

# 气象要素极值统计的稳定性、 多母体现象和区域极值\*

董双林

(北京应用气象研究所, 100081)

## 提 要

本文以 Gumbel 分布为例说明极大似然法估计参数可明显减小抽样误差的影响, 说明极值样本中个别极大点数据是引起统计结果不稳定的根本原因; 指出它们来自另外的统计母体, 即存在多母体现象。对包含极大点数据的稀遇极值用区域极值进行统计, 显著提高了结果的稳定性。为方便设计应用, 还导出了区域极值的单站发生概率。

**关键词:** 气象要素, 极值, 稳定性。

## 1 引言

根据实测样本, 拟合极值分布, 推断多年一遇的再现期值, 是极值统计的主要内容。它在能源、交通、水利、建筑、航天、军事等各方面有着广泛应用。气象要素极值不仅是环境条件的主要组成部分, 而且由于气候是一种可供开发利用的资源而日益显示其重要性。但由于气象要素极值的成因复杂, 观测资料年限较短, 抽样误差较大, 给极值分布线型的确定、参数估计方法的评价等提出了一些亟待解决的问题<sup>1)</sup>。其中一些问题集中反映在如何处理实测样本中的极大点数据上。如选择多参数分布线型来兼顾极大点数据和普通数据<sup>1)</sup>, 用高分位值的拟合优度作为确定分布线型的一个标准, 选择较充分反映极大点数据的影响、确保安全的参数估计方法……所有这些虽在一定程度上反映了真实极大点数据的影响, 但因高分位值的抽样误差较大, 使模型确定、参数估计等受到抽样误差的较大干扰。结果再现期值的空间分布出现虚假的不连续, 特别是随着样本资料中是否包含稀遇的极大点数据, 再现期值变化显著。因此, 极值统计的可靠性和稳定性一直是困惑人们的难题。而且由于存在气候变化, 使依据数十年资料推断万年一遇的极值不能令人信服。因此怀疑极值统计方法的可用性。例如用估计物理上可能最大暴雨和洪水来代替暴雨和洪水极值长再现期值的估计<sup>2)</sup>。只因估计物理上可能最大暴雨还存在较大的困难, 所以极值统计还是常用的方法。

本文是在对抽样误差计算、超拟合现象及其抑制<sup>2)</sup>、参数估计方法评价<sup>[3]</sup>、分布

\* 1989年9月28日收到原稿, 1991年7月2日收到最后修改稿。

1) 董双林, 1987, 极值统计的若干基本问题。

2) 董双林, 1987, 超拟合现象及其抑制, 全国第二届计算物理年会材料。

线型选择标准及途径<sup>1)</sup>、抽样误差较大时极大似然估计方法<sup>[4, 6]</sup>等初步研究基础上, 确认极大点数据与普通数据来自不同的统计母体, 对单站稀遇的极大点数据统计区域极值, 并用区域极值在单站的发生概率将单站长再现期值与区域极值的短再现期值联系起来, 从而显著提高了统计结果的稳定性、可靠性。

## 2 抽样误差及不同参数估计方法的响应函数

以常用的极值渐连分布 Gumbel 分布

$$F(x) = \exp(-\exp(-(x-x_0)/B)) \quad -\infty < x < \infty \quad (1)$$

为例, 次序统计量  $x_k$  的均值和标准差为

$$M(x_k) = n \binom{n-1}{k-1} \sum_{j=1}^{n-k+1} \binom{n-k}{j-1} (-1)^{j-1} \left[ \nu + \frac{x_0}{B} + \ln(k+j-1) \right] \frac{B}{k+j-1} \quad (2)$$

$$\sigma(x_k) = \left\{ n \binom{n-1}{k-1} \sum_{j=1}^{n-k+1} \binom{n-k}{j-1} (-1)^{j-1} \left[ \left( \ln(k+j-1) + \nu + \frac{x_0}{B} \right)^2 + \frac{\pi^2}{6} \right] \right\}^{1/2}$$

$$\frac{B^2}{k+j-1} \left[ n \binom{n-1}{k-1} \sum_{j=1}^{n-k+1} \binom{n-k}{j-1} (-1)^{j-1} \frac{B}{k+j-1} \left( \ln(k+j-1) + \nu + \frac{x_0}{B} \right)^2 \right]^{1/2} \quad (3)$$

