

考虑风向的极值风速估计的 Cook 方法改进

全涌, 刘磊, 顾明

(同济大学 土木工程防灾国家重点实验室, 上海 200092)

摘要: 基于对 Cook 方法理论缺陷的分析, 提出一种改进方法, 使得考虑风向的极值风速估计的概率模型由多元随机变量更准确地转化为一维随机变量. 以南京地区的极值风速为例, 对改进方法与 Cook 方法计算结果进行对比. 结果表明: 改进方法对极值风速概率分布形式无要求, 也不受 Cook 方法所需某些假设的限制, 从概率的角度来说是精确的方法.

关键词: 极值风速; 风向相关性; 概率模型

中图分类号: TU312.1

文献标志码: A

Improvement of Cook Method Considering Directional Extreme Wind Speed

QUAN Yong, LIU Lei, GU Ming

(State Key Laboratory for Disaster Reduction in Civil Engineering, Tongji University, Shanghai 200092, China)

Abstract: Based on the analysis of the limitations of the Cook method, an improved method was proposed. The method changed the multivariate probabilistic model of directional extreme wind speeds into an univariate model. Taking the extreme wind speeds in Nanjing area as example, the results of the method and Cook method are compared. Since the method has no limitations in the probability distribution and the assumptions of the Cook method, it has more accuracy in the probability way.

Key words: extreme wind speed; correlation of wind direction; probability model

风荷载的合理估计问题必须考虑风速和气动力系数的风向性^[1]. 这个问题的解决需要精确分析一个与风速极值和气动力系数概率密度函数有关的卷积. 由于两者的各方向概率分布差异很大, 所以不存在一般解. 考虑风向极值风速的准确估计是合理计算结构风荷载的重要基础. 为了解决这一问题, 学者

们提出了很多方法. Cook^[2]通过分析英国地区的极值风速数据, 认为各风向极值风速随超越概率的变化基本遵循同一比例, 即极值风速玫瑰图随超越概率的变化或放大或缩小, 但基本形状不变. 基于这一假设, 将考虑风向的极值风速问题由多元极值问题简化为一元极值问题来求解. 但并非所有地区的极值风速均满足这一假定, 此时应用 Cook 方法可能会使极值风速估算出现偏差. Haraguchi 等^[3]将二元 Logistic 概率模型推广到多元, 以探究各风向的极值分布与全风向的极值分布的关系. Itoi 等^[4]通过分析日本地区的风速数据肯定了考虑风向之间相关性的重要性. 然而这些方法都是从拟合风速数据到多元极值概率模型的角度来解决风向相关性问题的. 实际的风速数据是否会符合所采用的概率模型是较难考证的, 并且因涉及到多个参数的估计, 计算过程会非常复杂. 另外, Itoi 等^[4]和 Kasperski^[5]认为 Cook 方法没有将风向之间的相关性考虑在内, 这个观点有待进一步商榷.

本文对 Cook 方法^[2]进行改进, 弥补其理论缺陷, 并给出适用范围更广泛的方法.

1 方法概述

1.1 拟合概率模型方法

先假设考虑风向的极值风速概率分布函数已经得到, 表示为 $F_V(v_{d_1}, v_{d_2}, \dots, v_{d_n})$, 则

$$F_V(v_{d_1}, v_{d_2}, \dots, v_{d_n}) = \Pr(V_{d_1} < v_{d_1}, V_{d_2} < v_{d_2}, \dots, V_{d_n} < v_{d_n}) \quad (1)$$

式中: V_{d_i} 表示随机变量, 代表第 i 个风向的极值风速; v_{d_i} 表示第 i 个风向的极值风速取值. 如果需要计算各风向下 R 年重现期极值风速, 则

$$F_V(v_{d_1}, v_{d_2}, \dots, v_{d_n}) = \Pr(V_{d_1} < v_{d_1}, V_{d_2} < v_{d_2}, \dots,$$

收稿日期: 2014-05-07

基金项目: 国家自然科学基金(51278367); 土木工程防灾国家重点实验室自主研究课题基金(SLDRCE10-B-03); 中央高校基本科研业务费专项资金

第一作者: 全涌(1971—), 男, 副教授, 工学博士, 主要研究方向为结构抗风. E-mail: quanyong@tongji.edu.cn

$$V_{d_n} < v_{d_n}) = 1 - 1/R \quad (2)$$

将上式转化为多元函数求解问题,其解即为分位数线.显然,会有多条分位数线满足式(2),这是因为这里只有一个等式却有 n 个变量,还需要别的约束条件才能惟一确定分位数线.这个条件就是各个风向上的超越概率相等,即

$$F_{V_{d_i}}(V_{d_i}) = \Pr\{V_{d_i} < v_{d_i}\} = 1 - p^*, \quad i = 1, 2, \dots, n \quad (3)$$

式中: $F_{V_{d_i}}$ 表示第 i 个风向上极值风速的概率分布,也是 F_V 关于第 i 个分量的边缘分布函数; p^* 表示各个风向上共同的超越概率.

