

文章编号:1004-4574(2006)05-0076-07

中国大陆地区年最大平均风速的概率密度函数

李 杰, 陈建兵, 张琳琳, 韦 笠

(同济大学 土木工程学院, 上海 200092)

摘要:根据中国大陆地区 1951-2002 年的实测风速数据资料,采用密度演化方法进行了各站点年最大平均风速的概率密度函数估计,给出了大陆地区年最大平均风速均值和标准差的等值线分布图。与常用的基于拟合优度检验的概率密度函数估计和统计量分析不同,所提方法不需要进行先验分布类型的假定,可以通过直接计算给出与基本数据经验分布函数符合得较好的概率密度函数与概率分布函数。对比研究表明,该方法给出的统计结果是可信的,具有工程实用参考价值。

关键词:年最大平均风速;经验分布函数;概率密度函数;概率密度演化方法

中图分类号:TU312+.1

文献标识码:A

Probability density function of yearly maximum average wind speed in mainland of China

LI Jie, CHEN Jian-bing, ZHANG Lin-lin, WEI Li

(School of Civil Engineering, Tongji University, Shanghai 200092, China)

Abstract: Based on surveying data obtained in the mainland of China from 1951 through 2002, the probability density function of the yearly maximum average value of the wind speed is estimated through the probability density evolution method. In contrast to what is done in the widely used test-of-goodness-of-fit based statistical methods, a priori probability density function is not needed in the presented method. The probability density function that is directly obtained by the probability density evolution method accords well with the empirical distribution function. The investigations indicate that the probability density functions of the basic wind speed and the yearly maximum wind speed distribution are reliable and valuable in engineering practice.

Key words: yearly maximum average wind speed; empirical distribution function; probability density function; probability density evolution method

基本风速(亦称标准风速)是风的一个重要统计特征,对确定结构作用风力大小具有决定性的意义。在工程结构设计中,年最大平均风速的概率分布对确定基本风速是不可缺少的重要信息^[1]。例如,在我国荷载规范和国际可靠度与安全度联合委员会发布的概率模式规范中,均在给定(假定)年最大平均风速的分布类型的基础上,给出了基本风速和基本风压的统计量^[2,3]。针对不同地区或不同工程场址的需要,研究者们采用多种方法进行了风速资料的统计分析,例如文献[4]~[6]的工作。这些工作的基本思想一般是首先假定一个分布类型,进行该分布类型情况下的统计量分析,并采用某种准则进行拟合优度检验。显然,有限的规则分布类型难以满足实际应用的需要,为此,一些学者进行了统计分布与统计方法的改进^[7,8]。然而,在

收稿日期:2006-05-08; 修订日期:2006-08-10

基金项目:国家自然科学基金委创新研究群体项目(50325803)

作者简介:李杰(1957-),男,河南沈丘人,教授,主要从事地震工程、生命线工程抗灾研究。

总体上,现有的方法均不能逾越给定先验分布类型、然后进行拟合优度检验这一基本思想的藩篱。这本质上是给风速分布强加了某种限制条件。并且,在实际统计中,人们发现,要找到能够通过较为严格界限的拟合优度检验的分布类型是颇为困难的。近年来,我们发展了基于密度演化理论的随机变量概率密度函数估计方法^[9,10]。在该方法中,不需要进行概率分布或概率密度函数类型的先验假定,可以直接根据基本数据给出概率密度函数的估计。本文采用这一方法,根据我国大陆地区 52 a 中 740 个站点的实测风速数据,进行了我国大陆地区年最大平均风速的概率密度函数统计,并给出了年最大平均风速的均值和标准差等值线图。

1 风速基本资料

中国气象局国家气象中心气象资料室提供了我国大陆地区 740 个地面国际交换站 1951 - 2002 年 52 a 实测风速资料(共计 1 184.35 万组数据)。这些风速资料是每日 4 次定时 2 min 平均风速的平均值 U_2 。考虑到在结构设计中主要采用 10 min 平均风速作为基本依据,需要首先将 4 次定时 2 min 平均风速的日平均值 U_2 转化为自记 10 min 平均风速 U_{10} ,亦即进行双重次时换算。

