

江志红,刘冬,刘渝,等.导线覆冰极值的概率分布模拟及其应用试验[J].大气科学学报,2010,33(4):385-394.

Jiang Zhi-hong, Liu Dong, Liu Yu, et al. Study on the probability distribution simulation of conductor icing and its application test[J]. Trans Atmos Sci, 2010, 33(4): 385-394.

导线覆冰极值的概率分布模拟及其应用试验

江志红¹, 刘冬¹, 刘渝², 熊海星², 鞠晓慧³

(1. 南京信息工程大学 气象灾害省部共建教育部重点实验室, 江苏 南京 210044;

2. 中国电力工程顾问集团西南电力设计院, 四川 成都 610000; 3. 国家气象信息中心, 北京 100081)

摘要:利用南方地区多个气象站和电力部门观冰站的导线覆冰逐日冰厚资料,将广义极值分布和广义帕雷托分布引入导线覆冰的概率模型研究中,通过超限门覆冰次数的泊松分布拟合检验,结合Hill图解,提出了基于超限门峰值法门限值的确定方法;对两种分布在导线覆冰极值模型拟合的适用性研究表明,广义帕雷托分布对各站覆冰冰厚极值的拟合精度最高;重现期冰厚极值估计随样本长度的变化分析表明,广义帕雷托分布模型极值估计的稳定性比广义极值分布强,一般样本容量达到25 a左右时,广义帕雷托分布重现期冰厚极值的估计趋于稳定,可以作为短序列下估计导线覆冰极值的较好方法。

关键词:广义极值分布模型;广义帕雷托分布模型;导线覆冰;重现期

中图分类号:P429 文献标识码:A 文章编号:1674-7097(2010)04-0385-10

Study on the Probability Distribution Simulation of Conductor Icing and Its Application Test

JIANG Zhi-hong¹, LIU Dong¹, LIU Yu², XIONG Hai-xing², JU Xiao-hui³

(1. Key Laboratory of Meteorological Disaster of Ministry of Education, NUIST, Nanjing 210044, China;

2. CPECC Southwest Electric Power Design Institute, Chengdu 610000, China;

3. National Meteorological Information Center, Beijing 100081, China)

Abstract: The generalized extreme value distribution (GEV) and generalized Pareto distribution (GPD) are used to study the conductor icing probability model with the daily standard icing thickness data in the south of China. By the simulation test of Poisson distribution of the number of exceedance events and Hillplot, the method of identifying threshold is proposed based on Peaks Over Threshold (POT). Numerical simulations show that GPD shows the best accuracy in simulating extremes of each site. The change of icing thickness extremes of Return period under different sample lengths demonstrates that GPD has a better stability than GEV. When the sample length is about 25 years, the estimation of GPD tends to be stable which could be considered as a good method for short array data.

Key words: generalized extreme value distribution (GEV); generalized Pareto distribution (GPD); conductor icing; return period

0 引言

电线覆冰灾害是电网建设与安全运行中影响最

大的安全隐患之一,随着我国强大电网的快速建设发展,特别是我国特高压、大容量、超长距离主干线路的快速建设发展,越来越多的主干线路不得不穿

越严重覆冰的山地区域,科学合理地解决线路设计覆冰问题,这是我国电网规划设计行业所面临的一项十分迫切而重要的任务。

线路设计覆冰问题的关键是科学估计重现期的导线覆冰极值,即预估使用期间可能出现的最大导线覆冰冰厚。实际工作中主要有两类方法,一类偏经验,如利用全部观测资料建立经验频率曲线;另一类则以极值理论为基础,国内外主要采用极值 I 型、极值 II 型和皮尔逊 III 型(P-III 型)研究导线覆冰极值模型,估计重现期冰厚极值。如前苏联等国家认为覆冰冰厚符合 II 型极值分布,欧美国家在覆冰概率分析中采用 I 型极值分布^[1]。国内谭冠日等^[2]研究发现,II 型极值分布的拟合效果较好,且由 15 a 重现期计算的冰厚极值大于 I 型和 P-III 型分布的计算结果。研究表明对于不同的地区选取的模型不尽相同,且不同的模型计算的重现期冰厚极值存在着一定的差异,显然,是否存在通用的、较合适的覆冰极值概率分布模式值得深入研究。

另一方面,建模过程中面临最大的问题是资料缺失,目前主要是利用气象观测资料,但气象观测站大部分在平坝地区,位于山地重冰区域的输电线路基本上没有覆冰观测。近期电力部门也在极个别输电线路附近设立观冰站,但观测成本高,资料完整性差,序列长度不足 10 a。此外值得注意的是前述极值模型采用的抽样方法都是年最大值 AM(annual maximum)统计法,即每年取一个最大值构成系列的统计方法。鉴于覆冰极值出现的随机性很大,有的年份由于结冰条件适宜,可以出现多个极大值,而有的年份则一次极值过程也没有,因此对于短序列每年抽取一个最大值无法充分利用有限观测信息,重现期极值估计的精度受到影响。针对上述问题,如何充分利用有限样本信息,寻求适合短序列的稳健极值分布估计方法也是覆冰极值估计中的重要问题。

20 世纪 50 年代, Jenkinson 将三种类型的极值分布经过适当变换,提出了包含三种极值分布的广义极值分布(generalized extreme value distribution, GEV)^[3-6]。70 年代, Pickands 指出超出阈值的数据在阈值增加时,其分布函数的极限形式趋向广义帕雷托分布(generalized pareto distribution, GPD),即选取某门限值以上的数据进行分析(称这种抽样方法为“超门限峰值 POT”(peaks over threshold)抽样),其样本数据的概率分布为广义帕雷托分布^[7-9],显然 POT 抽样有效增加了极值的样本量。

近 20 ~ 30 a 与上述两种极值分布相关的应用研究受到了应用数学及其各应用领域研究人员的高度重视。目前,它们在水文气象、暴雨、洪水水位、极端气候事件等领域得到了广泛的应用^[10-15]。

本文在导线覆冰冰厚极值估计中引进广义极值分布和广义帕雷托分布两种新的分布模型,在广义帕雷托分布中,通过超门限覆冰次数的泊松分布拟合检验,结合 Hill 图解,提出 POT 抽样门限的确定的方法,研究上述两种分布在导线覆冰冰厚极值估计中的适用性,分析重现期冰厚极值随样本长度的变化,为利用短序列进行覆冰重现期冰厚的极值估计提供科学依据。

1 资料和方法

1.1 资料

选取南方五省 1957—2009 年的庐山、咸宁、峨眉山、绿葱坡、金佛山和三穗 6 个气象站以及电力部门设立的二郎山观冰站 2002—2009 年的导线覆冰标准冰厚(以下简称冰厚)资料作为研究对象。由于各气象站导线覆冰均存在缺测,庐山、咸宁、峨眉山、绿葱坡、金佛山和三穗 6 个气象站有样本的年份分别为 51、43、39、37、37 和 20 a。考虑到二郎山观冰站的南北向和东西向覆冰差异较大,特分两个方向分别进行研究。