当  $n$  大而  $k$  小时, 计算精度受机器字长的限制。计算  $M(x_k)$  和  $\sigma(x_k)$  的另一种方法是 Monte Carlo 模拟。给定母体参数  $B (= 1.50)$ 、 $x_0 (= 35.0)$ , 抽取容量为  $n (= 25)$  的  $M (= 100)$  个个例。用方法  $l (= J, R, G, S)$ , 分别表示矩法、回归法、Gumbel 法和极大似然法) 估计参数  $B_{li}$ ,  $x_{lio}$ 。真值  $x_k$ 、抽样值  $x_{ik}$ 、拟合值  $x_{lik}$  分别为

$$x_k = x_0 - B \ln \left( \ln \frac{(n+1)}{k} \right) \quad k=1, 2, \dots, n \quad (4)$$

$$x_{ik} = x_0 - B \ln \left( -\ln(P_{li}) \right) \quad i=1, 2, \dots, M \quad (5)$$

$$x_{lik} = x_{lio} - B_{li} \ln \left( \ln \frac{(n+1)}{k} \right) \quad l=J, R, G, S \quad (6)$$

这里  $P_{li}$  是对分布函数的均匀抽样<sup>[8]</sup>。则抽样误差  $\sigma_k$ 、拟合样本误差  $e_{ik}$ 、与真值误差 (估计误差)  $\sigma_{lik}$  的平方分别为

$$\sigma_k^2 = \sum_{i=1}^M (x_{ik} - x_k)^2 / (M-1) \quad (7)$$

$$e_{ik}^2 = \sum_{i=1}^M (x_{lik} - x_{ik})^2 / (M-1) \quad (8)$$

$$\sigma_{lik}^2 = \sum_{i=1}^M (x_{lik} - x_k)^2 / (M-1) \quad (9)$$

1) 董双林, 1988, 极值分布的线型问题, 北京气象学会“五四”报告会交流材料。

令

$$r_{ik} = \frac{\sum_{i=1}^M (x_{i,k} - \bar{x}_{i,k})(x_{i,k} - \bar{x}_k)}{\left[ \sum_{i=1}^M (x_{i,k} - \bar{x}_{i,k})^2 \sum_{i=1}^M (x_{i,k} - \bar{x}_k)^2 \right]^{1/2}} \quad (10)$$

则

$$\sigma_{ik}^2 = \sigma_k^2 + e_{ik}^2 + 2r_{ik}\sigma_k e_{ik} \quad (11)$$

上式表明,估计误差 $\sigma_{ik}$ 与抽样误差 $\sigma_k$ 、拟合样本误差 $e_{ik}$ 及它们的相关系数 $r_{ik}$ 有关。模拟计算结果(表1)表明,高分位点 $e_{ik}$ 较小且 $e_{ik}$ 最小的回归法, $\sigma_{ik}$ 却最大。说明拟合了较多的抽样误差,存在超拟合现象。高分位点 $e_{ik}$ 最大的极大似然法, $\sigma_{ik}$ 却最小,主要是极大似然法的 $r_{ik}(<0)$ 的绝对值最大(图1),较大的拟合样本误差起到了抵消抽样误差的作用。若定义方法 $l$ 对抽样误差的响应函数(图2)为

$$R_{ik} = \sigma_{ik} / \sigma_k \quad (12)$$

则可见,极大似然法与真值的误差小于抽样误差,且高分位点 $R_{ik}$ 更小。由此可见,由于存在抽样误差,不能单从对样本的拟合程度,特别是高、低分位点的拟合程度来选择分布线型和参数估计方法。

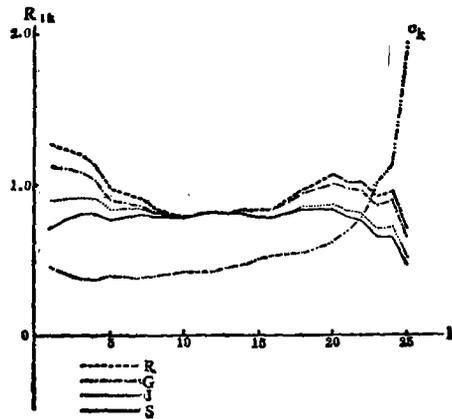
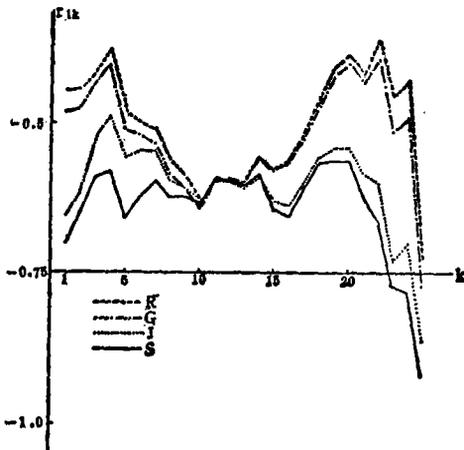


