

# 基于 Monte Carlo 法的极值分布类型及其参数估计方法比较

卢安平, 赵林, 郭增伟, 葛耀君

(同济大学 土木工程防灾国家重点实验室, 200092 上海)

**摘要:** 为探讨极值概率分布模型(极值I型、II型、III型分布)在工程场地重现期内极值风速预测中的适用性,以工程上常用的风速分布模型为依据,采用 Monte Carlo 数值模拟法分别产生服从指数分布、正态分布、瑞利分布、威布尔分布、对数正态分布、广义极值分布的伪风速母样. 基于以上伪风速母样分别使用极大似然参数估计法和概率矩法对极值I型、II型、III型分布模型中的参数进行估计,并与伪风速母样的理论极值分布进行对比分析. 结果表明:极大似然参数估计法适用性较强,参数估计精度较高;重现期内极值风速的估算对风速母样分布类型敏感,服从不同分布的风速母样,应有针对性选用极值分布类型估算重现期内极值风速.

**关键词:** Monte Carlo 法; 伪风速母样; 极值分布类型; 参数估计; 极大似然参数估计法; 极值风速

**中图分类号:** TU973.32; O212.1      **文献标志码:** A      **文章编号:** 0367-6234(2013)02-0088-08

## A comparative study of extreme value distribution and parameter estimation based on the Monte Carlo method

LU Anping, ZHAO Lin, GUO Zengwei, GE Yaojun

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

**Abstract:** To explore the applicability of extreme value distribution models, known as type I, II and III extreme value distribution, to the extreme wind velocity evaluations of project site during the return period, and according to the wind speed distribution model commonly used in engineering fields, Monte Carlo simulation method is used to produce the pseudo wind speed samples, respectively fitting exponential distribution, normal distribution, Rayleigh distribution, Weibull distribution, lognormal distribution and the generalized extreme value (GEV) distribution. The maximum likelihood parameter estimation method and probability weighted moment method are used to estimate the parameters of the type I, II and III extreme value distributions. A comparative study of different parameter estimation methods for every extreme value distribution has been done by comparing with the theoretical extreme value distribution of the pseudo wind speed. The analysis shows that the maximum likelihood parameter estimation method is optimal for the type I, II and III extreme value distribution. The extreme value wind velocity evaluations of the pseudo wind speed samples during the return period are sensitive to the parent distribution of the pseudo wind speed samples, so different parent distributions of wind speed should be targeted to the selection of different type of extreme value distribution.

**Key words:** Monte Carlo simulation method; pseudo wind speed sample; extreme value distribution; parameter estimation; maximum likelihood estimation; extreme wind speed

风工程中,建筑结构不但要承受过去某一段

时间的风速,还要保证在某一规定的时间期限内安全可靠地承受可能经受的风速.自然界中的风速具有随机性,不同时间有不同的规律,因此有必要根据数理统计的方法来求出建筑结构的设计风速,尤其是对一些重要的对风敏感的结构如输电

收稿日期: 2012-04-16.

基金项目: 国家自然科学基金资助项目(90715039,90815028).

作者简介: 卢安平(1986—),男,硕士研究生;

赵林(1974—),男,副研究员,博士生导师.

通信作者: 赵林, zhaolin@tongji.edu.cn.

塔,桥梁,桅杆等<sup>[1-2]</sup>. 大多数荷载规范只能较好地用于建筑结构的风荷载静力分析或是拟静力分析,所以,估算工程场地重现期内的极值风速是工程抗风设计的首要任务.

根据国内外风荷载规范<sup>[3-5]</sup>,目前常用的适用于推算工程场地重现期内的极值风速的概率统计方法有极值 I 型、II 型和 III 型分布. 针对上述 3 种极值分布类型,目前已经发展了多种参数估计方法<sup>[6-8]</sup>,普遍的方法有矩法、极大似然参数估计法和概率矩法等,特定的方法有最小二乘法、可靠性矩法和线性回归法等.

本文通过 Monte Carlo 方法,产生分别服从指数分布、正态分布、瑞利分布、威布尔分布、对数正态分布、广义极值分布共 6 组伪风速母样,分别采用极值 I 型、II 型和 III 型分布推算得到的不同重现期内极值风速,并与由伪风速母样理论极值分布函数得到的不同重现期内的极值风速相比,获得 3 种极值分布的最佳适用条件. 同时对 3 种极值分布的几种常用参数估计方法进行比较研究,得到不同极值分布类型的优选参数估计方法.

## 1 极值分布类型及参数估计方法

### 1.1 极值分布类型

设  $X_1, X_2, \dots$  是独立同分布的随机变量,分布函数为  $F(x)$ , 对自然数  $n$ , 令  $M_n = \max\{X_1, X_2, \dots, X_n\}$ , 表示  $n$  个随机变量的最大值, 则

$$P(M_n \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = F^n(x), x \in R.$$

如已知分布函数  $F(x)$ , 可根据上式, 精确求出最大值的分布函数. 但在实际应用中, 随机变量的原始分布函数  $F(x)$  往往未知, 为求得原始分布样本的极值分布, 必须考虑  $n \rightarrow \infty$  时极值分布函数的情况, 即极值渐进分布, 简称极值分布.

根据 Fisher-Tippett<sup>[9]</sup> 的极值类型定理, 无论随机变量的原始分布具有何种形式, 如果极大值渐进分布存在, 都可用 3 种分布类型描述:

1) 极值 I 型分布 (又称 Gumbel 分布)

$$F_1(x; a, b) = \exp\left\{-\exp\left(-\frac{x-b}{a}\right)\right\}, -\infty < x < +\infty.$$

2) 极值 II 型分布 (又称 Frechet 分布)

$$F_2(x; a, b, \gamma) = \begin{cases} 0, & x \leq b; \\ \exp\left\{-\left(\frac{x-b}{a}\right)^{-\gamma}\right\}, & x > b. \end{cases}$$

3) 极值 III 型分布 (又称 Weibull 分布)

$$F_3(x; a, b, \gamma) = \begin{cases} \exp\left\{-\left(-\frac{x-b}{a}\right)^\gamma\right\}, & x < b; \\ 0, & x \geq b. \end{cases}$$

式中:  $b$  为位置参数,  $a$  为尺度参数,  $\gamma$  为形状参数.