1.2 方法

1.2.1 GEV、GPD 概率分布模式及其参数估计

广义极值分布的分布函数为^[3,14,16]:

$$F(x) = \exp\left[-\left\{1 - k \frac{x - \beta}{\alpha}\right\}^{1/k}\right], \quad k \neq 0; \quad (1)$$

$$F(x) = \exp\left[-\exp\left\{-\frac{x - \beta}{\alpha}\right\}\right], \quad k = 0. \quad (2)$$

式中: α 为尺度参数; β 为位置参数; k 为形状参数。当 $k=0$ 时,为 I 型极值分布,即耿贝尔分布(Gumbel 分布); $k<0$ 时为 II 型极值分布; $k>0$ 时,则为 III 型极值分布,即韦布尔分布。

广义帕雷托分布的分布函数可写为^[7-8,15]

$$F(x) = 1 - \left[1 - k \left(\frac{x - \beta}{\alpha}\right)\right]^{1/k}, \quad k \neq 0, \beta \leq x \leq \frac{\sigma}{k}. \quad (3)$$

式中: k 是形状参数; α 是尺度参数; β 为门限值。

参数估计采取 L-矩估计的方法。L-矩估计方法最大的优点是对序列的极大值和极小值没有常规矩敏感,求得的参数估计值比较稳健^[14]。概率权重矩的三阶权重矩可写为^[17-18]:

$$b_0 = \bar{x}, \quad (4)$$

$$b_1 = \sum_{j=1}^{n-1} \frac{(n-j)}{n(n-1)(n-2)} x_j, \quad (5)$$

$$b_2 = \sum_{j=1}^{n-2} \frac{(n-j)(n-j-1)}{n(n-1)(n-2)} x_j. \quad (6)$$

计算上述三式的线性组合,即L矩,分别为:

$$\lambda_1 = b_0, \quad (7)$$

$$\lambda_2 = 2b_1 - b_0, \quad (8)$$

$$\lambda_3 = 6b_2 - 6b_1 + b_0. \quad (9)$$

根据文献[19],即可求得GEV分布参数的估计公式为:

$$k = 7.8590c + 2.9554c^2, \quad (10)$$

$$c = \frac{2}{3 + \lambda_3/\lambda_2} - \frac{\ln 2}{\ln 3}, \quad (11)$$

$$\alpha = \frac{\lambda_2 k}{(1 - 2^{-k})\Gamma(1+k)}, \quad (12)$$

$$\beta = \lambda_1 - \frac{\alpha[1 - \Gamma(1+k)]}{k}. \quad (13)$$

根据文献[20],GPD分布参数 α, k 的通式为:

$$k = \frac{\lambda_1 - \beta}{\lambda_2} - 2, \quad (14)$$

$$\alpha = (\lambda_1 - \beta) \left(1 + \frac{\lambda_1 - \beta}{\lambda_2} - 2\right). \quad (15)$$

1.2.2 概率分布拟合效果检验

通过计算相关系数(R)、均方误差(S_s)和科尔莫哥洛夫检验法(K-S)对模型的拟合效果进行检验。其中:

$$R = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (16)$$

$$S_s = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}}. \quad (17)$$

式中: x_i 为理论频率; y_i 为经验频率; \bar{x} 为理论频率的平均值; \bar{y} 为经验频率的平均值。

科尔莫哥洛夫检验法的具体步骤是根据假设和经验分布函数 $F_n(x)$,计算样本点上的偏差 $d = |F_0(x) - F_n(x)|$,找出这些偏差中的最大值 D_n 。若 n 很大,则可以认为 $D_n \sqrt{n}$ 近似地服从分布函数 $\Phi(\lambda)$,这样就可以根据显著性水平 α ,找到满足 $\Phi(\lambda) = 1 - \alpha$ 的临界值 λ_α ,然后比较 $D_n \sqrt{n}$ 和 λ_α ,若 $D_n \sqrt{n} < \lambda_\alpha$,则接受原假设;否则,拒绝原假设^[6,21-22]。

1.2.3 基于泊松分布和Hill图的门限值选取方法

门限值的选取是广义帕雷托分布应用研究中的关键问题,门限值确定过高,超门限样本数就少,有可能损失独立样本信息,反之,门限值过低,超门限样本数就多,但有可能出现超门限样本间的不独立,估计偏差加大;考虑到超门限覆冰极值的出现为一小概率事件,根据泊松分布(Poisson分布)的性质,相互独立的超门限覆冰极值出现的次数应符合泊松分布。具体方法为:假定超门限覆冰极值出现次数服从泊松分布,则每年发生超门限的次数 k 的概率为^[23]:

$$P_k(K = k) = \frac{\lambda^k}{k!} e^{-\lambda}, (k = 0, 1, 2, \dots). \quad (18)$$

其中: λ 为Poisson分布的参数,在这里就表示年交叉率。

通过 χ^2 检验法对其进行拟合优度检验,判断超门限值发生次数是否符合泊松分布,并由此判断通过POT抽样得到的各次超门限极值是否独立。

Hill图是基于Hill估计量的一种阈值图形分析法。由文献[24],Hill估计量为:

$$\gamma(m) = \frac{1}{m} \sum_{j=1}^m [\ln x_j - \ln x_m]. \quad (19)$$

其中: x_m 为门限值; x_j 为大于门限值的变量。考察Hill估计量随门限值的演变情况,取Hill估计量趋于稳定时对应的数值为最佳门限值。

1.2.4 重现期极值推算

极值分布模式建立的根本目的在于准确地推断重现期概率下极值变量的可能取值,在实际电网工程应用中,GEV分布 T 年重现期的冰厚极值估计式为^[25]:

$$x_T = \beta + \frac{\alpha}{k} \left\{ 1 - \left[-\ln \left(1 - \frac{1}{T} \right) \right]^k \right\}, k \neq 0; \quad (20)$$

$$x_T = \beta - \alpha \ln \left[-\ln \left(1 - \frac{1}{T} \right) \right], k = 0. \quad (21)$$

GPD分布重现期冰厚极值的估算公式为^[25]:

$$x_T = \beta + \frac{\alpha}{k} [1 - (\lambda T)^{-k}], k \neq 0; \quad (22)$$

$$x_T = \beta + \alpha \ln(\lambda T), k = 0. \quad (23)$$

式中: λ 为年交叉率,表示为 $\lambda = m/n$,其中 m 为超过门限的极值数量, n 为资料记录的总年数。

作为样本统计量,重现期冰厚极值将随样本容量变化出现波动,但不同概率模型,对样本长度的敏感性不一,本文通过数值分析,进一步探讨了重现期冰厚极值随样本长度的变化,当然,重现期冰厚极值作为随机变量,还将具有抽样分布,抽样误差也将对