图1 不同参数估计方法的 $r_{ik}$ 随 $k$ (或分位点)的变化 图2 抽样误差及不同参数估计方法的响应函数

### 3 单站极值再现期值的稳定性与多母体现象

用Gumbel分布对全国2180个测站日降水量年极值(多数测站记录年限为20—30a)统计表明,很多测站由于出现由稀遇暴雨造成的极大点数据,再现期值很不稳定,且与周围其它测站有显著差别。表2给出1975年8月河南暴雨的上蔡、西平,1963年8月河北暴雨的邯郸,1971年8月山东暴雨的微山等4站再现期为1000a的再现期值的不稳定程度 $D_{1000}$ 。序列中包含和不包含极大点数据,再现期值变化很大。而不稳定度随再现

表1 模拟的抽样误差 $\sigma_s$ 及不同参数估计方法对抽样误差的响应

| 次序<br>统计量<br>$x_k$ 的<br>序号<br>$k$ | 模拟<br>的抽<br>样误<br>差 $\sigma_s$ | 矩 法                         |                            | 回 归 法                       |                            | Gumbel法                     |                            | 极大似然法                       |                            |
|-----------------------------------|--------------------------------|-----------------------------|----------------------------|-----------------------------|----------------------------|-----------------------------|----------------------------|-----------------------------|----------------------------|
|                                   |                                | 拟合样<br>本误差<br>$\sigma_{js}$ | 与真值<br>误差<br>$\sigma_{js}$ | 拟合样<br>本误差<br>$\sigma_{js}$ | 与真值<br>误差<br>$\sigma_{js}$ | 拟合样<br>本误差<br>$\sigma_{js}$ | 与真值<br>误差<br>$\sigma_{js}$ | 拟合样<br>本误差<br>$\sigma_{js}$ | 与真值<br>误差<br>$\sigma_{js}$ |
| 1                                 | 0.459                          | 0.520                       | 0.409                      | 0.615                       | 0.679                      | 0.539                       | 0.512                      | 1.365                       | 0.330                      |
| 2                                 | 0.406                          | 0.440                       | 0.371                      | 0.525                       | 0.501                      | 0.460                       | 0.447                      | 0.309                       | 0.311                      |
| 3                                 | 0.373                          | 0.339                       | 0.349                      | 0.452                       | 0.450                      | 0.387                       | 0.405                      | 0.241                       | 0.302                      |
| 4                                 | 0.366                          | 0.278                       | 0.334                      | 0.374                       | 0.412                      | 0.317                       | 0.375                      | 0.216                       | 0.298                      |
| 5                                 | 0.391                          | 0.225                       | 0.325                      | 0.359                       | 0.382                      | 0.306                       | 0.352                      | 0.216                       | 0.297                      |
| 6                                 | 0.381                          | 0.228                       | 0.319                      | 0.330                       | 0.359                      | 0.281                       | 0.336                      | 0.215                       | 0.298                      |
| 7                                 | 0.377                          | 0.222                       | 0.316                      | 0.303                       | 0.342                      | 0.263                       | 0.325                      | 0.220                       | 0.302                      |
| 8                                 | 0.393                          | 0.229                       | 0.316                      | 0.284                       | 0.331                      | 0.251                       | 0.319                      | 0.230                       | 0.307                      |
| 9                                 | 0.400                          | 0.235                       | 0.318                      | 0.273                       | 0.325                      | 0.248                       | 0.318                      | 0.241                       | 0.313                      |
| 10                                | 0.416                          | 0.222                       | 0.323                      | 0.237                       | 0.325                      | 0.223                       | 0.321                      | 0.229                       | 0.322                      |
| 11                                | 0.410                          | 0.237                       | 0.331                      | 0.240                       | 0.330                      | 0.236                       | 0.329                      | 0.241                       | 0.331                      |
| 12                                | 0.422                          | 0.225                       | 0.340                      | 0.228                       | 0.341                      | 0.225                       | 0.341                      | 0.223                       | 0.342                      |
| 13                                | 0.444                          | 0.256                       | 0.353                      | 0.258                       | 0.357                      | 0.259                       | 0.358                      | 0.246                       | 0.355                      |
| 14                                | 0.455                          | 0.260                       | 0.367                      | 0.270                       | 0.378                      | 0.271                       | 0.379                      | 0.240                       | 0.369                      |
| 15                                | 0.495                          | 0.281                       | 0.385                      | 0.309                       | 0.405                      | 0.306                       | 0.404                      | 0.249                       | 0.385                      |
| 16                                | 0.524                          | 0.304                       | 0.405                      | 0.342                       | 0.437                      | 0.335                       | 0.434                      | 0.266                       | 0.403                      |
| 17                                | 0.532                          | 0.362                       | 0.428                      | 0.415                       | 0.475                      | 0.402                       | 0.468                      | 0.309                       | 0.423                      |
| 18                                | 0.539                          | 0.387                       | 0.456                      | 0.453                       | 0.520                      | 0.433                       | 0.508                      | 0.325                       | 0.446                      |
| 19                                | 0.571                          | 0.399                       | 0.488                      | 0.470                       | 0.572                      | 0.440                       | 0.555                      | 0.338                       | 0.473                      |
| 20                                | 0.604                          | 0.458                       | 0.525                      | 0.536                       | 0.633                      | 0.498                       | 0.610                      | 0.398                       | 0.504                      |
| 21                                | 0.694                          | 0.494                       | 0.571                      | 0.604                       | 0.706                      | 0.552                       | 0.675                      | 0.439                       | 0.541                      |
| 22                                | 0.783                          | 0.510                       | 0.627                      | 0.603                       | 0.796                      | 0.537                       | 0.756                      | 0.517                       | 0.587                      |
| 23                                | 1.000                          | 0.588                       | 0.700                      | 0.656                       | 0.912                      | 0.571                       | 0.860                      | 0.677                       | 0.646                      |
| 24                                | 1.128                          | 0.873                       | 0.805                      | 0.838                       | 1.076                      | 0.773                       | 1.007                      | 1.050                       | 0.730                      |
| 25                                | 1.949                          | 1.749                       | 0.985                      | 1.203                       | 1.355                      | 1.335                       | 1.257                      | 2.206                       | 0.876                      |
| $\Sigma( )^*$                     | 11.276                         | 6.639                       | 5.680                      | 6.189                       | 8.703                      | 5.714                       | 7.781                      | 8.358                       | 4.960                      |