通常,人们认为年最大平均风速在统计中具有重要意义,表 1 中列出了中国大陆地区年最大平均 2 min 风速与年最大自记 10 min 平均风速的换算公式^[11,12]。本文应用该表中的经验公式,完成实测平均风速资料的次时换算,得到了中国大陆地区年最大自记 10 min 平均风速转换数据,并将其作为统计分析的基本数据。

表 1 自记 10 min 与 4 次定时 2 min 风速转换关系式

Table 1 Relationship between different wind speeds

编号	地 区	回归方程式	应用范围
1	华 北	$U_{10} = 0.822U_2 + 7.82$	华北(内蒙古、山西、河北、北京、天津)
2	东 北	$U_{10} = 1.04U_2 + 2.20$	东北(黑龙江、吉林、辽宁)
3	西 北	$U_{10} = 1.004U_2 + 1.57$	西北(新疆、甘肃、宁夏、山西)和西藏
4	青 海	$U_{10} = 0.845U_2 + 5.21$	青海
5	云 贵	$U_{10} = 0.751U_2 + 6.17$	云南、贵州
6	四 川	$U_{10} = 1.24U_2$	四川
7	湖 北	$U_{10} = 0.732U_2 + 7.0$	湖北、湖南、江西
8	湖 南	$U_{10} = 0.68U_2 + 9.54$	湖北、湖南、江西
9	广 东	$U_{10} = 1.03U_2 + 4.15$	广东、广西、福建、台湾
10	江 苏	$U_{10} = 0.78U_2 + 8.41$	上海、江苏、浙江、安徽
11	山 东	$U_{10} = 1.03U_2 + 2.19$	山东
12	渤海面	$U_{10} = 0.93U_2 + 0.35$	海面

2 年最大平均风速的概率密度函数估计

2.1 经验分布函数与直方图估计

对于一组实测得到的样本 $(z_1, z_2, \dots, z_{n_{sp}})$,其中 n_{sp} 为样本容量,经验分布函数是对于所考察随机变量真实分布的稳健逼近,它是总体分布函数的相容估计。事实上,根据经验分布函数的定义,可以在已知经验分布函数的情况下完全反映原始的样本数据,即:在经验分布函数里,包含了样本的所有信息。正是基于这一背景, $K-S$ 检验以经验分布函数作为基准,以待检验分布函数与经验分布函数的偏差作为拟合优度的度量^[13,14]。

若样本 $(z_1, z_2, \dots, z_{n_{sp}})$ 由小到大的次序统计量是 $(z_{(1)}, z_{(2)}, \dots, z_{(n_{sp})})$,则经验分布函数为

$$F_{n_{sp}}(x) = \frac{1}{n_{sp}} \sum_{j=1}^{n_{sp}} u(x - z_{(j)}) \quad (1)$$

其中 $u(\cdot)$ 为单位阶跃函数。

在统计学中,有诸多方法可以在直观上给出概率密度函数的粗略估计^[14],其中最常用的是绘出数据的频率直方图。如所周知,由于直方图的形状往往随着划分区间的大小和数目而产生很大的差异,因此,它仅能提供一个相当粗疏的估计。事实上,这一问题与 X^2 检验中遇到的问题本质上是相同的。虽然如此,统计

经验表明,当选取的分段数目适当的时候,直方图可以大致反映概率密度函数的粗略形状。鉴于此,在本文中给出密度函数估计的同时,亦绘出频率直方图以资参照。根据已有大量统计分析给出的经验,本文中的直方图区间划分数目 k 取为^[14]