重现期冰厚极值的不确定性有重要影响,限于篇幅,我们将另文讨论。

2 导线覆冰极值概率分布模型的拟合及其应用试验

2.1 覆冰极值 GEV 概率分布型的拟合及其重现期估计

利用庐山、威宁、峨眉山、绿葱坡、金佛山、三穗和二郎山 7 个站的覆冰年冰厚极值观测资料,进行 GEV 概率分布模型拟合。图 1 给出了庐山站、威宁站、三穗站和二郎山南北向 GEV 概率分布模型的拟合情况(其他站图略),通过累积频率的理论值与经验值对比可以看到模型拟合较好,对于样本容量较大的庐山站和威宁站模型的拟合效果要优于三穗站和二郎山南北向。且 7 个站全部通过了柯尔莫洛夫检验,表明导线覆冰冰厚极值符合 GEV 分布。

表 1 列出了基于三种检验指标的 GEV 模型拟合情况,为验证 GEV 模型的可用性,表中还给出了

Gumbel 模型的拟合精度。可以看到,GEV 模型的相关系数全部在 0.97 以上,均方误差和 K-S 统计量也很小,与 Gumbel 模型相比,三种检验指标的差异均不大,表明这两种模型的拟合精度基本一致。

概率分布模型拟合的主要应用是估计不同重现期的覆冰极值。图 2 给出了 GEV 模型和 Gumbel 模型的 15 a、30 a、50 a 和 100 a 重现期的冰厚极值。从图 2 可以看出:利用 GEV 模型计算的重现期冰厚极值与 Gumbel 模型的计算结果基本相同,庐山站、峨眉山站、金佛山站、绿葱坡站和二郎山东西向,GEV 模型的计算结果略小于 Gumbel 模型,而威宁站、三穗站和二郎山南北向的计算结果 GEV 模型略大于 Gumbel 模型。庐山站、峨眉山站、金佛山站两种模型之间的差异均小于 5%,其他站的差异在 15% 以内,且重现期越小,GEV 模型与 Gumbel 模型之间的差异就越小。因此,对实际工程应用而言,GEV 模型计算的重现期冰厚极值可以满足需要。

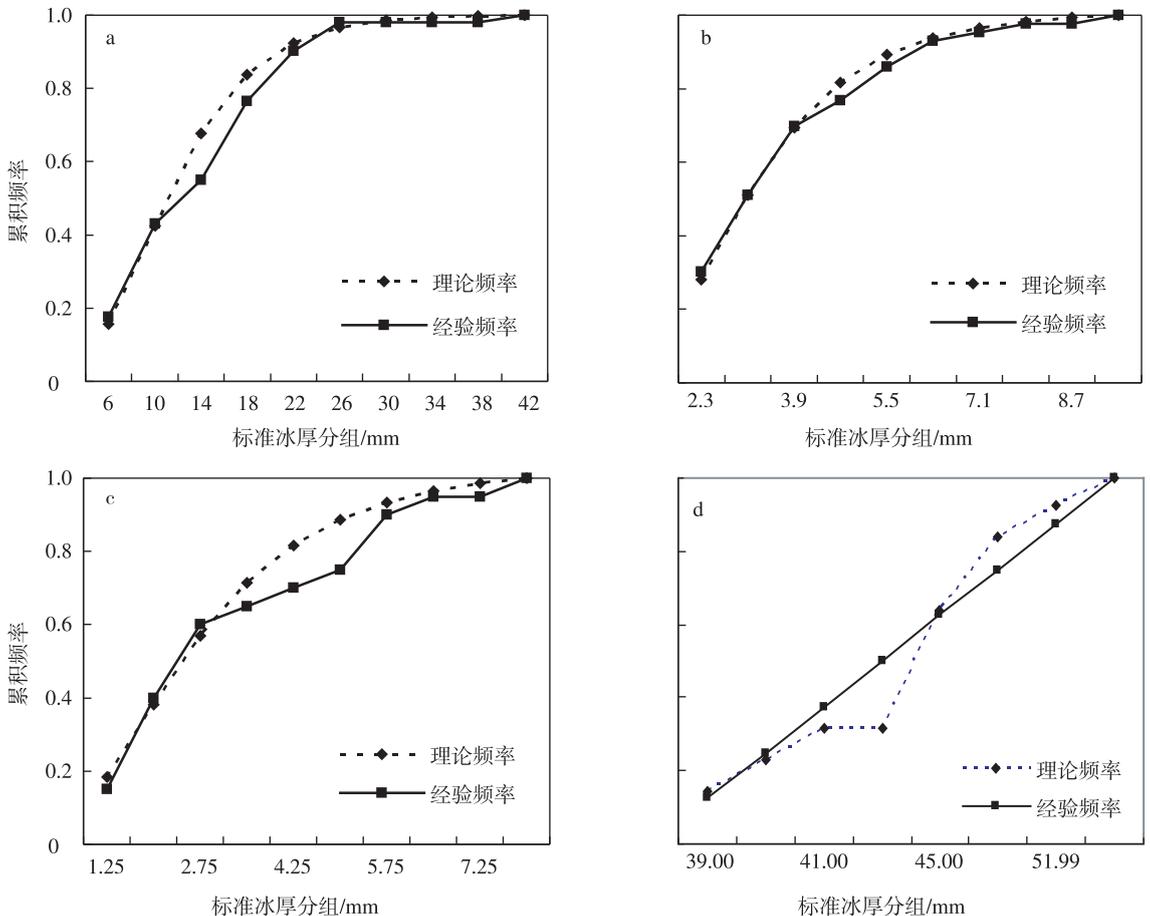


图 1 GEV 概率分布模型的拟合曲线 a. 庐山;b. 威宁;c. 三穗;d. 二郎山南北向

Fig. 1 The simulated curve of GEV (a, b, c and d are Lushan, Weining, Sansui and NS Erlangshan, respectively)

表 1 GEV 模型的拟合精度及其与 Gumbel 模型的对比

Table 1 The simulated accuracy of GEV and the comparison with Gumbel

站名	样本容量	GEV			Gumbel		
		R	K-S	S_s	R	K-S	S_s
庐山	51	0.987 9	0.127 8	0.048 4	0.987 3	0.130 7	0.049 3
威宁	43	0.997 4	0.048 7	0.020 9	0.997 9	0.032 7	0.021 7
峨眉山	39	0.993 4	0.092 1	0.043 2	0.993 0	0.094 1	0.045 4
金佛山	37	0.996 9	0.052 6	0.025 6	0.997 0	0.051 8	0.025 1
绿葱坡	37	0.998 8	0.028 6	0.018 6	0.996 7	0.085 6	0.032 1
三穗	20	0.981 2	0.136 7	0.064 5	0.980 8	0.118 1	0.0574
二郎山南北向	8	0.972 8	0.184 0	0.078 2	0.969 0	0.198 9	0.081 0
二郎山东西向	8	0.974 0	0.119 0	0.070 4	0.964 9	0.119 5	0.076 7