表2 日降水量单站极值再现期值及其不稳定程度\*  $D_{1000} = \frac{x_{1000} - x'_{1000}}{x'_{1000}}$ 

| 站名 | 样本量<br>n | 矩法         |            | 回归法        |            | Gumbel法    |            | 极大似然法      |            |
|----|----------|------------|------------|------------|------------|------------|------------|------------|------------|
|    |          | $x_{1000}$ | $D_{1000}$ | $x_{1000}$ | $D_{1000}$ | $x_{1000}$ | $D_{1000}$ | $x_{1000}$ | $D_{1000}$ |
| 上蔡 | 20       | 873        | 284%       | 1422       | 453%       | 1005       | 294%       | 414        | 67%        |
|    | 19       | 227        |            | 257        |            | 255        |            | 248        |            |
| 西平 | 20       | 743        | 210%       | 1150       | 316%       | 864        | 218%       | 403        | 82%        |
|    | 19       | 239        |            | 276        |            | 271        |            | 222        |            |
| 邯郸 | 26       | 548        | 133%       | 793        | 199%       | 619        | 138%       | 307        | 41%        |
|    | 25       | 235        |            | 265        |            | 260        |            | 217        |            |
| 微山 | 24       | 602        | 103%       | 844        | 140%       | 686        | 107%       | 344        | 43%        |
|    | 23       | 296        |            | 351        |            | 331        |            | 240        |            |

\* $x'_{1000}$ 为剔除最大值后估计的1000年一遇再现期值。

表3 日降水量年极值序列中极大点数据的抽样误差 $e_n$ 和再现期 $N, N'$ 

| 站名 | 样本量<br>n | $\hat{\beta}$ | $\hat{x}$ | $M(x_n)$ | $\sigma(x_n)$ | $x_{max}$ | $e_n$ | $N'_n$<br>( $10^4 a$ ) | $N_n$<br>( $10^4 a$ ) |
|----|----------|---------------|-----------|----------|---------------|-----------|-------|------------------------|-----------------------|
| 上蔡 | 20       | 48.31         | 80.55     |          |               | 755       | 18.0  | 118                    | 3934                  |
|    | 19       | 25.58         | 72.06     | 162      | 32.8          | 140       | -0.7  | 0.0013                 |                       |
| 西平 | 20       | 44.05         | 69.39     |          |               | 638       | 16.5  | 39                     | 556                   |
|    | 19       | 23.20         | 62.01     | 143      | 29.7          | 152       | 0.3   | 0.0048                 |                       |
| 邯郸 | 26       | 34.84         | 66.74     |          |               | 518       | 12.8  | 42                     | 6.6                   |
|    | 25       | 22.47         | 61.85     | 147      | 28.8          | 175       | 1.0   | 0.0158                 |                       |
| 微山 | 24       | 38.02         | 81.92     |          |               | 538       | 12.3  | 16                     | 3.0                   |
|    | 23       | 23.64         | 76.77     | 164      | 30.3          | 208       | 1.4   | 0.0260                 |                       |