根据经典极值理论, 上述 3 种极值分布的一般形式可归结为

$$F(x; a, b, \xi) = \exp\left\{-\left(1 + \xi \cdot \frac{x-b}{a}\right)^{-\frac{1}{\xi}}\right\}, \\ 1 + \xi \cdot \frac{x-b}{a} > 0.$$

式中:  $\xi$  为形状参数,  $-\infty < \xi < +\infty$ . 改变  $\xi$  值, 可获得不同极值分布类型. 当  $\xi > 0$ , 为位置参数等于  $b - a/\xi$ 、尺度参数等于  $a/\xi$  的极值 II 型分布, 对应母体分布为柯西型分布; 当  $\xi = 0$ , 为位置参数等于  $b$ 、尺度参数等于  $a$  的极值 I 型分布, 对应母体分布为指数型分布; 当  $\xi < 0$ , 为位置参数等于  $b + a/\xi$ 、尺度参数等于  $-a/\xi$  的极值 III 型分布, 对应母体分布为有界型极值分布<sup>[10]</sup>.

如果年最大风速  $x$  超过强风风速  $x_T$  的事件平均  $T$  年发生一次, 则  $T$  称为风速  $x_T$  的重现期. 由于年最大风速相互独立, 因此  $x \geq x_T$  事件的发生可看作贝努力试验, 那么当前  $n-1$  未发生强风, 但在第  $n$  年发生时, 则事件的发生概率为

$$f_N(1, n-1) = p \cdot (1-p)^{n-1}.$$

式中:  $p$  为每年发生强风的概率. 由于风速重现期  $T$  是发生一次强风时间间隔  $N$  年的期望值, 故

$$T = E(N) = \sum_{n=1}^{\infty} n \cdot f_N(1, n-1) = \frac{1}{p} = \frac{1}{1-F(x_T)}.$$

对于确定的极值分布函数  $F(x_T)$ , 可以求得任意重现期内的强风风速的期望值  $x_T$ .

### 1.2 PPCC 检验方法

极值风速样本能否拟合为某一分布类型, 主要依赖于统计检验. 常用统计检验方法有:  $\chi^2$  检验、K-S 检验和广义似然比检验, 但由于风速实测数据一般数量较少, 常采用可用于较小样本的概率曲线相关系数检验方法 (PPCC)<sup>[11]</sup>. 概率曲线相关系数法基本原理是一个待检样本的实际分布服从某假定分布  $F(x)$ , 则排序后的观测值  $X_i$  与假设分布理论值的概率点存在近似线性关系. 概率曲线相关系数法表达式为

$$\gamma_D = \frac{\sum (X_i - \bar{X}) \cdot [F^{-1}(P_i) - \overline{F^{-1}(P)}]}{\sqrt{\sum (X_i - \bar{X})^2 \cdot \sum [F^{-1}(P_i) - \overline{F^{-1}(P)}]^2}}$$

式中:  $\bar{X} = \frac{1}{n} \sum X_i$ ,  $X_i$  是原样本重新按升序排列;  $P_i$  为与  $X_i$  对应的经验概率值, 参照 Gringorten<sup>[12]</sup> 提出的  $P_i = (i - 0.44)/(n + 0.12)$ ,  $i$  为  $X_i$  按升序排列中排第  $i$  位置,  $n$  为样本容量;  $F^{-1}(x)$  为分布函

数  $F(x)$  的反函数,  $\overline{F^{-1}(P)} = \frac{1}{n} \sum F^{-1}(P_i)$ . 可知, 在已知样本和概率分布函数情况下, 概率曲线相关系数  $\gamma_D$  只与形状参数有关.

### 1.3 参数估计方法

#### 1.3.1 极大似然参数估计法

极大似然函数法<sup>[8]</sup>在已知变量累积分布函数的前提下, 被广泛用于各种参数估计, 并且具有良好的效果. 根据极大似然原理, 首先构造极值分布一般形式的似然函数, 函数式如下:

$$L(a,b) = \prod_{i=1}^n \frac{dF(x)}{dx} = \left(\frac{1}{a}\right)^n \prod_{i=1}^n \left(1 + \xi \cdot \frac{x_i - b}{a}\right)^{\frac{\xi+1}{\xi}} \cdot \exp\left\{-\sum_{i=1}^n \left(1 + \xi \cdot \frac{x_i - b}{a}\right)^{-\frac{1}{\xi}}\right\}.$$

根据极大似然原理:

$$\frac{\partial L(a,b)}{\partial a} = 0, \frac{\partial L(a,b)}{\partial b} = 0.$$

求解得出 3 种极值分布参数估计计算公式.