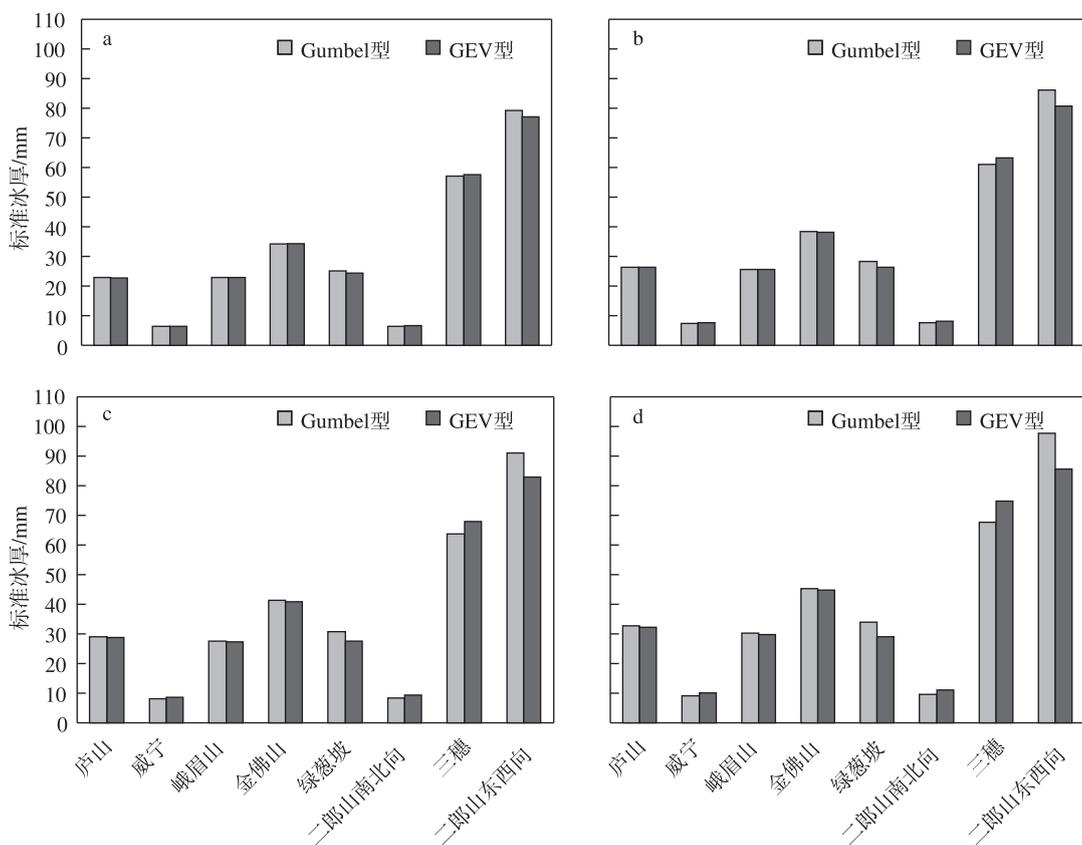


图 2 GEV 和 Gumbel 模型的重现期冰厚结果对比 a. 重现期 = 15 a; b. 重现期 = 30 a; c. 重现期 = 50 a; d. 重现期 = 100 a

Fig. 2 The comparison of GEV and GPD in icing thickness of Return period a. Return period = 15 years; b. Return period = 30 years; c. Return period = 50 years; d. Return period = 100 years

通过两种模型的拟合精度和重现期冰厚极值的对比分析发现,两种模型的计算结果差异不大,总体来说,GEV 模型理论与实际频率的相关系数大多略高于 Gumbel 模型,K-S 统计量和均方误差基本小于 Gumbel 模型,可见 GEV 分布的拟合精度略高于

Gumble 分布;鉴于 GEV 模型涵盖了三种极值分布,在使用上不需要多次计算,也避免了使用单一模型的不足,所以,对于采用 AM 抽样方法的极值模型来说,选用 GEV 分布拟合导线覆冰冰厚极值比 Gumble 分布更具有实用性,计算的重现期冰厚极值

精度略高。

2.2 GPD 概率分布型的拟合及其门限值的确定

2.2.1 门限值的选取

超门限峰值 (POT) 抽样方法的关键是确定门限值, 根据 1.2.3 介绍的方法, 利用 7 个站资料, 选取不同的门限值, 得到相应超门限极值次数序列。由于气象站资料的年序列较长, 因此将各站的年交叉率 λ 控制在 1~6 的范围内, 而二郎山观冰站由于序列较短且覆冰严重, 一年内出现多次极值的可能性更大, 因此将年交叉率 λ 控制在 10 以内。对不同的超门限极值次数序列进行 Poisson 分布拟合优度检验, 发现三穗站、金佛山站和二郎山站全部通过了 0.05 信度的显著性检验, 而庐山站、威宁站、峨眉山站和绿葱坡站除了在 $\beta=3, \beta=0.2, 0.5, \beta=9, 10, \beta=7$ 时序列没有通过检验而其他均通过检验, 证明了超门限序列的年发生次数符合泊松分布, 没有通过检验的序列可能是由于门限值较低, 超过门限值的年样本数较多, 样本之间的独立性受到了制约。再利用公式 (19) 计算不同门限值下的 Hill 估计量,

考察 Hill 估计量随门限值的演变情况, 图 3 给出了庐山站、威宁站、三穗站和二郎山南北向的 Hill 图 (其他站方法相同, 图略)

通过泊松分布的 χ^2 检验, 结合 Hill 图中 Hill 估计量随门限变化趋于稳定的点, 可以确定各站的最佳门限值: 庐山站为 8 mm 左右, 威宁站为 2.1 mm 左右, 峨眉山站为 11 mm 左右, 绿葱坡站为 8 mm 左右, 金佛山站为 14 mm 左右, 二郎山南北向为 19 mm 左右, 二郎山东西向为 29 mm 左右, 三穗站在取不同门限值下的 Hill 估计量波动较大, 结合考虑超门限发生次数符合泊松分布, 取三穗站的门限值为 1 mm。可见, 通过超门限值发生次数是否符合泊松分布, 结合 Hill 图, 可以基本确定 POT 抽样的最佳门限。