期的变化相对较小,如上蔡,用表3中的 $\hat{\beta}$ ,  $\hat{x}$ 值代入式(4), 计算得到,  $D_{10000} = 71\%$ ,  $D_{1000} = 60\%$ 。表3给出上述4站极大似然法估计的参数及极大点数据的抽样误差和再现期。其中 $M(x_n)$ 、 $\sigma(x_n)$ 是式(2)、(3)中令 $k=n$ 的值, 极大点数据的抽样误差为

$$e_n = (x_{max} - M(x_n)) / \sigma(x_n) \quad (13)$$

再现期为满足下式的 $N$ ,

$$x_{max} = \hat{x}_0 - \hat{\beta} \ln(-\ln(N, / (N, + 1))) \quad (14)$$

$N'_n$ 是用剔除极大点数据后的分布参数计算的 $N_n$ 。

计算 $N'_n$ 时使用近似公式

$$\ln\left(\frac{N'_n + 1}{N'_n}\right) \approx \frac{1}{N'_n}$$

以克服计算机(器)字长的限制。

对于同一母体, 抽样误差应有一定范围。例如正态母体, 样本点与均值的差超过1, 2, 3, 4.99倍 $\sigma$ 的概率分别是31.7%, 4.5%, 0.27%,  $6 \times 10^{-7}$ 。而表3中上蔡的极大点数据超过18倍 $\sigma$ , 显然大大超出了抽样误差的范围。从再现期来看, 在20年样本中抽到数千百年一遇的值, 从小概率事件一次抽样不可能发生的原理看, 同一母体是不可能出现这种现象的。

除日降水量外, 积雪深度、电线积冰直径与重量、冰雹块直径等要素极值也存在

多母体现象。其判断标准是, 给定置信水平 $\alpha$  (例如 $\alpha = 0.01$ ), 若 $N_r$  (或 $N_r^*$ )  $\geq n/\alpha$ , 就认为存在多母体现象。在单站极值统计中应剔除这些极大点数据, 而对极大点数据, 用区域极值进行统计推断。

#### 4 稀遇暴雨日降水量的区域极值

上述极大点数据是由稀遇暴雨形成的。稀遇暴雨在固定测站多年一遇, 而在一个适当的区域内每年都会发生。我们定义一定区域内所有测站单站日降水量的最大值为日降水量的区域极值。则对日降水量的区域极值的统计表明, 当去掉序列中的最大值时, 全国 37 个区域 (见图 3) 中, 再现期值改变不大于 16%, 且大多数区域小于 5% (见表 4)。

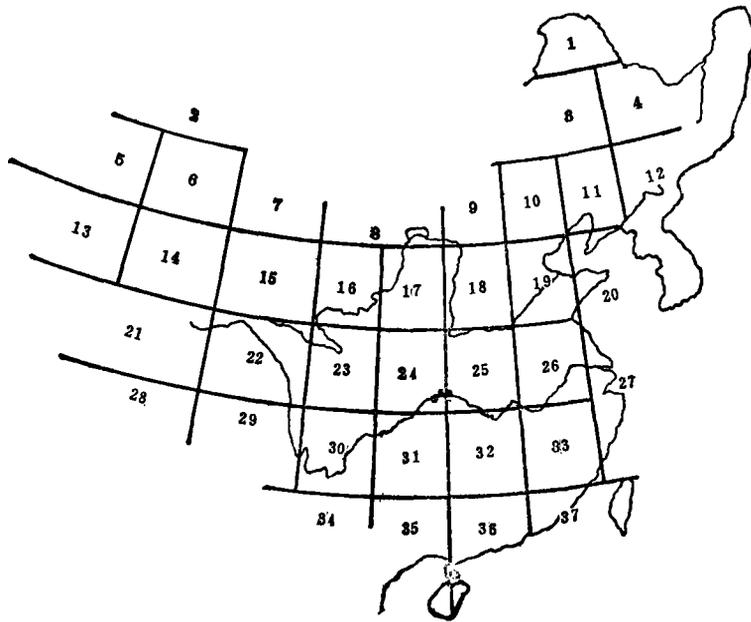


图 3 日降水量区域极值的分区

注意表 4 中的不稳定性是用区域极值 250 年再现期值计算的, 由于区域极值在单站的出现概率 (见下节) 较小, 相应的单站稀遇极值的再现期会大于 1000a。即表 4 中的不稳定性远小于表 2 中的  $D_{1000}$ , 说明稳定性有显著提高。这反过来证明了稀遇暴雨日降水量区域极值可以构成一个统计母体。