1) 极值 I 型极大似然参数估计计算公式:

$$\sum_{i=1}^n x_i \exp\left(-\frac{x_i}{\hat{a}}\right) - (\bar{x} - \hat{a}) \cdot \sum_{i=1}^n \exp\left(-\frac{x_i}{\hat{a}}\right) = 0, \tag{1a}$$

$$\hat{b} = -\hat{a} \cdot \ln\left[\frac{1}{n} \cdot \sum_{i=1}^n \exp\left(-\frac{x_i}{\hat{a}}\right)\right]. \tag{1b}$$

2) 极值 II 型极大似然参数估计计算公式:

$$\sum_{i=1}^n (x_i - \hat{b})^{-(\gamma+1)} - \frac{\gamma+1}{n\gamma} \sum_{i=1}^n (x_i - \hat{b})^{-\gamma} \cdot \sum_{i=1}^n (x_i - \hat{b})^{-1} = 0, \tag{2a}$$

$$\hat{a} = \left[\frac{1}{n} \cdot \sum_{i=1}^n (x_i - \hat{b})^{-\gamma}\right]^{\frac{1}{\gamma}}. \tag{2b}$$

3) 极值 III 型极大似然参数估计计算公式:

$$\sum_{i=1}^n (\hat{b} - x_i)^{\gamma-1} - \frac{\gamma-1}{n\gamma} \sum_{i=1}^n (\hat{b} - x_i)^{\gamma} \cdot \sum_{i=1}^n (\hat{b} - x_i)^{-1} = 0, \tag{3a}$$

$$\hat{a} = \left[\frac{1}{n} \cdot \sum_{i=1}^n (\hat{b} - x_i)^{\gamma}\right]^{\frac{1}{\gamma}}. \tag{3b}$$

极值 I 型由式(1)可直接进行参数估计; 对于极值 II 和 III 型, 应先确定形状参数  $\gamma$ , 而  $\gamma$  取值应该满足概率曲线相关系数  $\gamma_D$  渐进于 1 的原则, 则最优的  $\hat{\gamma}$  值应使  $\gamma_D$  取得最大值, 即

$$\gamma_{D,\max} = f(\hat{\gamma}). \tag{4}$$

但计算中常面对两个困难: 首先, 相关系数  $\gamma_D$  随着形状参数  $\gamma$  单调递增, 因而不存在可使  $\gamma_D$  取最大值的  $\hat{\gamma}$  值, 表现为  $\gamma$  值上溢; 其次, 满足式(4)的  $\hat{\gamma}$  值代入式(2)和式(3), 无法求得估计参

数, 表现为极大似然法失效. 本文采用文献[8]提出的基于极大似然函数法的逐步迭代估计法, 能解决参数估计中溢出及失效的问题, 充分发挥极大似然参数估计法的最优估计效率. 不同的是, 为了提高计算效率, 本文在求解极大似然参数方程时, 将概率矩法求解得到的参数估计值作为极大似然方程迭代法求解的初值, 从而获得极大似然函数法参数估计值.

#### 1.3.2 概率矩参数估计法

概率矩法<sup>[10]</sup>是利用样本的概率加权矩来估计母体概率加权值而获得参数估计的一种方法. 对极值 I 型、II 型、III 型 3 种极值分布采用概率矩法估计参数值, 其  $r$  ( $r = 0, 1, 2, \dots$ ) 阶概率矩值  $k_r$  一般计算公式为

$$k_r = E[x(F(x;a,b,\gamma))^r] \quad (r = 0, 1, 2, \dots),$$

样本的前三阶概率加权矩值的无偏量估计值为

$$\hat{k}_0 = \frac{1}{n} \sum_{i=1}^n x_i, \quad \hat{k}_1 = \sum_{i=1}^{n-1} \frac{(n-i)x_i}{n(n-1)},$$

$$\hat{k}_2 = \sum_{i=1}^{n-2} \frac{(n-i)(n-i-1)x_i}{n(n-1)(n-2)}.$$

根据  $k_r = \hat{k}_r$  的无偏量估计原则, 分别获得极值 I 型、II 型和 III 型分布的参数估计计算公式:

1) 极值 I 型概率矩法参数估计计算公式为

$$\hat{a} = 2.8854 \cdot \hat{k}_1 - 1.4427 \cdot \hat{k}_0, \quad \hat{b} = 1.8327 \cdot \hat{k}_0 - 1.6654 \cdot \hat{k}_1.$$

2) 极值 II 型概率矩法参数估计计算公式为

$$-\frac{1}{\hat{\gamma}} = 7.859z + 2.9559z^2, \quad \hat{a} = \frac{(2\hat{k}_1 - \hat{k}_0)}{\Gamma\left(1 - \frac{1}{\hat{\gamma}}\right)\left(2^{\frac{1}{\hat{\gamma}}} - 1\right)}, \hat{b} = \hat{k}_0 - \hat{a} \cdot \Gamma\left(1 - \frac{1}{\hat{\gamma}}\right).$$

式中:  $\Gamma\left(1 - \frac{1}{\gamma}\right)$  表示伽马 (Gamma) 函数,

$$\Gamma\left(1 - \frac{1}{\gamma}\right) = \int_0^\infty e^{-u} \cdot u^{\left(-\frac{1}{\gamma}\right)} du, \quad z = \frac{2\hat{k}_1 - \hat{k}_0}{3\hat{k}_2 - \hat{k}_0} - \frac{\ln 2}{\ln 3}$$

3) 极值 III 型概率矩法参数估计计算公式为

$$\frac{1}{\hat{\gamma}} = 7.859z + 2.9559z^2, \quad \hat{a} = -\frac{(2\hat{k}_1 - \hat{k}_0)}{\Gamma\left(1 + \frac{1}{\hat{\gamma}}\right)\left(2^{\frac{1}{\hat{\gamma}}} - 1\right)}, \hat{b} = \hat{k}_0 + \hat{a} \cdot \Gamma\left(1 + \frac{1}{\hat{\gamma}}\right).$$

$$\text{式中: } \Gamma\left(1 + \frac{1}{\gamma}\right) = \int_0^\infty e^{-u} \cdot u^{\frac{1}{\gamma}} du, \quad z = \frac{2\hat{k}_1 - \hat{k}_0}{3\hat{k}_2 - \hat{k}_0} - \frac{\ln 2}{\ln 3}$$

## 2 Monte Carlo 方法产生伪风速母样

基于极值理论和现有风速母样分布的研究成果<sup>[6,8,14]</sup>, 风速母样分布的类型有指数分布、正态

分布、瑞利分布、威布尔分布、对数正态分布、广义极值分布. 简单介绍一下不常见的瑞利分布、威布尔分布、对数正态分布.