2.2.2 覆冰极值 GPD 概率分布型的拟合及其重现期极值估计

根据以上确定的门限值, 利用 7 个站的逐日冰厚资料, 进行 GPD 模型的拟合, 图 4 给出了各站 GPD 模型拟合精度的三种检验指标及其与 GEV 模

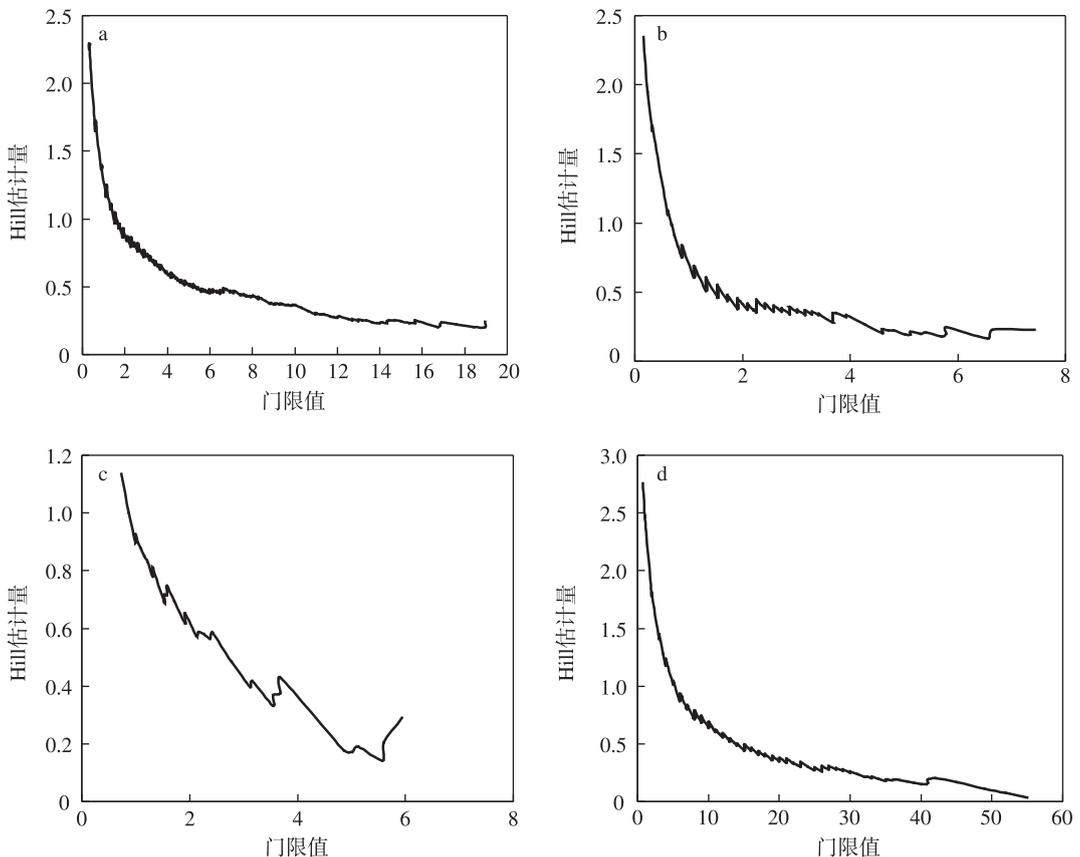


图 3 逐日冰厚的 Hill 图 a. 庐山; b. 威宁; c. 三穗; d. 二郎山南北向

Fig. 3 Hill plot of daily icing thickness (a, b, c and d are Lushan, Weining, Sansui and NS Erlangshan, respectively)

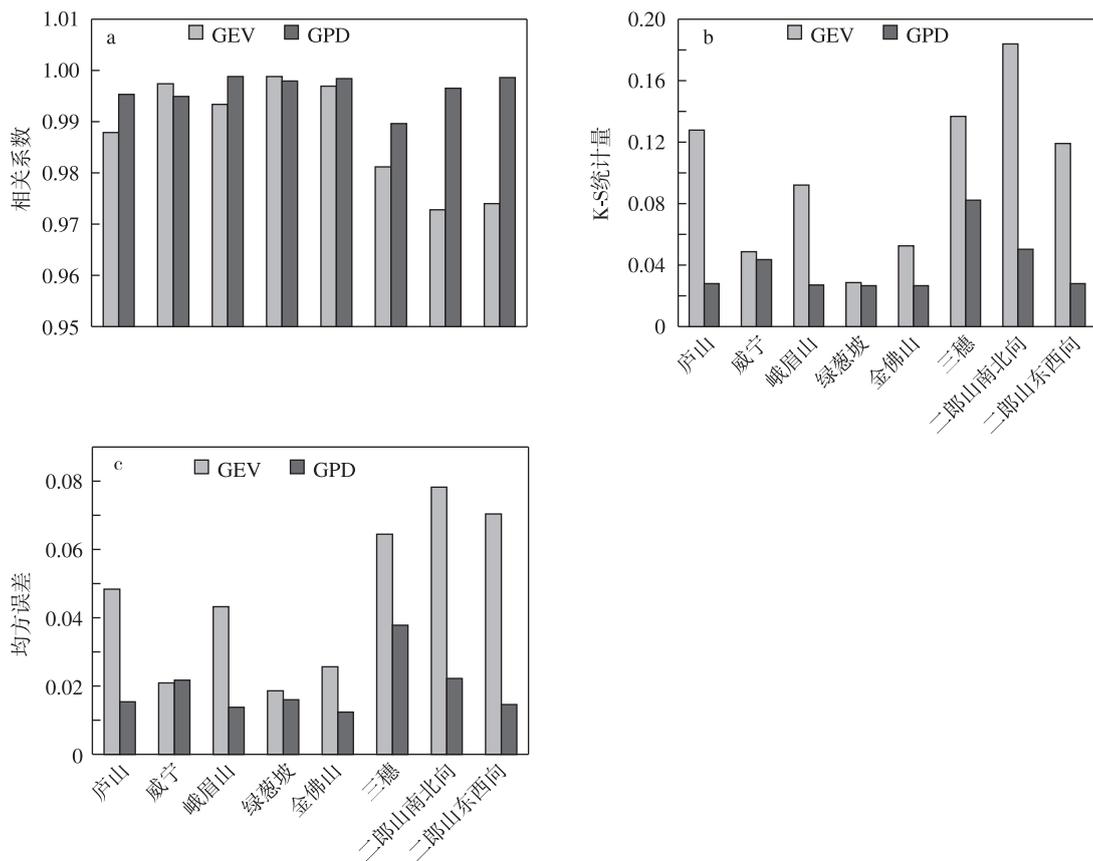


图 4 GPD 和 GEV 模型拟合精度比较

Fig. 4 The comparison of GPD and GEV in the numerical simulations

型的对比,从图 4 中可以看出威宁站和绿葱坡站 GEV 模型的相关系数略高于 GPD 模型,但两者差异不大,其他各站 GPD 模型的相关系数均高于 GEV 模型;7 个站的 K-S 统计量 GPD 模型均明显小于 GEV 模型;绝大多数站点 GPD 模型的均方误差明显小于 GEV 模型。可见,GPD 模型的拟合精度较之 GEV 模型均有大幅度的提高。