这样式(2)和(3)组成一组含有 $v_{d_1}, v_{d_2}, \dots, v_{d_n}$, p^* 的 $(n+1)$ 个变量的 $(n+1)$ 组多元函数组,它在理论上有着惟一的解.但是,这个多元非线性方程组的求解是困难的,并且由 F_V 得到 $F_{V_{d_i}} (i=1, 2, \dots, n)$ 需涉及多元函数的积分问题,求解也很困难.

1.2 Cook 方法

Cook 方法从边缘概率分布出发,将上述多元极值的联合概率分布转化为一元极值分布,避免了多元函数求解的问题.

假定各个风向下极值风速和全风向的极值风速均服从极值 I 型分布.为方便下面的表述,以 $F_i (i=1, 2, \dots, n)$ 表示 i 风向的风速极值分布, F^* 表示全风向的极值风速分布. Cook 方法的具体步骤如下:

(1) 分别求得各风向的 R 年重现期的风速极值 v_{d_i} 和全风向的风速极值 v_d^* .

$$F_i(v_{d_i}) = 1 - 1/R, \quad i = 1, 2, \dots, n$$

$$F^*(v_d^*) = 1 - 1/R$$

(2) 计算出各个风向下 R 年重现期的风速极值与全风向时的风速极值之比.

$$k_i = v_{d_i}^* / v_d^*, \quad i = 1, 2, \dots, n$$

(3) 将 i 风向下的风速样本 V_i 乘以 k_i .

$$U_i = V_i k_i, \quad i = 1, 2, \dots, n$$

(4) 将经过转换后的样本 U_i 选取新的极值样本,重新拟合为全风向极值风速的极值 I 型分布 F^{**} , 然后计算 v_d^{**} .

$$F^{**}(v_d^{**}) = 1 - 1/R$$

(5) 计算比例因子.

$$k = v_d^{**} / v_d^*$$

(6) 计算 R 年重现期各风向下风速极值 $v_{d_i}^*$.

$$v_{d_i}^* = v_d^* k, \quad i = 1, 2, \dots, n$$

1.3 对 Cook 方法的改进

1.3.1 方法的原理

为了方便讨论,将式(2)和(3)写成如下方程组

的形式:

$$\begin{cases} F_V(v_{d_1}, v_{d_2}, \dots, v_{d_n}) = \Pr(V_{d_1} < v_{d_1}, V_{d_2} < v_{d_2}, \dots, \\ V_{d_n} < v_{d_n}) = 1 - 1/R \\ F_{V_{d_1}}(v_{d_1}) = \Pr\{V_{d_1} < v_{d_1}\} = 1 - p^* \\ F_{V_{d_2}}(v_{d_2}) = \Pr\{V_{d_2} < v_{d_2}\} = 1 - p^* \\ \vdots \\ F_{V_{d_n}}(v_{d_n}) = \Pr\{V_{d_n} < v_{d_n}\} = 1 - p^* \end{cases} \quad (4)$$

式(4)中第 1 个等式,即 n 个风向上极值风速的联合概率分布 F_V , 很难直接给出的.但其余等式,即各个风向上的极值风速分布 $F_{V_{d_i}}$, 可以直接由风速风向的长期观测数据方便地统计给出.另外,全风向的极值风速分布 F_{V_d} 也可以直接由风速风向观测数据方便地统计给出.

$$F_{V_d}(v_d) = \Pr\{\max(V_{d_1}, V_{d_2}, \dots, V_{d_n}) < v_d\} = F_V(v_d, v_d, \dots, v_d) \quad (5)$$