$$k = 1 + \log_2 n_{sp} \quad (2)$$

例如,当样本量 $n_{sp} = 52$ 时,划分区间数目可取 6。

2.2 年最大平均风速的密度函数估计

2.2.1 虚拟随机过程与密度演化方法

为了获取一组数据所来自的总体 Z 的概率密度函数 $p_Z(z)$, 可以构造一个虚拟随机过程

$$X(\tau) = \phi(Z, \tau) \quad (3)$$

其中 τ 为虚拟时间参数。原则上,式(3)的函数形式只需使得 Z 为 $X(\tau)$ 在 $\tau = \tau_c$ 时刻的截面随机变量即可,亦即

$$Z = X(\tau) |_{\tau=\tau_c} = \phi(Z, \tau_c) \quad (4)$$

通常,为方便起见,式(3)的初始值可以取为

$$X(\tau) |_{\tau=0} = \phi(Z, \tau = 0) = 0 \quad (5)$$

根据计算经验,式(3)取如下形式具有较好的效果:

$$\phi(Z, \tau) = Z \sin(\omega\tau), \tau_c = 1 \quad (6)$$

显然,当 $\omega = 2.5\pi$ 时,式(6)满足式(4), (5)中给出的条件。

对式(3)关于 τ 求导可得 $X(\tau)$ 的速率

$$\dot{X}(\tau) = \dot{\phi}(Z, \tau) \quad (7)$$

根据密度演化理论, (X, Z) 构成的随机系统概率守恒,因而可以导出其联合概率密度函数 $p_{XZ}(x, z, \tau)$ 满足的广义密度演化方程^[9,10]

$$\frac{\partial p_{XZ}(x, z, \tau)}{\partial \tau} + \dot{\phi}(z, \tau) \frac{\partial p_{XZ}(x, z, \tau)}{\partial x} = 0 \quad (8)$$

由式(5)可知式(8)的初始条件为

$$p_{XZ}(x, z, \tau) |_{\tau=0} = \delta(x) p_Z(z) \quad (9)$$

边界条件可取 $p_{XZ}(x, z, \tau) |_{x \rightarrow \pm\infty} = 0$ 。

求解偏微分方程初-边值问题式(8)~(9)可以给出联合概率密度函数 $p_{XZ}(x, z, \tau)$, 进而积分可得

$$p_X(x, \tau) = \int_{-\infty}^{\infty} p_{XZ}(x, z, \tau) dz \quad (10)$$

注意到式(4), 即可获得随机变量 Z 的概率密度函数

$$p_Z(z) = p_X(x = z, \tau) |_{\tau=\tau_c} \quad (11)$$

2.2.2 实施方法

若有实测样本 $(z_1, z_2, \dots, z_{n_{sp}})$, 欲以此为基本数据估计总体的概率密度函数。首先将方程(8)关于 z 离散, 并分别取已经获取的各个样本值, 可得到一组偏微分方程

$$\frac{\partial p_{XZ}(x, z_j, \tau)}{\partial \tau} + \dot{\phi}(z_j, \tau) \frac{\partial p_{XZ}(x, z_j, \tau)}{\partial x} = 0 \quad (j = 1, 2, \dots, n_{sp}) \quad (12)$$

在通常情况下, 可以认为数据实测过程是独立重复进行的, 因此, 获取各个样本值的概率相同, 故初始条件(9)可取为

$$p_{XZ}(x, z_j, \tau) |_{\tau=0} = \delta(x) \frac{1}{n_{sp} |\Delta z|} \quad (13)$$

为方便计, 权取 $|\Delta Z| = (z_{\max} - z_{\min}) / n_{sp}$ 。

在初始条件式(13)下求解方程式(12), 可获得解答 $p_{XZ}(x, z, \tau)$ ($j = 1, 2, \dots, n_{sp}$), 进而, 对式(10)进行数值积分可给出

$$p_X(x, \tau) = \sum_{j=1}^{n_{sp}} p_{XZ}(x, z_j, \tau) |\Delta z| \quad (14)$$

结合式(13),(14)不难看出, $|\Delta z|$ 的选取事实上对最终结果不产生任何影响, 因此, 在实际分析中, 可直接取 $|\Delta z| = 1$ 。

式(12)可以采用差分格式求解。根据计算经验, 在概率密度函数估计中, 采用单边差分格式与具有 TVD 性质的修正的 Lax - Wendroff 格式构成的加权组合格式可取得较好的效果。具体实施步骤见图 1。