图 5 为庐山站 $\beta = 8 \text{ mm}$ 、威宁站 $\beta = 2.1 \text{ mm}$ 、三穗站 $\beta = 1 \text{ mm}$ 和二郎山站南北向 $\beta = 19 \text{ mm}$ 的 GPD 模型累积频率曲线图(其他站图略)。从图 5 中可以看出各站的理论累积频率曲线与实测累积频率曲线非常接近,尤其是对于样本很少的二郎山站,拟合效果明显优于 GEV 模型。

图 6 比较了 GPD 和 GEV 模型的 15 a、30 a、50 a 和 100 a 重现期的冰厚极值。可以看出,与 GEV 模型相比,随着重现期的增加,两种模型计算的冰厚极值之间的差异也变大;与 GEV 模型相比,庐山站、威宁站、峨眉山站、绿葱坡站、金佛山站和三穗站四种重现期计算的结果差异不大,而样本容量较少的二郎山站两种模型的计算结果差异较大。由

于 GPD 模型采用 POT 抽样,有效增大了极值的样本量,其概率分布模型的拟合精度又高于 GEV 模型,可以认为由 GPD 模型估算的重现期冰厚极值比 GEV 模型更为可靠。

2.3 样本长度对重现期极值估计的影响

样本长度对重现期冰厚极值估计的影响是重现期冰厚极值估计中经常遇到的问题,也关系到极值统计推断是否稳健。为此,计算不同重现期水平下冰厚极值随样本长度改变的变化,选取样本长度较多的庐山站、威宁站、峨眉山站、绿葱坡站和金佛山站的冰厚年极值,采用正推方式以 1 递增依次建立 GEV 和 GPD 模型,考虑到 GPD 建模还受门限值的影响,且在序列较短时,高门限值会导致短序列的样本容量过少,无法分析极值估计随样本长度的变化,故适当降低各站最佳门限值,取庐山站为 5.5 mm、威宁站为 1 mm、峨眉山站为 5 mm、绿葱坡站为 8 mm、金佛山站为 7 mm。

图 7 给出了庐山站和威宁站 GEV 模型和 GPD 模型 100、50、30、15 a 一遇重现期冰厚极值随样本长度的变化情况。由图 7 可见,当样本容量较少时,

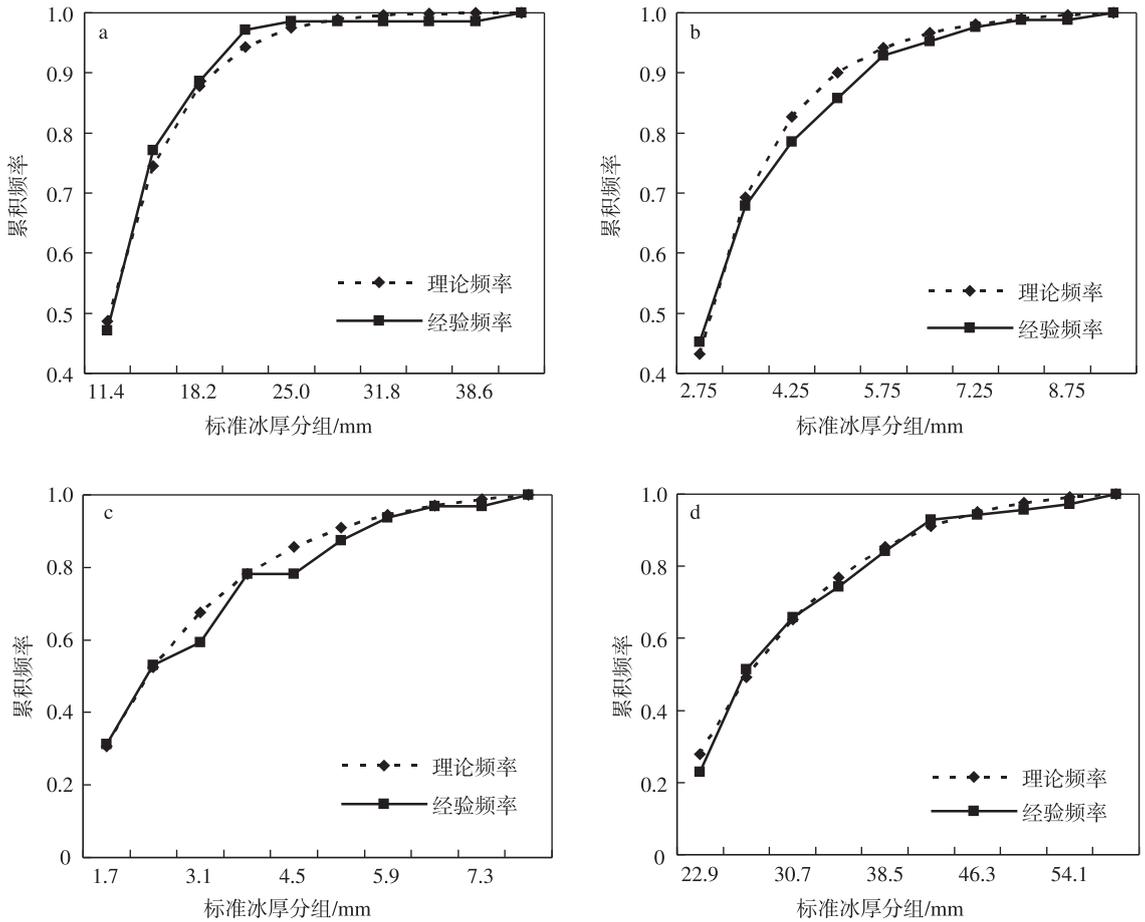


图 5 GPD 概率分布模型的拟合曲线 a. 庐山; b. 威宁; c. 三穗; d. 二郎山南北向

Fig. 5 The simulated curve of GPD (a, b, c and d are Lushan, Weining, Sansui and NS Erlangshan, respectively)

GEV 模型重现期冰厚极值变化幅度较大, 庐山站在样本容量少于 29 时, 百年一遇重现期冰厚极值在 19.5 ~ 50.2 mm 之间变动, 大于 29 个样本时, 变动范围为 28.8 ~ 32.6 mm; 威宁站在样本容量少于 28 时, 百年一遇重现期冰厚极值在 9.3 ~ 13.0 mm 之间变动, 大于 28 个样本时, 变动范围为 10.0 ~ 11.9 mm, 表明随着序列长度的增加, 重现期冰厚极值趋向稳定, 且趋于稳定的样本长度与重现期基本无关, 一般样本容量达到 30 a 左右时, GEV 模型重现期冰厚极值的估计趋于稳定, 值得注意的是, 随着重现期的增大, 极值的变化幅度也变大, 而在样本容量较短时, 波动明显, 对重现期冰厚极值的结果影响较大。因此, 在采用 AM 方法抽样建立极值模型时, 样本长度一般要 30 a (其他站结果类似, 图略)。