瑞利分布的概率密度函数和累积分布函数分别为

$$f(x; \sigma^2) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right), x \geq 0;$$

$$F(x; \sigma^2) = 1 - \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

式中  $\sigma$  为瑞利分布参数.

威布尔分布的概率密度函数和累积分布函数分别为

$$f(x; \lambda, k) = \begin{cases} \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} \cdot \exp\left(-\left(\frac{x}{\lambda}\right)^k\right), x \geq 0; \\ 0, & x < 0. \end{cases}$$

$$F(x; \lambda, k) = 1 - \exp\left(-\left(\frac{x}{\lambda}\right)^k\right).$$

式中:  $\lambda$  为尺度参数,  $k$  为形状参数.

对数正态分布的概率密度函数和累积分布函

数分别为

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right), x > 0;$$

$$F(x; \mu, \sigma) = \int_0^x f(x; \mu, \sigma) dx = \Phi\left(\frac{\ln x - \mu}{\sigma}\right).$$

式中:  $\mu$  为位置参数,  $\sigma$  为尺度参数,  $\Phi(\cdot)$  为标准正态分布函数.

以日最大风速作为样本, 假定生成 1961 年 1 月 1 日到 2000 年 12 月 31 日的容量为  $365 \times 40 = 14\ 600$  的伪风速母样. 具体方法是根据乘同余数法, 采用著名的 Park&Miller<sup>[15]</sup> 随机数产生等式产生 14 600 个  $[0, 1]$  区间均匀分布的伪随机数  $\theta_i$ , 再由逆变换方法分别转换成满足风速母体分布的伪随机数  $x_i$ .

$$x_i = F^{-1}(\theta_i).$$

式中  $F$  为指数分布、正态分布、瑞利分布、威布尔分布、对数正态分布、广义极值分布. 参照实测风速数据拟合结果<sup>[6,8,13-14]</sup>, 假定各母样分布函数的参数见表 1, 产生的伪风速母样和对应的伪风速母样分布曲线见图 1.

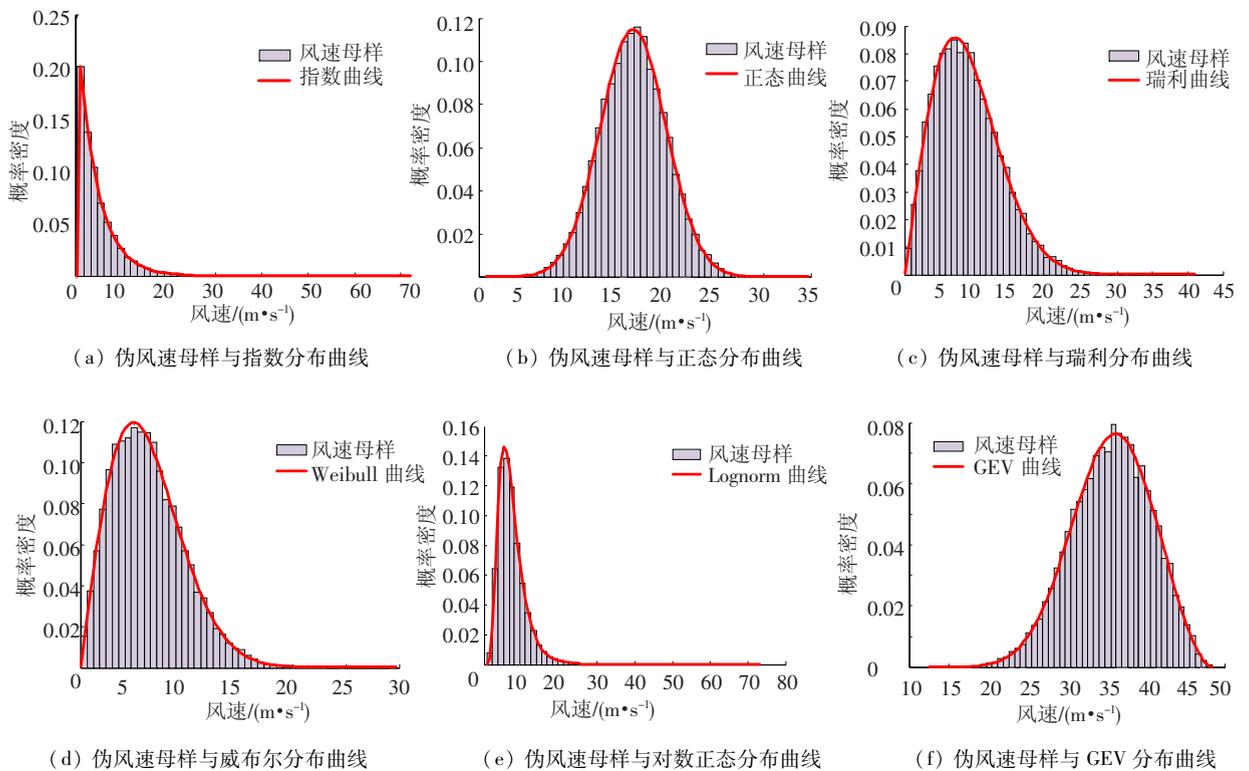


图 1 伪风速母样和对应的分布曲线

表 1 风速母样分布参数

分布类型	分布参数	取值	分布类型	分布参数	取值
指数分布	$\lambda$	0.233	Lognorm 分布	$\mu$	1.90
	$\mu$	16.26		$\sigma$	0.45
正态分布	$\sigma$	3.47		$b$	34.00
瑞利分布	$\sigma$	7.07	GEV	$a$	5.40
Weibull 分布	$\lambda$	0.021		$\zeta$	-0.36
	$k$	1.97			