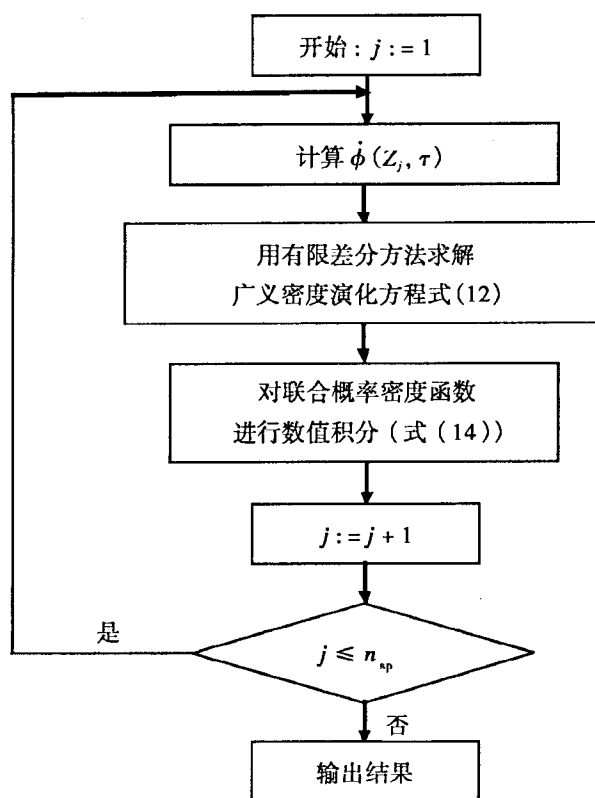


图 1 概率密度函数估计的流程图

Fig. 1 Flowchart of estimation of probability density function

2.2.3 年最大平均风速的密度函数估计

如前所述, 根据实测数据, 可以获取每一站点 52 a 中的每年实测最大平均风速, 亦即具有 52 个年最大平均风速的样本, 因而, 可以直接以这些数据(经双重次时转换以后)作为第 2.2.2 节中的样本 $(z_1, z_2, \dots, z_{n_{sp}})$, 进而可以获得每一站点年最大平均风速的概率密度函数估计。在这一概率信息的基础上, 获取结构设计基准期内的最大风速分布以及具有一定重现期的设计风速是相当方便的^[13]。

如前指出, 以往进行风速分布的统计, 往往着重于进行统计检验, 以某一预设的分布形式进行拟合优度检验, 并相应进行基本参数的统计^[1,2]。由于实际风速分布的复杂性, 特别是当采用比较严格的检验标准的时候, 往往难以找到符合统计检验要求的分布形式。采用基于密度演化理论的概率密度函数估计, 可以直接给出概率密度函数估计的数值解答, 尽管可能不是某一种特殊的分布, 但就实测获得的数据而言, 它是更为真实合理的分布。这一分布与经验分布函数的最大偏差

$$d = \sup_{-\infty < x < \infty} | \hat{F}(x) - F_{n_{sp}}(x) | \quad (15)$$

可以作为概率密度函数估计的精度度量, 这里,

$$\hat{F}(x) = \int_{-\infty}^x p_z(z) dz = \int_{-\infty}^x p_x(x, \tau_c) dx \quad (16)$$

是密度演化方法给出的概率分布函数估计, $F_{n_{sp}}(x)$ 是式(1)中定义的经验分布函数。

显然, 式(15)给出的偏差事实上就是统计检验中的 $K-S$ 检验统计量。所不同者, 在 $K-S$ 检验中, 是先设定分布函数的形式, 而在本文中, 不需人为设定分布函数的形式。

3 典型站点的年最大平均风速分布统计结果

3.1 典型站点的年最大平均风速概率密度函数估计

采用第 2.2.3 节中所述的方法,进行了 740 个国际交换站 1951 - 2002 年 52 a 的实测年最大平均风速的概率密度函数估计。这里仅选择其中两个站点的分析结果作为示例。

