

流域平均極値雨量のAMSとPOTによる確率評価

ANNUAL MAXIMUM SERIES AND PEAKS OVER THRESHOLD FREQUENCY ANALYSIS OF EXTREME VALUE OF CATCHMENT AVERAGED RAINFALL IN JAPAN

田中茂信

Shigenobu TANAKA

正会員 工修 (独) 土木研究所水災害・リスクマネジメント国際センター
(〒305-8516 つくば市南原1-6)

This paper shows characteristics of Annual Maximum Series (AMS) and Peaks Over Threshold (POT) from time series of catchment averaged rainfall. It uses time series of 3-hour, 6-hour, 12-hour and 24-hour rainfall of A-class rivers in Japan. It is said that AMS and POT give same quantiles in the range of small exceedance probability theoretically. Further, it is also said that the shape parameter of Generalized Extreme Value (GEV) distribution for AMS and the shape parameter of Generalized Pareto (GP) distribution for POT can be used interchangeably. This paper compares the mean shape parameters of GEV and GP among regions in Japan. It also shows their systematic tendency by region. In addition, L-skewness and L-kurtosis relations of above data are also shown.

While the AMS procedure has been commonly used for hydrological frequency analysis because of easiness of data collection, it is emphasized that the POT analysis is indispensable.

Key Words : AMS, POT, GEV, GP, shape parameter, L-Moments, L-skewness, L-kurtosis

1. はじめに

わが国では治水計画に適用する水文頻度解析においては年最大値資料 (Annual Maximum Series, AMS) が一般的に用いられてきた。年最大日雨量や年最大洪水ピーク流量は極値資料であり、その分布には極値理論が適用できると考えられている。一方、極値を扱うもう一つの方法に閾値を超過する独立のピークを抽出する閾値超過資料 (Peaks Over Threshold, POT) を用いる方法がある。POTにおいて閾値を超える事象がポアソン分布にしたがって生起する場合、AMSの分布関数 $F(x)$ は

$$F(x) = \exp[-\lambda \{1 - G(x)\}] \quad (1)$$

で表される¹⁾。ここに、 $G(x)$ はの分布関数、 λ はPOTの年平均生起回数である。式(1)によりAMSとPOTが理論的に結ばれているので、本来はどちらを用いてもよいはずである。AMSは年間の最大値のみを抽出すればよいので資料整理が容易であるが、他の年の年最大値より大きな年間2位、3位などの極値データがある場合にこれらを有効に使うことができない。一方、POTには閾値の選定や独立なピークの抽出において課題がある²⁾。角屋はAMSとPOTの

関係を示し、地点雨量を対象にPOTを用いなくてもAMSから年に2度以上起こる事象の生起確率なども簡単に推測できると述べている³⁾。また、星も理論的にリターンペリオドが大きい場合、POTよりも資料収集が容易なAMSによればよいと述べている⁴⁾。田中・宝はわが国の一級河川の洪水資料を対象にAMSとPOTの比較を行い、両者が異なる確率水文量を与えることがあることを示した⁵⁾。本論文は一級水系の流域平均雨量の同一の時系列資料からAMSとPOTを抽出し、両者について極値理論に基づく分布モデルをあてはめて比較を行い、両者で確率水文量が異なる場合の対処方法を提案するものである。なお、わが国では極値理論に基づく分布モデル以外に対数正規分布他の分布も慣用されてきているが、ここでは極値資料を対象に極値分布に基づく分布モデルに焦点をあててAMSとPOTの両面から議論する。POTをGumbel確率紙上でAMSと比較する際には式(1)を用いて変換している。

