水文量の非毎年確率の適用に関する検討

A rsearch on non-annual probability value of hydrological events

北海学園大学　○学生員　豊田 尚正（Naomasa Toyota）
北海学園大学　　学生員　鍋島 靖裕（Yasuhiro Nabesima）
北海学園大学　　正員　　許士　達広(Tatuhiro Kyoshi)

1.　まえがき

水文確率値の算定は一般的に毎年確率値が用いられている。しかしデータ数が少ないときなど、毎年確率の年間第１位の値のみで無く､全体の分布傾向を知るうえで第２位第３位の値も含めて考えた確率が望ましい場合がある。また対象とする規模の現象が必ずしも毎年発生しない場合など毎年確率値があてはめにくい時には、非毎年確率を検討する必要が生ずる。河川計画の分野においては､非毎年確率は近年毎年確率の適合度が低い場合の検討事項として取り上げられ始めているが､あまり利用されてこなかったため､詳細については十分に解明されていない部分がある。ここでは非毎年確率分布の課題である閾値と確率値の関係､及び分布の適合度について報告する。

２．非毎年確率分布

降雨量を x として非毎年データの非超過確率の分布として以下のようなものが用いられている。

１）指数分布

累積分布関数　$F(x) = 1 - \exp\left(-\dfrac{x-c}{a}\right)$　・・・１）

c＝x $_0$（閾値）　$a = \dfrac{1}{M}\sum_{i=1}^{M}(x_i - x_0)$

２）一般化パレート分布

累積分布関数　$F(x) = 1 - \left[1 - \dfrac{k}{a}(x-c)\right]^{\frac{1}{k}}$　・・・２）

c＝x $_0$（閾値）　$k = \dfrac{\lambda_1}{\lambda_2} - 2$　　$a = \lambda_1(1+k)$

$$\lambda_1 = \beta_{0=} = \frac{1}{M}\sum_{i=1}^{M}\left[x(i) - x_0\right] \qquad \lambda_2 = 2\beta_1 - \beta_0$$

$$\beta_1 = \frac{1}{M(M-1)}\sum_{i=1}^{M}(i-1)\left[x_i - x_0\right]$$

ただし $x_0 - x_i > 0$ である。

これを確率年で表現するために以下の方法を用いる。ある閾値ｘ $_0$ を越えた洪水ピークの発生間隔がポアソン分布に従うものとし､ある事象がおこる年平均生起回数をλとすれば、年間の生起回数Ｘがｍである確率は

$P(X = m) = e^{-\lambda_*}\lambda_* / m!$　　（ｍ＝０，１，２・・・）

ここで x ミリの降雨以上の降雨が発生する年平均生起回数を $\lambda_* = \lambda(1 - G(x))$　　とおく。

λ＝Ｍ／Ｎ　　Ｍ：Ｎ年間で閾値ｘ $_0$ を越えた降雨数

年超過確率は年平均λ $_*$ 回起こる事象が１回以上起こることであるから、年非超過確率はm=0 として

$F(x) = e^{-\lambda^*} = \exp\{-\lambda[1 - G(x)]\}$　・・・３）

となる。ｘミリを越える非毎年の年非超過確率は３）式にＧ（ｘ）として 1),２)式のＦ（ｘ）を用いることにより求めることができる。Ｇ（ｘ）が指数分布のときはグンベル分布､一般化パレート分布のときはＧＥＶ分布の形状となる。

３．閾値と確率値の関係

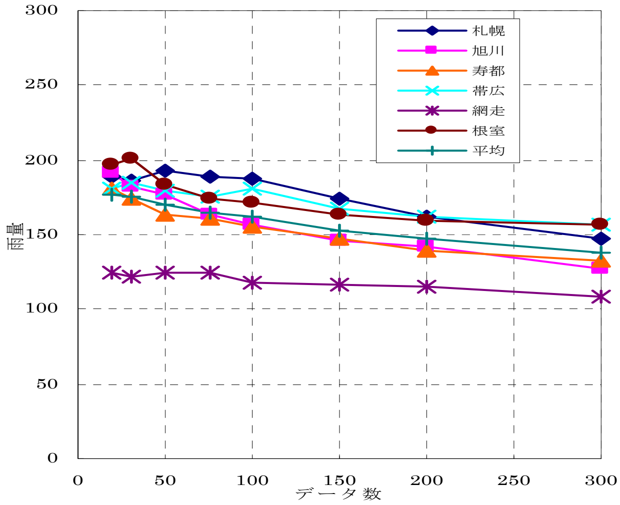
非毎年確率は一般に閾値が用いられている。非毎年確率分布の他の定数は､各データ値の分布に適合するようにいろいろな手法で決めることが出来る。しかし閾値はデータそのものの範囲を変えるので､定める手法に明確なものが無く、かつ確率値にも大きく影響する。

閾値の違いによる閾値以上のデータ数について､道内６箇所の雨量観測所の６０ヵ年の日雨量でみると表―１のようになる。

表―1　各地点のデータ数と閾値の関係（データ年数 60 年）

データ数	札幌	旭川	寿都	帯広	網走	根室
20	92.5	75	78	93.6	65.1	92.6
30	82.5	66	69.5	82	58.9	78.6
50	65.5	53.5	60	70	49	69.5
75	55	46.5	50.9	61	42	62.5
100	47.5	41.5	45.5	52.5	39	56.5
150	40	35.5	39.4	46.1	33.5	49.5
200	36.1	30.5	35.5	41	29.3	44.5
300	31	26.4	30.1	34	25	37