图 2、图 3 分别是北京站和成都站年最大平均风速的概率密度函数与概率分布函数估计结果。为了比较,两图中的(a)还同时绘出了具有相同均值与标准差的极值 I 型分布的概率密度函数以及分段数目由式(2)确定的数据频率直方图。从图中不难看到,直接采用密度演化方法计算获得的概率密度函数曲线在趋势上与直方图最为符合。两图的(b)中同时绘出了经验分布函数、密度演化方法计算获得的概率分布函数(式(16))和极值 I 型分布函数,从中可见,密度演化方法计算获得的概率分布函数与经验分布函数的偏差较小,在对可靠度问题影响较大的尾部更是如此。

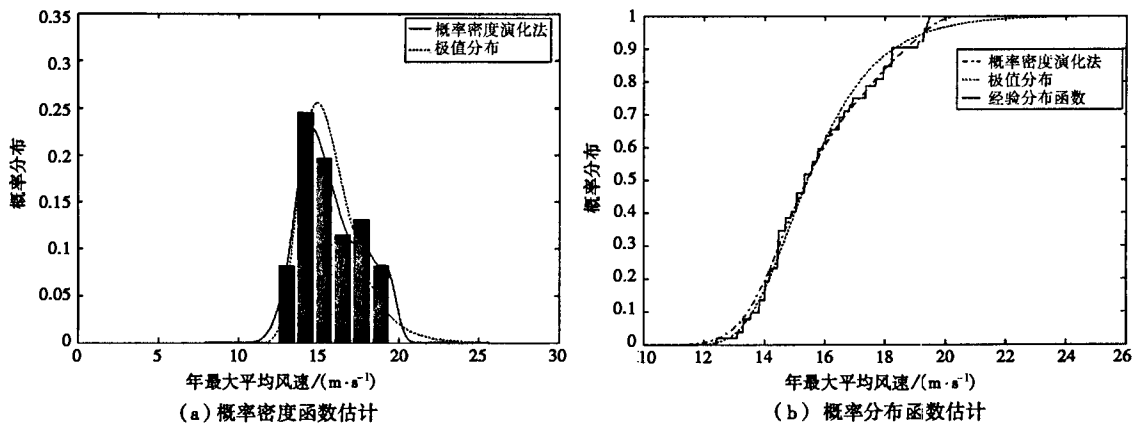


图 2 北京站年最大平均风速概率密度函数统计

Fig. 2 PDF of yearly maximum average wind speed at Beijing Station

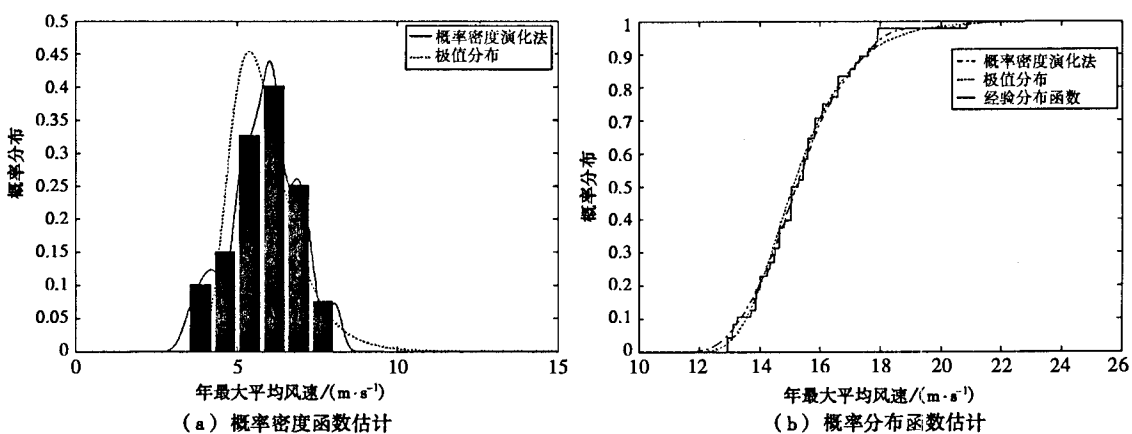


图 3 成都站年最大平均风速概率密度函数统计

Fig. 3 PDF of yearly maximum average wind speed at Chengdu Station

3.2 年最大平均风速均值、标准差统计

在给出概率密度函数分布的同时,可以容易地获得年最大平均风速的均值与标准差。图 4 为我国大陆地区年最大平均风速均值的等值线分布,图 5 是年最大平均风速标准差的等值线分布图。两图中的底图,即中国国界线图,下载自国家基础地理信息系统网站(<http://nfgis.nsd.gov.cn/>),且已经该网站登记,该国家基础地理数据版权归国家测绘局所有。据此,若需要任意地区或工程场址年最大平均风速的均值与标准差,则可以直接从图中查取相关的数据。



图 4 我国大陆地区(不包括港、台和领海海域)年最大平均风速的均值等值线分布图
Fig. 4 Contour of mean value of yearly maximum average wind speed in mainland of China



图 5 我国大陆地区(不包括港、台和领海海域)年最大平均风速的标准差等值线分布图
Fig. 5 Contour of variance of yearly maximum average wind speed in mainland of China

4 结论与讨论

采用基于密度演化理论的概率密度函数估计方法,分析了我国大陆地区 740 个地面国际交换站 1951 - 2002 年共 52 a 的基本风速实测数据。给出了各个交换站年最大平均风速的概率密度函数估计与概率分布函数估计。与传统的基于拟合优度检验的规则分布函数的统计方法比较,本文方法不需要概率分布类型或密度函数类型先验分布的假定。研究表明:本文统计给出的结果是可信的。