### 3 参数估计方法比较研究

若风速母样分布为  $F(x)$ , 年风速观测值样本  $X = \{x_1, x_2, \dots, x_n\}$ , 现假定服从风速母样分布的风速观测值相互独立 (实际中大自然界风速观测值之间存在相关性), 根据概率统计理论, 则年风速理论极值分布函数为

$$P(X \leq x) = P(x_1 \leq x, \dots, x_n \leq x) = F^n(x).$$

式中:当风速观测值样本  $X$  为日极值风速时,  $n = 365$ . 由于根据乘同余数法和相应母样分布函数产生的伪风速母样样本相互独立,因此可以将年风速理论极值分布函数(母样分布 365 次方)和不同参数估计方法得到的极值 I 型、II 型和 III 型分布曲线进行对比,获得极值 I 型、II 型和 III 型分布极大似然估计法和概率矩法参数估计精度.

### 3.1 极值 I 型参数估计方法比较

基于容量为 14 600 的满足上述 6 种概率分布的 6 组伪风速母样,采用年极值采样法和极值 I 型分布,分别对每组伪风速母样进行参数估计,获得极值 I 型分布的位置参数  $b$  和尺度参数  $a$  的估计值见表 2,对应的极值 I 型分布曲线见图 2.

表 2 各伪风速母样极值 I 型分布的参数估计值

风速母样分布类型	参数估计方法	极值 I 型分布		风速母样分布类型	参数估计方法	极值 I 型分布	
		$b$	$a$			$b$	$a$
指数分布	极大似然函数法	26.45	4.91	Weibull 分布	极大似然函数法	17.79	1.56
	概率矩法	28.91	1.52		概率矩法	18.56	0.43
	矩法(规范)	22.37	12.85		矩法(规范)	16.71	3.64
正态分布	极大似然函数法	26.11	1.16	Lognorm 分布	极大似然函数法	24.35	4.31
	概率矩法	26.68	0.31		概率矩法	26.52	1.55
	矩法(规范)	25.33	2.65		矩法(规范)	20.04	12.78
瑞利分布	极大似然函数法	24.69	2.14	GEV 分布	极大似然函数法	47.33	0.86
	概率矩法	25.74	0.58		概率矩法	46.97	0.09
	矩法(规范)	23.21	4.96		矩法(规范)	47.03	0.86

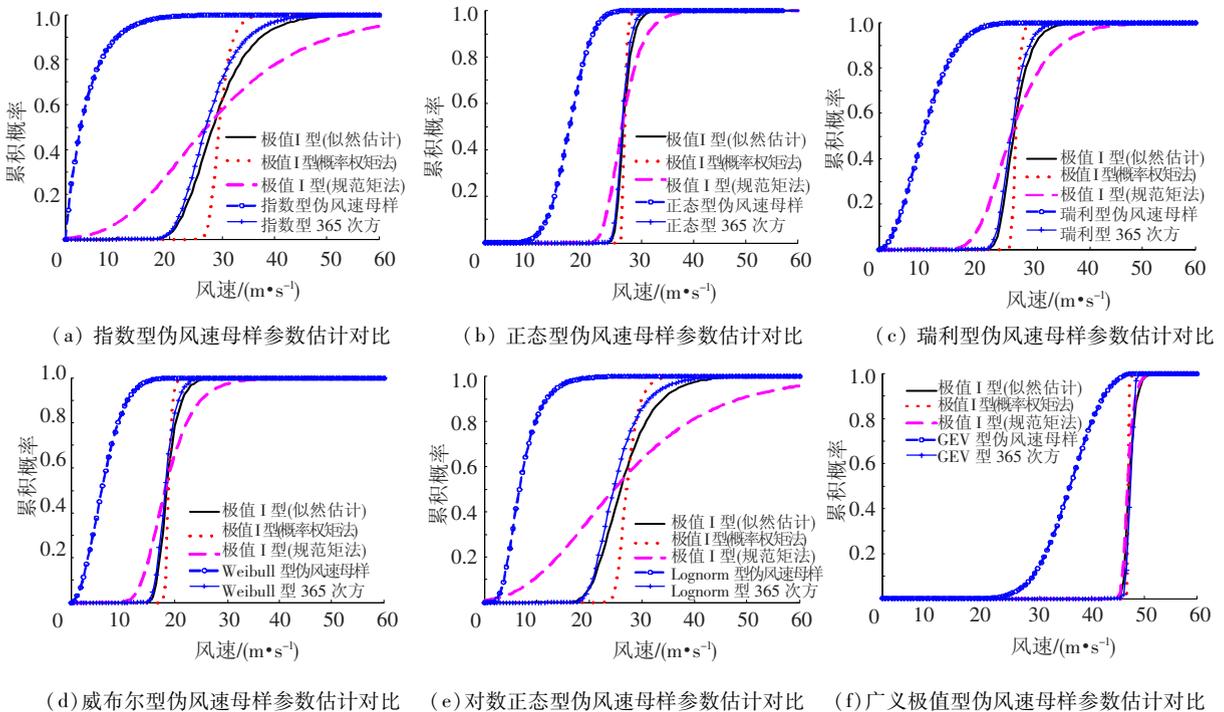


图 2 各伪风速母样极值 I 型参数估计对比

从表 2 看,对极值 I 型分布参数估计,极大似然参数估计法、概率矩法和我国建筑结构荷载规范<sup>[3]</sup>采用的矩法 3 种参数估计方法对位置参数  $b$  估计的稳定性明显优于对尺度参数  $a$  的估计. 3 种参数估计方法获得的位置参数和尺度参数估计值相差较大,特别是尺度参数估计值. 这结论与文献<sup>[6]</sup>的不同种方法得到的位置参数估计结果明显好于尺度参数估计结果的结论相同. 从图 2 看,

满足上述 6 种母样分布的伪风速母样,采用概率矩法和矩法获得的极值 I 型分布函数与年风速理论极值分布函数(母样分布 365 次方)相差较大,极大似然函数法对参数的估计值精度较高. 对极值 I 型分布,规范<sup>[3]</sup>采用的矩估计法获得重现期内的极值风速偏保守,概率矩法偏危险.