设

$$U_{d_i} = V_{d_i} / k_i, \quad i = 1, 2, \dots, n \quad (6)$$

其中

$$k_i = v_{d_i} / v_d, \quad i = 1, 2, \dots, n \quad (7)$$

则 n 个风向极值风速的联合概率分布 F_V 可变形为

$$\begin{aligned} F_V(v_{d_1}, v_{d_2}, \dots, v_{d_n}) &= \Pr(V_{d_1} < v_{d_1}, V_{d_2} < v_{d_2}, \dots, V_{d_n} < v_{d_n}) = \\ &= \Pr(V_{d_1} < k_1 v_d, V_{d_2} < k_2 v_d, \dots, V_{d_n} < k_n v_d) = \\ &= \Pr(U_{d_1} < v_d, U_{d_2} < v_d, \dots, U_{d_n} < v_d) = \\ &= F_U(v_d, v_d, \dots, v_d) = \\ &= \Pr\{\max(U_{d_1}, U_{d_2}, \dots, U_{d_n}) < v_d\} = \\ &= F_{U_d}(v_d) \end{aligned} \quad (8)$$

式中: F_U 为新随机变量 $U_{d_1}, U_{d_2}, \dots, U_{d_n}$ 的联合分布函数; F_{U_d} 为 n 个新随机变量最大值的分布函数,即 $U_d = \max(U_{d_1}, U_{d_2}, \dots, U_{d_n})$; v_d 为全风向 R 年重现期下的极值风速,是已知量.

从式(8)可以看出,联合概率分布可以与一组新随机变量的最大值的分布函数建立等式关系,所以如果能确定新随机变量的最大值所服从的分布,那么相当于知道了式(4)中的 n 个风向的联合概率分布,这就将多元极值问题转换成一元极值问题.

从式(6)可知,如果 $k_i (i=1, 2, \dots, n)$ 值确定,那么新随机变量与原随机变量存在确定的比例关系.如果将 i 风向的风速数据分别除以 k_i , 那么经过转化后的样本将服从分布函数 F_U 的分布,进而全风向极值样本也将服从分布函数 F_{U_d} 的分布.由于已经获得了原随机变量的样本,也就是风速风向数据,因

此相当于获得了这组新随机变量最大值的样本。

但是,考虑风向的联合分布函数的类型未知,很难证明转化后的样本的年极值与原样本的年极值服从同一类型的概率分布。然而,Cook方法中默认新样本与原样本一样,依然服从极值型分布。基于下文给出的南京地区的数据分析结果,如图1和2所示,本文认为其假定是基本合理的。虽然分布类型是模糊的,但当样本量足够时,至少可以用经验分布函数来代替真实的分布函数。从这个角度来说, F_{U_d} 是可以得到的。

根据以上分析,式(4)可转化为

$$\begin{cases} F_{U_d}(v_d) = \Pr\{U_d < v_d\} = 1 - 1/R \\ F_{V_{d_1}}(v_{d_1}) = \Pr\{V_{d_1} < v_{d_1}\} = 1 - p^* \\ F_{V_{d_2}}(v_{d_2}) = \Pr\{V_{d_2} < v_{d_2}\} = 1 - p^* \\ \vdots \\ F_{V_{d_n}}(v_{d_n}) = \Pr\{V_{d_n} < v_{d_n}\} = 1 - p^* \end{cases} \quad (9)$$

上式仍然有 $(n+1)$ 个变量的 $(n+1)$ 组多元方程,只是第1个等式的形式由包含 n 个变量 U_d 的函数方程式转化为与变量的分布函数有关的等式。这个分布函数实际上是 p^* 的函数,它满足这样的条件: $F_{U_d}(v_d)$ 的值随 p^* 的值增大而单调增大。所以,可以通过调试 p^* 的值,使得 $F_{U_d}(v_d) \approx 1 - 1/R$,此时的 $v_{d_i} (i=1, 2, \dots, n)$ 就是式(8)的解,也是 R 年重现期下各风向的风速极值。

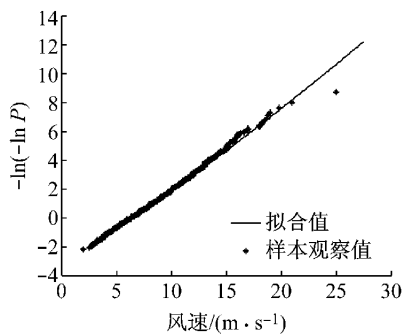


图1 原风速样本分布

Fig.1 Distribution of original samples of wind speed

在调试 p^* 值时,可以先应用下面的理论确定 p^* 的范围。若假定各风向的极值风速完全相互独立,则对应于 R 年重现期下全风向极值风速的 p^* 值为 $(1 - 1/R)^n \approx 1 - 1/(Rn)$ 。若假定各风向的极值风速完全相关,则对应于 R 年重现期下全风向极值风速的 p^* 值为 $(1 - 1/R)$ 。在求解过程中,由于 $F_{U_d}(v_d)$ 是 p^* 的单调增函数,因此可以在区间上 $(1 -$