GPD 模型重现期冰厚极值随样本长度的变化小于 GEV 模型, 庐山站在样本容量少于 20 时, 百年一遇重现期冰厚极值在 29.1 ~ 49.2 mm 之间变动, 大于 20 个样本时, 变动范围为 38.1 ~ 43.9 mm; 威宁站在样本容量少于 23 时, 百年一遇重现期冰厚极

值在 9.8 ~ 16.2 mm 之间变动, 大于 23 个样本时, 变动范围为 8.2 ~ 9.9 mm, 表明随着序列长度的增加 GPD 模型重现期冰厚极值趋向稳定, 一般在样本容量达到 25 a 左右时, 趋于稳定, 在样本容量较短时, 波动明显, 对重现期冰厚极值的结果影响较大。可见样本长度对 GPD 模型对重现期冰厚极值估计的影响略低于 GEV 模型, 重现期冰厚极值推算的稳定性相对稍高一些 (其他站结果类似, 图略)。

进一步计算不同样本长度下, GEV 和 GPD 模型形状参数 (k)、尺度参数 (α) 以及重现期冰厚极值估计的标准差, 探讨样本长度对重现期冰厚极值估计的影响。通过不同样本长度各站两种模型统计估计量标准差的对比, 发现大部分站 GPD 模型参数标准差和重现期冰厚极值标准差均小于 GEV 模型, 威宁站由于结冰较弱两种模型计算的标准差差异较小。表明 GPD 模型重现期冰厚极值随样本长度变化的幅度小于 GEV 模型, 极值估计稳定性较强, 是较短序列样本下进行覆冰冰厚极值估计的较好方法。

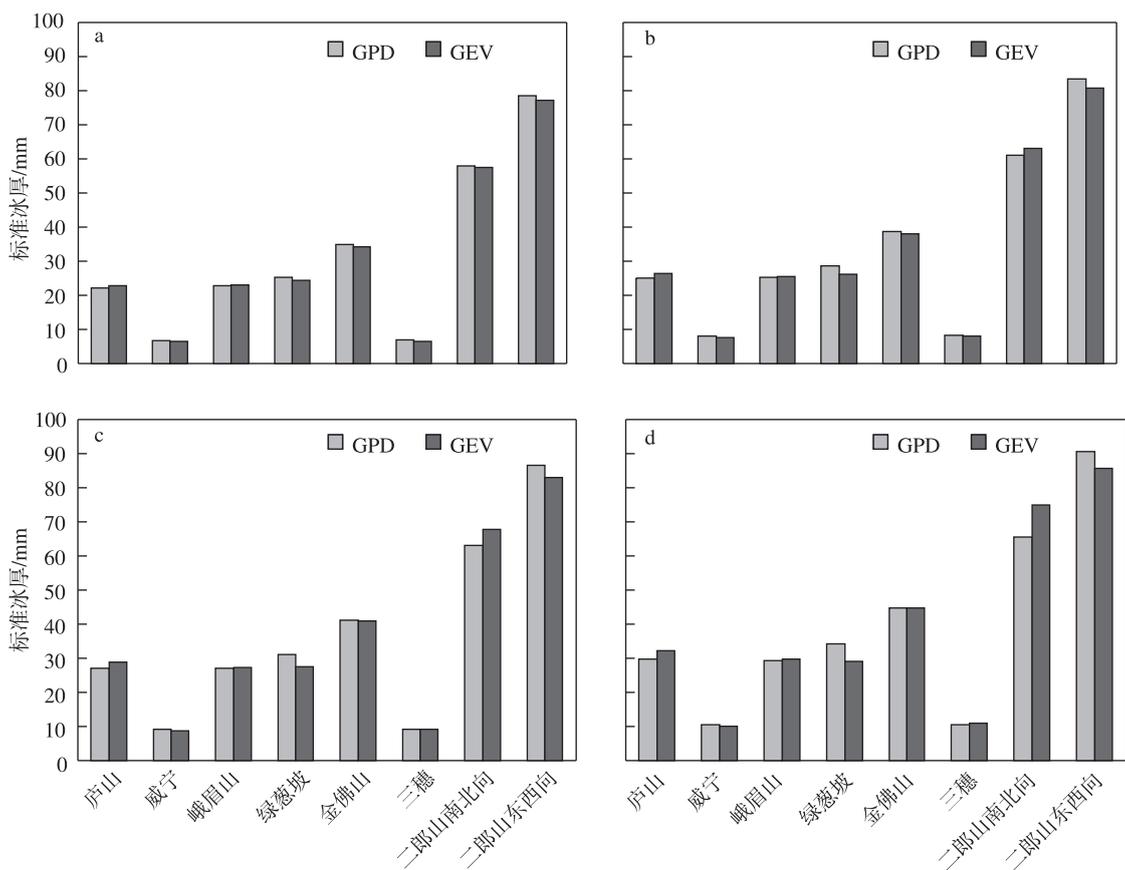


图6 GPD和GEV模型的重现期冰厚结果比较 a. 重现期=15 a; b. 重现期=30 a; c. 重现期=50 a; d. 重现期=100 a

Fig. 6 The comparison of GPD and GEV in icing thickness of Return period a. Return period = 15 years; b. Return period = 30 years; c. Return period = 50 years; d. Return period = 100 years

3 结论

本文对庐山站、威宁站、峨眉山站、绿葱坡站、金佛山站、三穗站和二郎山站导线覆冰冰厚数据进行了GEV和GPD模型的拟合,对比不同模型的拟合精度及其不同重现期下的冰厚极值,并对不同模型的计算结果进行了比较,同时分析了GPD模型和GEV模型的参数和重现期冰厚极值随样本长度的变化情况,得到如下结论:

(1)利用超门限次数的泊松分布拟合检验,结合Hill图解,可以较好地解决超门限抽样中门限值的确定问题,为广义帕雷托分布在极值研究中的应用提供了重要技术支持。

(2)GEV模型和GPD模型对导线覆冰冰厚极值概率分布模拟的适用性表明,我国西南地区的导线覆冰冰厚极值的概率分布均符合GEV模型和GPD模型,其拟合精度高于以常用的Gumbel模型。而GPD模型与GEV模型拟合精度的对比表

明,GPD模型优于GEV模型,故GPD模型估算的重现期冰厚极值较GEV模型更为可靠,两种模型计算的重现期冰厚极值差异一般低于10%,个别短序列观冰站差异不超过20%。

(3)样本长度对GPD模型参数估计稳定性的影响小于GEV模型,一般样本容量达到25 a左右时,GPD模型重现期冰厚极值的估计趋于稳定,可以作为短序列下估计导线覆冰极值的较好方法。

参考文献:

- [1] 聂国一. 架空送电线路导线冰载的探讨[J]. 电力建设, 1992(7):46-54.
- [2] 谭冠日,王冰,秦剑,等. 导线积冰极值的推算[J]. 电力技术, 1983(2):8-14.
- [3] Jenkinson A F. The frequency distribution of the annual maximum(or minimum)values of meteorological elements[J]. Quart J Roy Meteor Soc, 1955, 81:158-171.
- [4] 丁裕国,申红艳,江志红,等. 气候概率分布理论及其应用新进展[J]. 气象科技, 2009, 37(3):257-262.