### 3.2 极值 II 型参数估计方法比较

基于上述 6 组伪风速母样,采用年极值采样

法和极值 II 型分布, 分别对每组伪风速母样进行参数估计. 利用极大似然函数参数估计法和概率矩法求解得到极值 II 型分布的位置参数  $b$ 、尺度参数  $a$  和形状参数  $\gamma$  估计值见表 3, 对应的极值 II 型分布曲线见图 3.

从表 3 看, 对极值 II 型分布的参数估计, 极大似然参数估计法的稳定性明显优于概率矩法; 概率矩法对服从 GEV 分布的伪风速母样参数

估计失效, 得到的尺度参数  $a$  和形状参数  $\gamma$  估计值都小于零; 极大似然参数估计法和概率矩法获得的 3 个参数的估计值相差较大, 特别是尺度参数估计值. 从图 3 看, 满足上述 6 种母样分布的伪风速母样, 采用概率矩法获得的极值 II 型分布函数与年风速理论极值分布函数(母样分布 365 次方)相差较大, 极大似然函数法对参数的估计值精度较高.

表 3 各伪风速母样极值 II 型分布的参数估计值

风速母样分布类型	参数估计方法	极值 II 型分布			风速母样分布类型	参数估计方法	极值 II 型分布		
		$b$	$a$	$\gamma$			$b$	$a$	$\gamma$
指数分布	极大似然函数法	16.182	8.922	2.944	Weibull 分布	极大似然函数法	13.133	4.418	3.298
	概率矩法	27.541	1.128	1.800		概率矩法	18.007	0.493	2.205
正态分布	极大似然函数法	21.191	4.733	5.300	Lognorm 分布	极大似然函数法	15.520	7.511	3.600
	概率矩法	26.251	0.384	2.280		概率矩法	25.469	0.796	1.576
瑞利分布	极大似然函数法	18.719	5.642	3.000	GEV 分布	极大似然函数法	44.514	2.338	12.911
	概率矩法	24.979	0.677	2.213		概率矩法	48.318	-0.84 < 0	-8.0 < 0

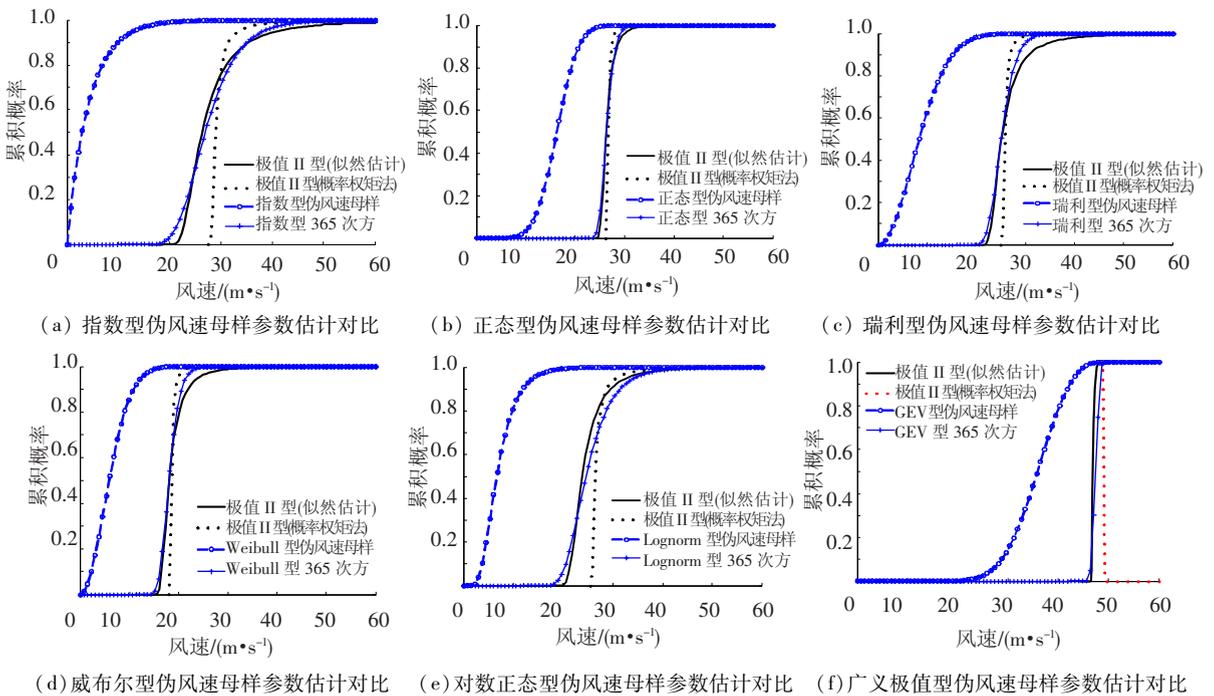


图 3 各伪风速母样极值 II 型参数估计对比

### 3.3 极值 III 型参数估计方法比较

对上述 6 组伪风速母样, 采用年极值采样法和极值 III 型分布, 分别对每组伪风速母样进行参数估计. 利用极大似然函数参数估计法和概率矩法求解得到极值 III 型分布的参数估计值见表 4, 对应的分布曲线见图 4.