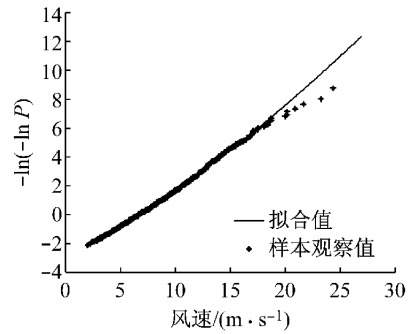


图2 新样本的分布($p^* = 0.996\ 96$)

Fig.2 Distribution of new samples($p^* = 0.996\ 96$)

$1/(Rn), 1 - 1/R)$ 反复采用二分法以逼近 p^* 的准确值。

1.3.2 改进方法的步骤

(1) 拟合各风向下极值风速所服从的分布 $F_{V_{d_i}}$ ($i=1, 2, \dots, n$)和全风向极值风速所服从的分布 F_{V_d} 。当样本量足够时,也可以用经验分布函数。

(2) 由等式 $F_{V_d}(v_d) = 1 - 1/R$ 得到 v_d 。

(3) p^* 变化的上下限初值分别为 $p_{up} = 1 - 1/(Rn)$, $p_{dn} = 1 - 1/R$,即 $p^* \in (p_{dn}, p_{up})$ 。

(4) p^* 值取其变化区间中点。

(5) 将 p^* 代入式(3)得到 $v_{d_i} (i=1, 2, \dots, n)$,进而通过式(7)计算得到 $k_i (i=1, 2, \dots, n)$ 。

(6) 由式(6)得到新样本,并拟合全风向的新样本极值风速所服从的分布 F_{U_d} 。

(7) 计算 $F_{U_d}(v_d)$,当 $\Delta = |F_{U_d}(v_d) - (1 - 1/R)|$ 小于需要的精度时,结束循环,进入下一步。否则,当 $F_{U_d}(v_d) < 1 - 1/R$ 时, p^* 变化区间的下限,更新 p^* 的当前值,即 $p_{dn} = p^*$;当 $F_{U_d}(v_d) > 1 - 1/R$ 时, p^* 变化区间的上限,更新 p^* 的当前值,即 $p_{up} = p^*$ 。转入上述步骤(4)。

(8) $v_{d_i} (i=1, 2, \dots, n)$ 即为所需要的结果。

2 Cook方法与改进方法的比较

本文继承了Cook方法中将多元极值问题转化成一元极值问题的思路,但在求解上与原方法有较大的不同。Cook通过上文中提到的假设将式(9)中随 p^* 变化的分布函数 F_{U_d} 转化为固定的分布函数,然后将问题转化为对分位数的求解。这一方法的假设通常是不成立的(关于“假设通常不成立”的说明见后文),因此这种方法算出的值与真实值之间存在偏差。

实际上,Cook方法中的 F_{U_d} 是当 $p^* = 1 - 1/R$

时新样本所服从的分布函数. 从式(8)分析可知, 当分位点为全风向 R 年一遇的风速极值时, 这个分布函数的值等于分位数线为各方向 R 年一遇的风速极值时的概率, 上述分析可写为

$$1 - 1/R > F_{U_d}(v_d^*) = F_V(v_{d_1}, v_{d_2}, \dots, v_{d_n}) \quad (10)$$

毫无疑问, 此时的超越概率是远大于 $1/R$ 的. 就本文的算例 $F_{U_d}(v_d^*) \approx 0.8334$, 也就是说此时总的超越概率 $p \approx 0.1666$. 按照 Cook 方法的假设, 分位数的形状不发生变化, 只将比例扩大 k 倍, 则式(10)变为

$$1 - 1/R = F_{U_d}(kv_d^*) = F_V(kv_{d_1}, kv_{d_2}, \dots, kv_{d_n}) \quad (11)$$

从第 1.3.1 节的分析可知, 当式(11)中第 1 个等号成立时, 第 2 个等号不一定成立. 另外, 总的超越概率从 $p \approx 0.1666$ 到 $p \approx 0.0200$ 变化时, 样本转化后的分布函数 F_{U_d} 也可能发生变化. 相比之下, 本文的方法是通过试算的方式寻找满足要求的分布函数, 期间没有用到特殊的假设和特定的分布函数, 当样本量足够时, 甚至可以用经验分布函数来代替.