区域的划分既要考虑稀遇暴雨的范围、出现频率等气候特征; 又要使区域极值的单站发生概率与设计使用的最长再现期的乘积大于估计次序统计量分布函数均值所需的最小样本数。为了确定适当的范围, 在依气候特征分区后, 再改变范围试算对比。将图 3 中 18 和 19 区合并, 25 和 26 区合并后试算表明, 并区后的区域过大。另外, 位于区域边界附近的测站, 单站稀遇极值可能有较大计算误差, 所以应使区域的位置随所考虑的测站而变化, 在考虑气候特征基础上, 尽量使所考虑的测站位于区域的中部。

表 4 区域极值再现期值及其稳定性

| 区号 | 站数  | 样本量<br>$N_i$ | $B_i$ | $\bar{x}_i$ | $q^{x_{i..}}$ | $q^{x_{i..}}$ | $q^{x_{i..}}$ | $q^{x_{i..}}$ | $q^{x_{i..}}$ | $\frac{q^{x_{i..}} - q^{x_{i..}}}{q^{x_{i..}}}$ |
|----|-----|--------------|-------|-------------|---------------|---------------|---------------|---------------|---------------|---|
| 1  | 8   | 29           | 14.62 | 50.43       | 95            | 108           | 118           | 131           | 125           | 4.5%  |
| 2  | 19  | 27           | 8.877 | 30.58       | 57            | 65            | 72            | 80            | 75            | 6.2%  |
| 3  | 18  | 30           | 13.99 | 86.66       | 129           | 142           | 151           | 164           | 158           | 3.6%  |
| 4  | 57  | 30           | 18.52 | 96.18       | 152           | 169           | 182           | 198           | 189           | 4.5%  |
| 5  | 29  | 30           | 11.47 | 32.35       | 67            | 77            | 85            | 96            | 93            | 3.1%  |
| 6  | 27  | 30           | 13.57 | 41.31       | 82            | 95            | 104           | 116           | 110           | 5.1%  |
| 7  | 16  | 30           | 7.018 | 22.19       | 48            | 50            | 54            | 61            | 59            | 3.2%  |
| 8  | 16  | 27           | 18.56 | 48.78       | 105           | 122           | 134           | 151           | 137           | 9.2%  |
| 9  | 42  | 30           | 25.24 | 69.91       | 146           | 169           | 186           | 209           | 192           | 8.1%  |
| 10 | 38  | 30           | 56.02 | 113.96      | 283           | 334           | 372           | 423           | 396           | 6.1%  |
| 11 | 51  | 30           | 54.41 | 165.51      | 330           | 379           | 416           | 466           | 438           | 6.0%  |
| 12 | 47  | 30           | 24.50 | 121.36      | 195           | 217           | 234           | 257           | 249           | 3.1%  |
| 13 | 19  | 30           | 7.042 | 26.74       | 48            | 54            | 59            | 66            | 59            | 10.6%   |
| 14 | 3   | 27           | 4.334 | 6.91        | 20            | 24            | 27            | 31            | 26            | 16.0%   |
| 15 | 21  | 30           | 12.26 | 29.56       | 67            | 78            | 86            | 97            | 91            | 6.1%  |
| 16 | 50  | 30           | 16.75 | 64.17       | 115           | 130           | 141           | 157           | 148           | 5.7%  |
| 17 | 72  | 30           | 34.86 | 97.71       | 203           | 234           | 258           | 290           | 273           | 5.8%  |
| 18 | 129 | 30           | 54.11 | 159.76      | 323           | 372           | 409           | 459           | 415           | 9.5%  |
| 19 | 138 | 30           | 54.35 | 204.71      | 369           | 418           | 455           | 505           | 488           | 3.3%  |
| 20 | 36  | 30           | 50.00 | 175.22      | 326           | 371           | 406           | 451           | 417           | 7.5%  |
| 21 | 8   | 28           | 7.178 | 29.21       | 51            | 57            | 62            | 69            | 65            | 5.7%  |
| 22 | 28  | 30           | 9.815 | 39.28       | 69            | 78            | 85            | 94            | 90            | 4.2%  |
| 23 | 67  | 30           | 56.50 | 164.37      | 335           | 386           | 425           | 477           | 461           | 3.3%  |
| 24 | 121 | 30           | 44.39 | 170.34      | 304           | 344           | 375           | 415           | 396           | 4.5%  |
| 25 | 141 | 30           | 73.05 | 212.05      | 433           | 499           | 549           | 616           | 556           | 9.7%  |
| 26 | 151 | 30           | 53.79 | 238.32      | 401           | 449           | 486           | 535           | 499           | 6.7%  |
| 27 | 64  | 30           | 70.77 | 201.49      | 415           | 479           | 528           | 592           | 549           | 7.2%  |
| 28 | 11  | 28           | 14.78 | 41.14       | 86            | 99            | 109           | 123           | 114           | 7.3%  |
| 29 | 21  | 30           | 16.01 | 76.90       | 125           | 140           | 151           | 165           | 157           | 4.8%  |
| 30 | 105 | 30           | 41.02 | 189.87      | 314           | 351           | 379           | 416           | 404           | 2.8%  |
| 31 | 140 | 30           | 45.98 | 193.93      | 333           | 374           | 406           | 448           | 429           | 4.2%  |
| 32 | 126 | 30           | 38.71 | 196.82      | 314           | 349           | 374           | 411           | 399           | 2.9%  |
| 33 | 131 | 30           | 57.14 | 229.31      | 402           | 453           | 493           | 545           | 529           | 2.9%  |
| 34 | 68  | 30           | 42.59 | 131.60      | 260           | 299           | 328           | 367           | 358           | 2.4%  |
| 35 | 70  | 30           | 77.76 | 250.40      | 485           | 555           | 609           | 680           | 658           | 3.2%  |
| 36 | 71  | 30           | 84.46 | 302.69      | 558           | 634           | 692           | 769           | 737           | 4.1%  |
| 37 | 30  | 30           | 83.33 | 242.00      | 494           | 569           | 626           | 702           | 660           | 5.4%  |