从表 4 看, 对极值 III 型分布的参数估计, 同样极大似然参数估计法的稳定性明显优于概率矩法; 概率矩法对服从指数分布、正态分布、瑞利分布、威布尔分布和对数正态分布的伪风速母样参数估计失效, 得到的尺度参数  $a$  和形状参数  $\gamma$

估计值都小于零, 对服从 GEV 分布的伪风速母样参数估计有效, 正好与极值 II 型分布相反. 这是因为, 概率矩参数估计法是从样本数据出发, 忽略了极值分布本身所提供的信息, 而概率加权矩值的无偏量估计值只能唯一. 极大似然参数估计法和概率矩法两方法对服从 GEV 分布的伪风速母样参数估计, 获得的 3 个参数的估计值相差较大, 特别是尺度参数估计值. 从图 4 看, 满足上述 6 种母样分布的伪风速母样, 极大似然函数法对参数的估计值精度较高.

表 4 各伪风速母样极值 III 型分布的参数估计值

风速母样分布类型	参数估计方法	极值 III 型分布			风速母样分布类型	参数估计方法	极值 III 型分布		
		<i>b</i>	<i>a</i>	$\gamma$			<i>b</i>	<i>a</i>	$\gamma$
指数分布	极大似然函数法	131.23	104.51	18.5	Weibull 分布	极大似然函数法	39.998	22.089	12
	概率权矩法	27.541	-1.13 < 0	-1.8 < 0		概率权矩法	18	-0.49 < 0	-2.2 < 0
正态分布	极大似然函数法	42.397	16.2	12	Lognorm 分布	极大似然函数法	107.08	82.33	14.5
	概率权矩法	26.859	-0.38 < 0	-2.28 < 0		概率权矩法	25.469	-0.8 < 0	-1.58 < 0
瑞利分布	极大似然函数法	54.98	30.13	12	GEV 分布	极大似然函数法	50.16	2.867	5
	概率权矩法	24.98	-0.68 < 0	-2.21 < 0		概率权矩法	48.32	0.841	8.01

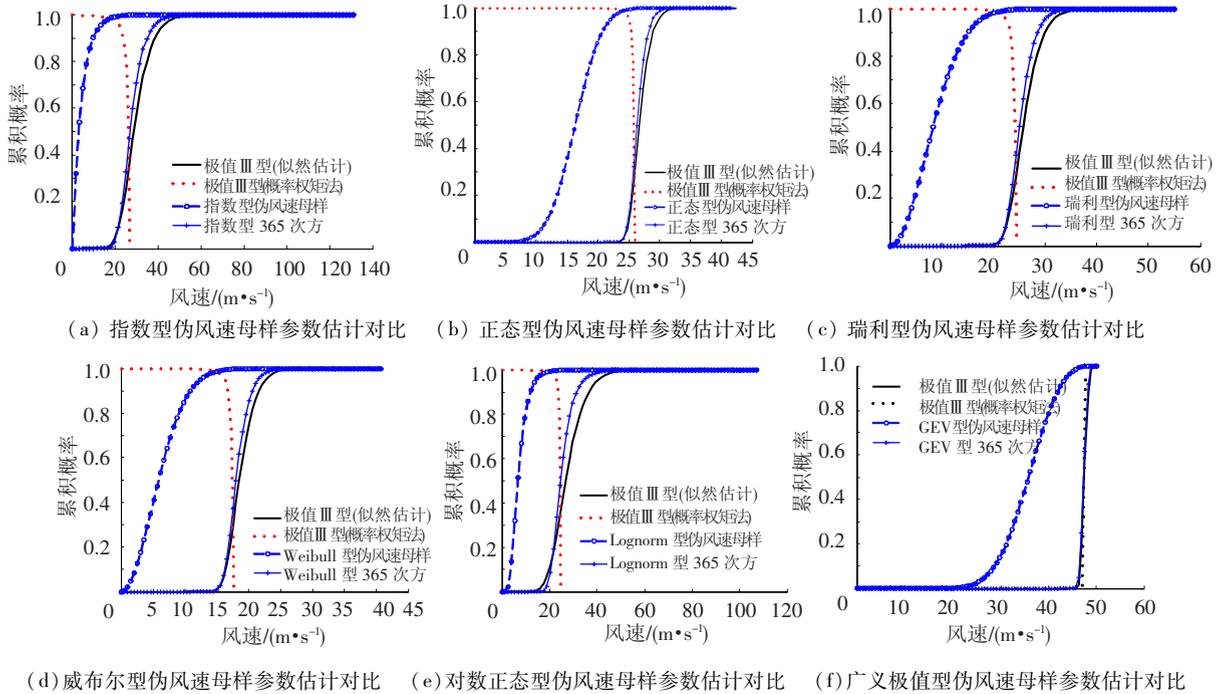


图 4 各伪风速母样极值 III 型参数估计对比

综上,极大似然函数估计法对极值 I 型、II 型和 III 型 3 种极值分布类型的参数估计都适用,估计精度都较高. 概率权矩法参数估计稳定性较差,特别是对极值 II 分布和极值 III 型分布,概率权矩法可能失效. 这是因为,概率权矩参数估计法是从样本数据出发,忽略了极值分布本身所提供的信息,而极大似然函数法在已知变量累积分布函数的前提下,结合样本数据进行的参数估计,因此极大似然函数法具有良好的效果.

### 4 极值分布类型比较研究

基于 6 组伪风速母样,采用极大似然函数参数估计法,将极值 I 型、II 型、III 型 3 种极值分布获得的不同重现期内的极值风速与基于风速母样分布函数得到的理论极值风速进行对比,获得 3 种极值分布的各自最佳适用条件,对比结果见表 5.