改进的方法需要的运算量比 Cook 方法稍有增大, 但是在计算机发达的今天, 这已经不是个难题.

3 算例

本算例采用了中国南京地区的风速风向数据, 截取的时间段为 1960 年 1 月 1 日到 2012 年 1 月 1 日. 其中, 删掉了 80 个没有风速数值记录的数据和 15 个没有风向的数据.

首先对数据选取风暴, 选取的方式在一定程度上类似于 Cook 独立风暴法中所用的方法. 如图 3 所示, 两个相邻的圆点之间为一个风暴, 五角星为这次风暴的最大值. 当不考虑风向时, 每个风暴只提取最大值作为样本, 转化后的新样本极值的选取也是如此; 考虑风向时, 则选取每个风暴中这个风向下数据的最大值. 将选取的样本均拟合成新的分布函数: $f(x) = (1 - \exp(-v/v_0)^w)^q$, 其中 v_0, w, q 为待估参数, v 为随机变量, 表示风速.

值得注意的是, 上述的分布函数先用来拟合风暴的母体分布, 然后推算得到风暴极值的分布函数. 如果直接用来拟合样本的极值, 它会与 Harris^[6] 提到的 penultimate 极值分布函数类似. 由于本文并不是讨论极值风速的分布问题, 所以对于上述样本选取的合理性和分布函数的适用性并不多做讨论. 使用新的分布函数对选取样本的拟合情况如图 1 和 2

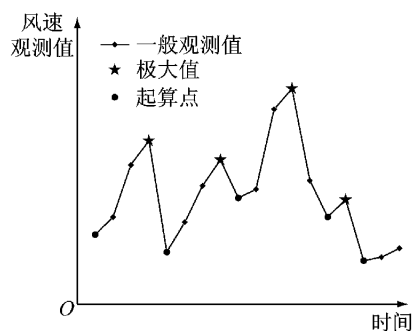


图 3 独立风暴的采样方法

Fig. 3 Sampling method of independent storm

所示.

以下分别运用 Cook 方法和改进方法求得 50 年重现期下各方向的风速极值. 与 Cook 方法^[2]所不同的是: 本算例中的 Cook 方法仅借用了如式(10)和(11)所示的假设, 而极值风速的分布函数与改进方法相同. 通过数据分析, 当式(9)中的 $p^* = 0.99696$ 时, 可以得到总的超越概率 $p = 0.020017$. 将上述 2 种方法所得的各风向的风速极值除以全风向风速极值可得到极值风速的风向折减因子, 如图 4 所示. 通过比较可以看出, 两种方法在 270° 风向 (西风方向) 的差别较明显, 差异率达到 9%.

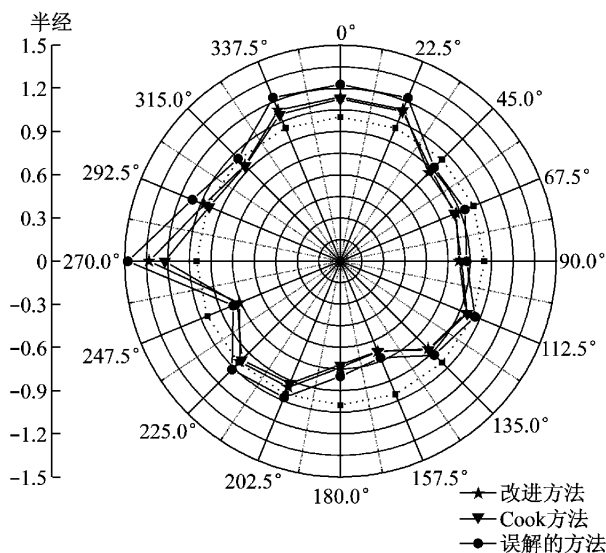


图 4 极值风速的风向折减因子对比

Fig. 4 Comparison of wind directional factors

另外, Cook 方法通常被认为将各个风向的极值风速看成是独立的, 且可将全风向的超越概率均分到每个方向上进行处理, 即式(4)中的 $p^* = 1 - 0.02/16 = 0.99875$, 这样得到的结果会使得全风向的超越概率小于 0.0200, 从而使得估计结果非常保守. 但实际上这并不是 Cook 方法的本意. 就本文算

(下转第 204 页)