前山地区韦伯分布及莱利分布皆通过检定。

$K-S$ 检定只能判别理论之分布是否适合描述观测之数据, 并无法判定使用之机率分布何者较佳。

研究中进一步计算误差绝对值之总和作为判定, 即将理论与观测累积机率分布相减, 取绝对值后加总, 所得数值较小者作为较为适宜之判定。结果为动差法计算韦伯分布参数所得之理论值与观测值之绝对误差总和为 0.136, 而以最大概似法所得误差总和为 0.107, 两种计算方式所得之误差值差异不大。而以动差法计算莱利分布参数所得之理论值与观测值之绝对误差总和为 0.130。三者之误差值均相当, 故以全部资料分析时, 韦伯分布及莱利分布均适用于华梵校区前山测之风速之机率分布。

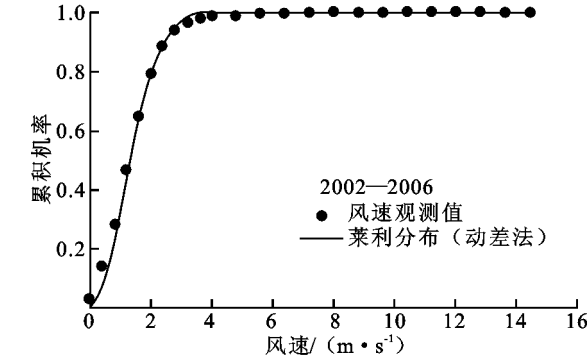


图 4 观测值与以动差法求得参数之莱利累积机率分布

表 3 $K-S$ 适合度检定(临界值 0.131)

气象站	机率分布	样本数	最大差值	检定结果
前山	韦伯分布 (动差法)	108	0.038	通过
	韦伯分布 (最大概似法)	108	0.044	通过
	莱利分布 (动差法)	108	0.080	通过

3.2 夏半年与冬半年风速之分析

由于风速具有随机变量的特性, 会因天气形态不同而产生变化, 台湾北部地区受季风影响甚大, 夏季常受西南风之影响, 冬季受东北季风影响, 故将气象站之观测风速分为夏半年与冬半年 2 种, 并逐年利用韦伯分布与莱利分布进行分析, 其参数仍使用动差法或最大概似法分别进行计算, 参数如表 4 和表 5。

为了进一步了解华梵校区夏半年及冬半年适合之风速机率, 故计算实际风速与理论风速之绝对误差值总和(如表 6 和表 7)。由表 6 中可知前山气象站之夏半年, 以最大概似法计算韦伯分布之参数, 所得误差总和最小, 其值为 0.751, 故夏半年以最大概似法求取参数之韦伯分布最为合适。由表 7 中可看出冬半年以动差法计算韦伯分布之参数, 所得误差总和最小, 其值为 0.658, 故冬半年以动差法求取参数之韦伯分布最为合适。综合论之, 韦伯分布为最佳描述夏半年及冬半年之机率分布, 而莱利分布在

夏半年与冬半年所得之误差值,除了 2004 年外,其余误差总和皆较韦伯分布大,显示莱利分布为不适用之机率分布。

表 4 2002– 2006 历年夏半年之韦伯与莱利分布参数

夏半年	韦伯分布				莱利分布
	动差法		最大概似法		动差法
	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>	<i>b</i>
2002	1.48	2.02	1.53	2.02	1.45
2003	1.68	1.51	1.69	1.53	1.08
2004	1.37	1.61	1.46	1.63	1.17
2005	1.27	1.47	1.35	1.50	1.09
2006	1.18	1.21	1.44	1.42	0.91

表 5 2002– 2006 历年冬半年之韦伯与莱利分布参数

冬半年	韦伯分布				莱利分布
	动差法		最大概似法		动差法
	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>	<i>b</i>
2002	2.12	1.70	2.10	1.65	1.16
2003	2.31	1.82	2.28	1.76	1.24
2004	1.86	1.75	1.85	1.74	1.23
2005	1.60	1.40	1.73	1.42	0.99
2006	1.86	1.60	1.98	1.58	1.11

表 6 夏半年之韦伯与莱利分布误差值

夏半年	韦伯分布		莱利分布
	动差法	最大概似法	动差法
2002	0.172	0.169	0.269
2003	0.077	0.102	0.203
2004	0.138	0.106	0.124
2005	0.155	0.124	0.221
2006	0.245	0.250	0.520
绝对误差总和	0.787	0.751	1.337

4 结 论

利用韦伯及莱利机率分布分析 2002 年 5 月至 2006 年 12 月华梵前山气象站之逐时风速数据,分布之参数是利用动差法与最大概似法求得,并采用 *K-S* 适合度检定判别其适合性,并以误差绝对值总和找出适用于华梵校区之机率分布。

由前山气象站之全部风速数据分析结果显示,

在韦伯分布中,以动差法及最大概似法计算参数,皆通过适合度检定,说明韦伯分布为适合描述华梵校区之风速机率分布。在莱利分布所得误差总和值与韦伯分布相近,故莱利分布亦适用于研究地区之风速机率分布。

表 7 冬半年之韦伯与莱利分布误差值

冬半年	韦伯分布		莱利分布
	动差法	最大概似法	动差法
2002	0.071	0.135	0.161
2003	0.078	0.124	0.261
2004	0.087	0.092	0.067
2005	0.206	0.186	0.299
2006	0.216	0.212	0.253
绝对误差总和	0.658	0.749	1.041

夏半年与冬半年风速之分析中,夏半年以最大概似法计算韦伯分布之参数,所得误差值最小。冬半年则以动差法计算韦伯分布之参数,所得误差值最小。而莱利分布除 2004 年外,其它部分误差值皆较大,故莱利分布较不适用于描述夏半年及冬半年。

经分析结果显示:韦伯分布虽因计算参数方法不同,而造成结果的差异,但整体来说,韦伯分布为最适用于华梵校区之风速机率分布,藉此可作为校区风蚀、防灾、生态、环境风场特性及评估风能潜势等研究之参考。

参考文献:

[1] Lun I Y F, Lam J C. A study of Weibull parameters using long-term wind observations[J]. Renewable Energy, 2000, 20: 145- 153.

[2] 庄月璇. 台湾地区风速机率分布之研究[D]. 台湾 桃园:“ 中央大学” 土木工程研究所, 2001.

[3] 张仓荣, 谢怡芳, 许华倚. 台湾地区风能密度与适当风力机选择之先期评估[J]. 农业工程学报, 2003, 49(1): 52-60.

[4] 丁明, 吴义纯, 张立军. 风电厂风速概率分布参数计算方法的研究[J]. 中国电机工程学报, 2005, 25(10): 107- 110.

[5] 陈永盛. 具风力发电机组之电力系统可靠度分析[D]. 台湾 桃园: 中原大学电机工程研究所, 2006.