定义相对偏差:

$$err = \frac{(\hat{x} - x)}{x} \times 100\%.$$

式中: $\hat{x}$  为重现期极值风速估计值; $x$  为根据母样分布函数获得的重现期内理论极值风速.

由表 5 可知,与基于风速母样分布函数得到的不同重现期内的理论极值风速相比,对服从不同母样分布的伪风速母样,极值 I 型、II 型和 III 型获得的不同重现期内的极值风速计算精度各不相同;对服从指数分布、正态分布、瑞利分布和 Weibull 分布的伪风速母样,极值 I 型和 III 型都能给出计算精度较高的极值风速预测;对服从 Lognorm 分布的伪风速母样,极值 II 型给出最优的极值风速估算;对服从 GEV 分布(形状参数  $\gamma < 0$ , 实际为极值 III 型)的伪风速母样,极值 III 型给出最优的极值风速估算. 上述结论与文献[8]结论相同,但文献[8]采用的风速母样分布函数、Monte Carlo 数值模拟方法与本文不同. 因此,重现期内极值风速的估算对风速母样分布类型敏感,服从不同母样分布的风速母样,应有针对性选用极值分布类型估算重现期内极值风速.

表5 不同样本和不同极值分布类型推算的极值风速

m · s<sup>-1</sup>

阶段极值 理论	指数分布				正态分布				瑞利分布			
	50年		100年		50年		100年		50年		100年	
	极值	偏差/%										
理论分布	42.1	0	45.2	0	29.7	0	30.3	0	31.3	0	32.4	0
极值 I 型	45.6	8.3	49.0	8.4	30.6	3.0	31.4	3.6	33.0	5.4	34.5	6.5
极值 II 型	49.8	18.3	58.7	29.9	31.1	4.7	32.5	7.3	39.4	25.9	44.9	38.6
极值 III 型	46.6	10.7	49.7	10.0	30.7	3.4	31.4	3.6	33.2	6.1	34.4	6.2

阶段极值 理论	Weibull 分布				Lognorm 分布				GEV 分布			
	50年		100年		50年		100年		50年		100年	
	极值	偏差/%	极值	偏差/%	极值	偏差/%	极值	偏差/%	极值	偏差/%	极值	偏差/%
理论分布	22.7	0	23.4	0	38.1	0	41.1	0	48.6	0	48.7	0
极值 I 型	23.9	5.3	25.0	6.8	41.2	8.1	44.2	7.5	50.1	3.1	50.7	4.1
极值 II 型	27.6	21.6	31.0	32.5	37.7	-1.0	42.5	3.4	47.7	-1.9	47.9	-1.6
极值 III 型	24.0	5.7	24.9	6.4	44.2	16.0	47.1	14.6	48.8	0.4	49.0	0.6

## 5 结 论

1) 对于极值 I 型分布,与年风速理论极值分布函数相比,规范采用的矩估计法获得重现期内的极值风速偏保守,概率矩法偏危险,极大似然函数法参数估计精度最高。

2) 极大似然函数参数估计法对 3 种极值分布的参数估计都适用,估计精度都较高. 概率矩法参数估计稳定性较差,特别是对极值 II 分布和极值 III 型分布,概率矩法参数估计可能失效。

3) 重现期内极值风速的估算对风速母样分布类型敏感,服从不同母样分布的风速母样,应有针对性选用极值分布类型估算重现期内极值风速。

## 参考文献

- [1] COLES S G, TAWN J A. Statistical methods for multivariate extremes: an application to structural design [J]. Applied Statistics, 1994, 43(1):1-48.
- [2] 李宏男,王杨,伊廷华. 极值风速概率方法研究进展 [J]. 自然灾害学报,2009,18(2):15-26.
- [3] GB50009—2001 建筑结构荷载规范[S]. 北京:中国建筑工业出版社,2002.
- [4] ASCE. Minimum design loads for buildings and other structures. ANSI/ASCE Standard 7-95[S]. Reston, VA,1998.
- [5] Canadian Commission on Building and Fire Code. National building code of Canada[S]. Ottawa:The National Research Council of Canada,2002.
- [6] 段忠东,欧进萍,周道成. 极值风速的最优概率模型 [J]. 土木工程学报,2002,35(5):11-16.
- [7] 段忠东,周道成. 极值概率分布参数估计方法的比较研究 [J]. 哈尔滨工业大学学报,2004,36(12):1605-1609.
- [8] 赵林,葛耀君,项海帆. 极值风速拟合优化策略 [J]. 同济大学学报:自然科学版,2003,31(4):383-388.
- [9] FISHER R A, TIPPETT L H C. Limiting forms of the frequency distribution of the largest of smallest member of a sample [J]. Proc Cambridge Philos Soc,1928,24:180-190.
- [10] 史道济. 实用极值统计方法 [M]. 天津:天津科学技术出版社,2006.
- [11] SIMIU E, FILLIBEN J J. Statistical analysis of extreme winds [R]. Technical Note 868. Washington D. C.: National Bureau of Standards, 1975.
- [12] GRINGORTEN I I. A plotting rule for extreme probability paper [J]. Journal of Geophysical Research, 1968,68(3):813-814.
- [13] 葛耀君. 宝山、川沙和龙华气象站日最大风速记录数据汇编 [R]. 上海:同济大学土木工程防灾国家重点实验室,1997.
- [14] 欧进萍,段忠东,陆钦年. 渤海海域的风特性统计分析 [J]. 海洋通报,1997,16(1):20-28.
- [15] PARK S K, MILLER K W. Random number generators: good ones are hard to find [J]. Communications of the ACM, 1988, 31(10):1192-1201.

(编辑 赵丽莹)