限于篇幅,本文仅给出了若干站点年最大平均风速统计结果的简单示例。为了工程实用的需要,可以将各个站点计算结果的概率密度函数编制成图册,并同时数值化的概率密度函数编制成数据表格以供查阅。

感谢中国气象局国家气象中心气象资料室提供我国大陆地区 740 个地面国际交换站 1951 - 2002 年 52 a 实测风速资料。

参考文献:

- [1] 张相庭. 工程结构风荷载理论及抗风计算手册[M]. 上海:同济大学出版社,1990.
- [2] 陈英俊,于希哲. 风荷载计算[M]. 北京:中国铁道出版社,1998.
- [3] JCSS, Probabilistic Model Code, Part 2: Loads[S].
- [4] 张淮水,刘安国,宋珊,吴术礼. 海面风的特征分析[S]. 青岛海洋大学学报,1989,19(2):48 - 54.
- [5] 陈伟,王海良,项海帆. 风速统计特性研究探讨[J]. 石家庄铁道学院学报,1999,12(2):51 - 55.
- [6] 陈正洪,向玉春,杨宏青,等. 深圳湾公路大桥设计风速的推算[J]. 应用气象学报,2004,15(2):227 - 233.
- [7] 张柄华,徐志春. 一种新的最大风速分布函数[J]. 同济大学学报,1989,17(1):81 - 87.
- [8] 董安正,赵国藩. 建筑结构基本风压统计方法的改进[J]. 工业建筑,2003,33(4):75 - 77.
- [9] 陈建兵,李杰. 随机结构动力反应的极值分布[J]. 振动工程学报,2004,17(4):382 - 387.
- [10] 陈建兵,李杰. 密度演化方法在随机变量概率分布估计中的应用研究[J]. 同济大学学报,2006,34(4):433 - 437.
- [11] 张相庭. 结构风压和风振计算[M]. 上海:同济大学出版社,1985.
- [12] 朱瑞兆. 风压计算的研究[M]. 北京:科学出版社,1976.
- [13] Ang AH S, Tang W H. Probability Concepts in Engineering Planning and Design[M]. Indianapolis:John Wiley & Sons, 1975.
- [14] Gentle J E. Elements of Computational Statistics[M]. New York:Springer Science + Business Media, Inc., 2002.