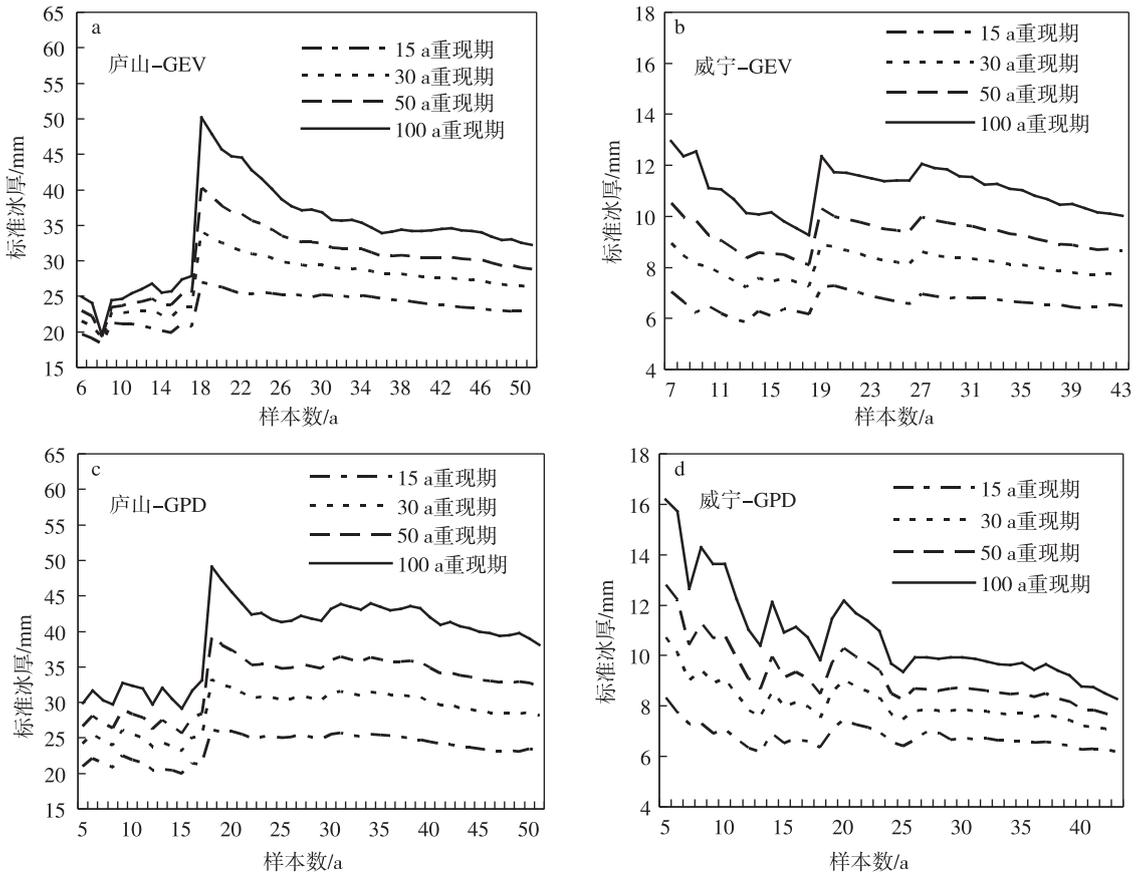


图 7 两种模型重现期冰厚极值随样本长度的变化

Fig. 7 The change of icing thickness of return period in two models under different sample lengths

- [5] 王大钧,程炳岩,丁裕国. 气候平均状况的变化对气候极值出现概率的影响[J]. 南京气象学院学报,2003,26(2):263-269.
- [6] 吴香华,秦伟良,王新蕾,等. 用最绝对偏差方法(LAD)估计极值分布参数的探讨[J]. 气象科学,2006,26(3):260-264.
- [7] Pickands J. Statistical inference using extreme order statistics. The Annals of Statistics[J]. Ann Stat,1975,3(1):119-131.
- [8] 江志红,丁裕国,朱莲芳,等. 利用广义帕雷托分布拟合中国东部日极端降水的试验[J]. 高原气象,2009,28(3):573-580.
- [9] 苏布达,姜彤,董文杰. 长江流域极端强降水分布特征的统计拟合[J]. 气象科学,2008,28(6):625-629.
- [10] 丁裕国. 探讨灾害规律的理论基础——极端气候事件概率[J]. 气象与灾害研究,2000,29(1):44-50.
- [11] 蔡敏,丁裕国,江志红. 我国东部极端降水时空分布及其概率特征[J]. 高原气象,2007,26(2):309-317.
- [12] 刘聪,秦伟良,江志红. 基于广义极值分布的设计基本风速及其置信限计算[J]. 东南大学学报:自然科学版,2006,36(2):331-334.
- [13] 吴香华,牛生杰,祁莉,等. 中国南方6省1月份降水和极端低温的分布特征[J]. 大气科学学报,2009,32(6):807-814.
- [14] 蔡敏,丁裕国,江志红. L-矩估计方法在极端降水研究中的应用[J]. 气象科学,2007,27(6):597-603.
- [15] 程炳岩,丁裕国,张金玲,等. 广义帕雷托分布在重庆暴雨强降水研究中的应用[J]. 高原气象,2008,27(5):1004-1009.
- [16] Coles S. An introduction to statistical modeling of extreme values[M]. London:Springer Verlag,2001.
- [17] Hosking J R,Wallis J R. Parameter and quantile estimation for the generalized Pareto distribution[J]. Technometrics,1987,29:339-349.
- [18] Coles S,Walshaw G D. Directional modelling of extreme wind speeds[J]. Appl Stat,1994,43:139-157.
- [19] Hosking J R,Wallis J R,Wood E F. Estimation of the generalized extreme value distribution by the method of probability-weighted moments[J]. Technometrics,1985,27:251-261.
- [20] Zhai P M,Sun A J,Ren F M,et al. Changes of climate extremes in China[J]. Climatic Change,1999,42:203-218.
- [21] 马开玉,张耀存,陈星. 现代应用统计学[M]. 北京:气象出版社,2004:91-92.
- [22] 谢志清,姜爱军,杜银,等. 长江三角洲强降水过程年极值分布特征研究[J]. 南京气象学院学报,2005,28(2):268-274.
- [23] 王善序. 洪水超定量系列频率分析[J]. 人民长江,1999,30(8):23-25.
- [24] 许月萍,任立新,黄艳,等. 水文极值计算及其不确定性[J]. 应用基础与工程科学学报,2009,17(2):172-178.
- [25] 丁裕国,江志红. 极端气候研究方法导论[M]. 北京:气象出版社,2009:84,92-93.

(责任编辑:张福颖)