## 5 区域极值在单站的发生概率及单站稀遇极值

工程设计要求的是稀遇暴雨在单站的再现期值——单站稀遇极值。设在 $j$ 站上工程设计要求的再现期为 $T_j$ 年，区域极值在 $j$ 站的发生概率为 $p_j$ ，则在 $j$ 站 $T_j$ 年一遇的单站稀遇极值就是区域极值 $T_v (= T_j \times p_j)$ 年一遇的再现期值

$$x_{T_i} = \hat{x}_{o_i} - \hat{B}_i \ln(-\ln(T_i/(T_i+1))) \tag{15}$$

式中  $\hat{x}_{o_i}$ 、 $\hat{B}_i$  是用极大似然法对区域极值的年极值样本(大小排序)

$$x_{1i} \geq x_{2i} \geq \dots \geq x_{N_i} \tag{16}$$

估计的分布参数。

设该区内  $j$  站的年极值样本为  $y_i, x_i = 1, 2, \dots, n_j$

当  $y_i < x_{N_i}$  时, 说明  $j$  站未出现区域极值; 当  $x_{li} \leq y_i < x_{(l-1)i}$  时, 区域极值在  $j$  站出现 概率正比于  $(N_i - l + 1)/(N_i + 1)$ , 后者为次序统计量分布的数学期望。令

$$M_i = \begin{cases} = N_i - l + 1 & \text{当 } x_{li} \leq y_i < x_{(l-1)i} \\ = 0 & \text{当 } y_i < x_{N_i} \end{cases} \tag{17}$$

则第  $j$  站出现区域极值的概率正比于

$$K_j = \sum_{i=1}^{n_j} M_i \tag{18}$$

在固定区域均匀分布且测站密度合适时, 区域极值随机地出现于各个测站, 且只要有限个测站就可以抽得极值统计的样本, 因此可以假定存在独立完备的测站组, 均匀分布时测站数为  $N_o$ 。这样,  $j$  站出现区域极值的概率为

$$p_j = K_j / \sum_i K_i \tag{19}$$

当  $N < N_o$  时, 估计的  $p_i$  偏大; 但因采样不够密使区域极值偏小, 单站稀遇极值估计偏差不致太大; 当  $N > N_o$  时, 估计的  $p_i$  偏小, 但区域极值不会因测站过密而增加。这时邻近测站间极值样本可能会有相关。如在  $\alpha$  水平下相关显著, 则进行相关订正

$$(p_j^*)^T = (\rho_{ij})(p_i) \tag{20}$$

这里  $(\rho_{ij})$  表相关矩阵, 上角标 T 表示转置。

对于测站分布不均的区域, 求  $p_i$  时应用测站  $j$  所代表的面积  $S_j$  加权

$$p_i = S_j K_j / \sum_j (S_j K_j) \tag{21}$$

给定再现期  $T_i = 1000$ , 表 5 给出上蔡等 4 站单站极值再现期(剔除稀遇极值后的估值)  $y_{T_i}$ 、单站稀遇极值  $x_{T_i}$  及用稀遇极值样本或极值样本进行单站出现概率相关订正后计算(即  $T_i$  不同)的  $x_{T_i}^*$ 、 $x_{T_i}^{**}$ 。 $x_{T_i}^{**}$  是区域合并后的  $x_{T_i}^{**}$ 。由于采样年限较短时单站稀遇极值样本太少, 所以推荐  $x_{T_i}^{**}$  作为设计极值。