2. 解析に用いた資料とその整理方法

解析に用いた資料は1999年時点における一級水系の全国103地点の3時間、6時間、12時間および24時間の流域平

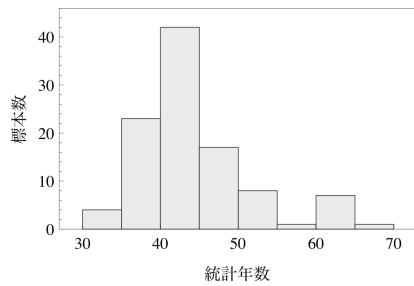


図-1 使用した資料の観測期間のヒストグラム

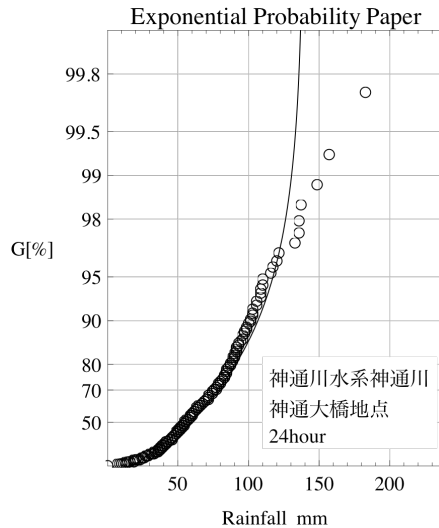


図-2 流域平均雨量のPOTの事例（神通川）

均雨量である。観測期間の分布は図-1のヒストグラムに示すとおりであり、最大67年、平均44年である。

式(1)の $G(x)$ の分布が一般パレート分布(GP)で、 $F(x)$ の分布が一般極値分布(GEV)で表される場合、それらの形状母数は互換的に使用できる²⁾ので、POTにGPをあてはめた場合の形状母数とAMSにGEVをあてはめた場合の形状母数を比較する。なお、POTは図-2に示すように大きい事象のデータと小さい事象のデータが必ずしも同じ傾向にないことが多い。この事例は神通川水系神通川の神通大橋地点上流における24時間の流域平均雨量のPOTを指数確率紙にプロットしたものである。POTにおける閾値（または上位個）選定については従来からいろいろな研究が行われている⁹⁾がまだ決定的なものはない。Tanakaら⁷⁾は種々の指標を比較した上で、POTから上位個を取り出した確率水文量の n に対する安定性が良好な判断規準となることを示している。この事例の場合、 $n=67$ を境に指数分布(Exp)とGPの確率水文量が発散・乖離し始める。このときの上位 n 番目の値は77mmであり、図-2においてこの付近を境にデータの傾向が異なっていることがわかる。ここではAMSとPOTの特性を見ることを主眼としているので閾値選定の問題を認識しつつ、POTについてはAMSと同じ標本サイズとなる非毎年最大値を使用して整理する。図-2の資料を非毎年最大値で整理するとその最小値は91mmであり、異なる傾向のデータが含まれる問題は回

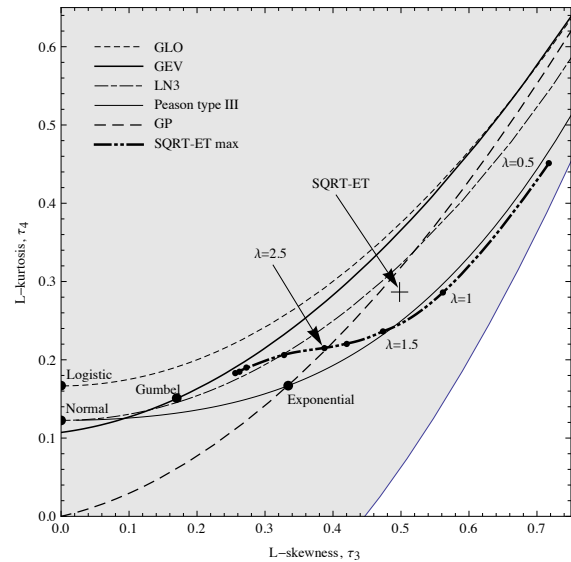


図-3 各種分布モデルのL-skewnessとL-kurtosisの関係

避でき、ほとんど直線状のプロットとなる。この場合、GPの形状母数は0.072であり、図-2の場合の0.574に比べかなり小さく、ほぼ0となりExpにより表される。

3. L積率比 L-skewnessとL-kurtosisの特性

(1) 分布モデルのL積率比

分布モデルを水文資料にあてはめるにはモデルの母数を推定する必要がある。一般に2母数の分布モデルでは平均と分散が、3母数の分布モデルでは、これに加えひずみ度がモデルの理論値と標本から得られる値が等しいと置いて母数を決定する。しかしながら、小標本の場合、高次の積率ほど標本誤差も大きくなり、偏り(bias)と変動の問題に直面する。積率法に代わって導入されたL積率法またはPWM(Probability Weighted Moments)法⁸⁾は順序統計量の線形和に基づく統計量であるため、上記の問題を緩和させるためにも有効である⁹⁾。宝らは、小標本の場合、3母数対数正規分布(LN3)、Gumbel分布、GEVにおいてL積率法が良好な母数推定法であることを示しており¹⁰⁾、わが国の水文極値資料の整理には適している方法である。

Hoskingらは従来用いられてきた分布モデルについてL積率による解法を整理し、L-skewnessとL-kurtosisの関係を表すグラフを示している¹¹⁾。これに江藤らが提案した平方根指数分布(SQRT-ET)および平方根指数型最大値分布¹⁾(SQRT-ET max)のL-skewnessとL-kurtosisの値を合わせて示したものが図-3である。なお、SQRT-ETおよびSQRT-ET maxの分布関数はそれぞれ $x \geq 0$ において次のとおりである。

$$G(x) = 1 - \left(1 + \sqrt{\beta x}\right) \exp\left(-\sqrt{\beta x}\right) \quad (2)$$

$$F(x) = \exp\left[-\lambda \left(1 + \sqrt{\beta x}\right) \exp\left(-\sqrt{\beta x}\right)\right] \quad (3)$$