図－１は 60 年間の非毎年データについて、閾値の大小によるデータ数と 100 年確率値の変化を見たものである。計算法は指数分布にポアソン分布を適用したものである。どの観測所もデータ数が多くなるほど確率値が小さくなることが分かる。ここには示していないが 30 年間、100 年間と比較すると年数が小さいほど、確率値のデータ数による減少割合が大きい。



図－１　各都市の１００年確率値と閾値の関係

この傾向は、他のプロッティングポジションを用いた計算法などにも見られ、閾値の適切な設定が重要であることを示している。

4．閾値と適合度の考え方

どのような閾値を選ぶべきかについては、信頼の高い確率値を得る上で、分布の適合度が一つの指標となると考えられる。適合度の指標として、多く用いられているものとして、SLSC と相関係数（COR）がある。

1）SLSC

採用分布形の理論クオンタイルと標本順序統計量との誤差度合を測る指標で、次式によって定義される。

$$SLSC = \frac{\sqrt{\xi^2}}{|s_{0.99} - s_{0.01}|} \qquad \xi^2 = \frac{1}{N}\sum_{i=1}^{N}(s_i - r_i)^2 \qquad \cdots 4)$$

$s_{0.99}$、$s_{0.01}$:それぞれの非超過確率 0.99 と 0.01 に対する当該確率分布の標準偏差

s_i:順序統計量データを推定母数で変換した標準変量

r_i:プロッティングポジションに対応した理論クオンタイルを推定母数で変換した標準変量

2）X－COR

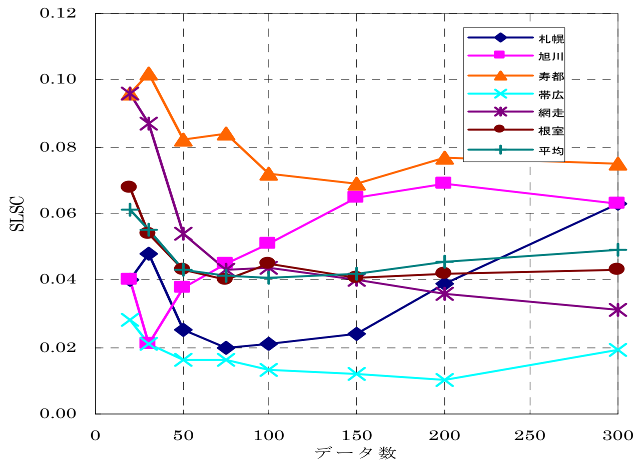
順序統計量 Xi（標本）に対応するプロッティングポジションにおける理論統計量 Yi について、以下の式で示される。

$$X-COR = \frac{\sum_{i=1}^{n}(Xi - \bar{X})(Y_i - \bar{Y})}{\left[\left\{\sum_{i=1}^{n}(X_i - \bar{X})^2\right\}\left\{\sum_{i=1}^{n}(Y_i - \bar{Y})^2\right\}\right]^{1/2}} \quad \cdots 5)$$

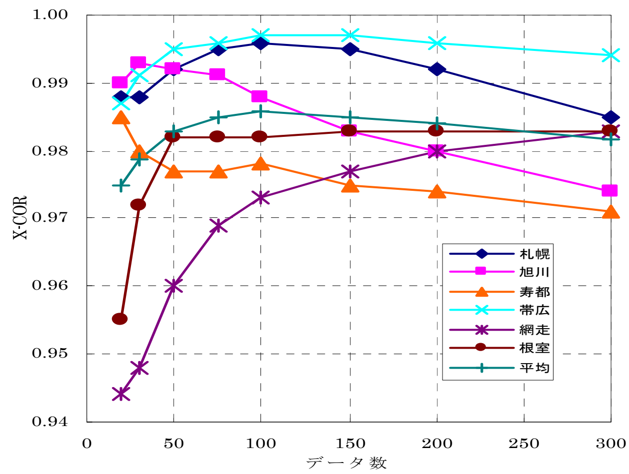
これらの値が閾値によりどのように変化するかを調べ、SLSC が小さく、X－COR が大きくなる閾値を採用する。

6. 計算例

図―２及び図―３は、指数分布にポアソン分布を用いた計算法の場合の SLSC および X-COR と、閾値によるデータ個数との関係をみたものである。他の年数の場合も含め、平均的には SLSC はデータ年数程度の個数、X-COR はデータ年数よりやや大きい個数で最適になっているが、観測所によるばらつきが大きい。



図－２　各都市のSLSCと閾値の関係



図－３　各都市のX-CORと閾値の関係

7．あとがき

非毎年確率は、近年毎年確率とあわせて比較検討されるようになっており、重要性が増しているが課題も多い。閾値の最適化は、確率分布の評価手法と密接な関係があり、さらに検討が必要である。

参考文献

1）星清：洪水ピークの評価方法について、開発土木研究所月報 No.539、pp34-47,1998

2）（財）国土開発技術研究センター：高水計画検討の手引き（案）、参考資料 pp20‐27