表5 再现期  $T_i = 1000a$  时单站极值再现期  $y_{T_i}$ 、单站稀遇极值  $x_{T_i}^{**}$

| 站名 | $\hat{B}_i$ | $\hat{x}_{o_i}$ | $Y_{1000}$ | $x_{1000}$ | $x_{1000}^*$ | $x_{1000}^{**}$ | $x_{1000}^{***}$ |
|----|-------------|-----------------|------------|------------|--------------|-----------------|------------------|
| 上蔡 | 73.046      | 212.05          | 248        | 406        | 533          | 571             | 588              |
| 西平 | 73.046      | 212.05          | 222        | 400        | 400          | 576             | 590              |
| 郟郟 | 54.112      | 159.78          | 217        | 325        | 435          | 453             | 532              |
| 徽山 | 53.792      | 238.32          | 240        | 415        | 415          | 458             | 513              |

需要指出的是,上蔡的,相应的区域极值再现期仅14a,相应的区域极值再现期为135a,都远小于1000a。用数十年区域极值样本估算135年一遇的极值,比用数十年单站极值样本估算1000年(甚至上万年)一遇的极值可信得多。事实上,单站极值的成千上万年的长再现期,并不是工程设备使用年限达上万年,而是由再现风险率较小造成的。区域极值在单站的较小的发生概率,基本抵消了小再现风险率延长再现期的作用。从而克服了对上万年的长再现期值需考虑气候变化的困难。

## 6 小结

在对极值样本的抽样误差进行可靠估计的基础上,依据在一次抽样中小概率事件不可能发生的原理,认定样本中的极大点数据来自另外的统计母体。对包含极大点数据的稀遇极值统计区域极值,不仅使同一气候区相邻的包含与不包含极大点数据的测站的统计结果(单站极值再现期值,单站稀遇极值)相接近,显著提高了时、空稳定性;而且区域极值在单站的较小的发生概率,使上百年一遇的区域极值相当于上万年一遇的单站稀遇极值,增加了统计结果的可信度。在合理划定测站相应的区域情况下,由导出的区域极值单站发生概率 $p_i^*$ 计算单站稀遇极值,满足了安全设计的需要。

对区域的划分和单站发生概率 $p_i$ 的相关订正还需要作深入研究。例如为避免虚假相关,似可在 $p_i$ 上乘一个空间相关函数 $\rho(r)$ 。它是测站间距离的函数,它的相关半径取决于稀遇暴雨的范围。为减少区域划分和调整的人工干预,需研究制定一些定量标准。

## 参考文献

- [1] 王守鹤等.几种随机概率模型对年洪水时序系列的适应性.水利水电科学研究院科学研究论文集.1988,28,1—6.
- [2] 詹道江,邹进上.可能最大暴雨和洪水.水利电力出版社,1983.
- [3] 董双林. Gumbel分布参数估计方法的统计分析.水利学报,1989,(11),35—42.
- [4] 董双林. Weibull分布函数形式的选取对参数似然估计的影响.应用数学与计算数学学报.1988,2(1),42—49.
- [5] 董双林. Gumbel分布函数形式的选取对参数似然估计的影响.科技通报,1987,3(4):11—14.

# ON STABILITY, MULTI-PARENT PHENOMENON AND REGIONAL EXTREMA IN EXTREMAL STATISTICS OF METEOROLOGICAL ELEMENTS

Dong Shuanglin

(Beijing Institute Of Applied Meteorology, 100081)

## Abstract

Taking Gumbel distribution as an example, it is demonstrated in this paper that the influence of sampling errors on the estimation of distributional parameters can be minimized if using the method of maximum likelihood, and the instability of extremal statistics is mainly caused by the isolated

maximum in the series of extremal samples. The further research shows that the isolated maximum come from other statistical parent, being the multi-parent phenomenon in the extremal samples. The statistical stability can be maximized if the scarce extrema including these isolated maxima are counted statistically as the regional extrema. In order to meet the needs of engineering design, the probabilities of regional extrema occurred in every stations are calculated.

**Key words,** Meteorological element, Extremum, Stability.