ここに λ は事象の年平均生起回数であり、 β は尺度母数である。SQRT-ETはL-skewnessが0.4979とExpの1/3に比べ大き

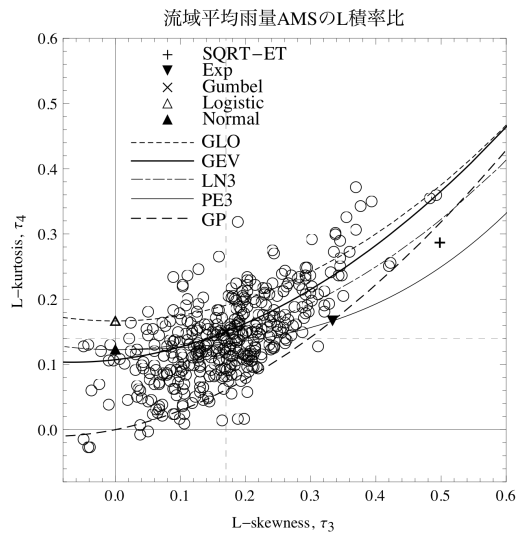


図-4 AMSのL-skewnessとL-kurtosisの関係

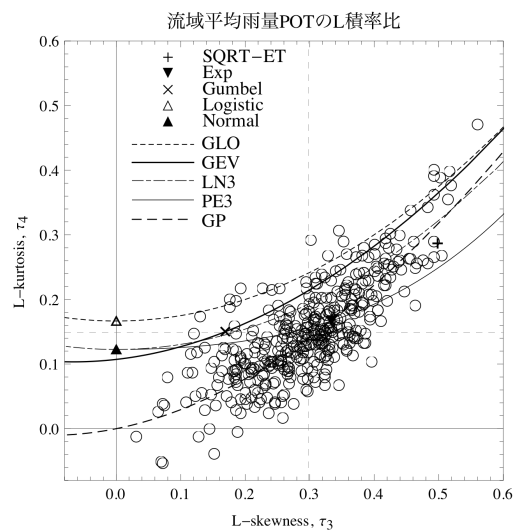


図-5 POTのL-skewnessとL-kurtosisの関係

く、極大方向に裾の長い分布であるのが特徴である。標本が有する歪みがこの理論値と同程度に大きい場合は適合度がよいが、歪みが小さいものや歪みが負の標本に対しては適合性はよくないと考えられる。ExpやGumbel分布等他の2母数の分布がこの図では点で表されるのに対し、SQR-ET maxは2母数の分布であるが、年間の事象の発生率 λ によってL-skewnessとL-kurtosisの値が変化する。ここでは $\lambda=0.5\sim 80$ について求めたL-skewnessとL-kurtosisの関係を示している。この図からわかるようにSQR-ET maxのL-skewnessとL-kurtosisの関係は、 $\lambda=0.5\sim 1.5$ の場合、Pearson type IIIの値とほぼ同じであり、2.5でGPの値とほぼ同じになり、4以上になるとLN3に近くなり、さらに λ の値が大きくなるとGEVに近づく。江藤らは λ が大きくなるとGumbel分布に漸近するかもしれないと述べているが、 $\lambda=300$ でもL-skewnessの値が0.242であり、GEVには近づいているもののGEVの特殊な場合であるGumbel分布に漸近するまでには至らない。本来、江藤らは λ が小さい場合は想定していないと思われるが、図-3にみられるように、 λ が大きい場合

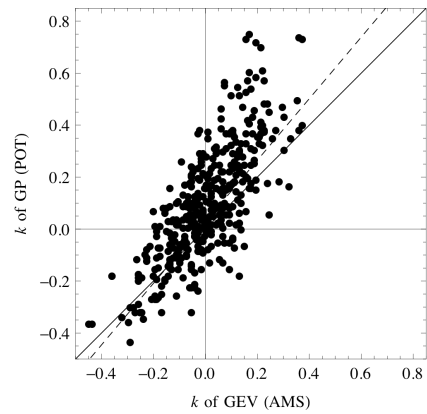


図-6 GEV (AMS) とGP (POT) の形状母数 k の比較

に比べ λ の値の少しの変化に対し敏感に対応して変化する。このため、 λ が0.5から2.5に変化する間にL-skewnessは0.718から0.388まで大きく変化する。いずれにしても、SQR-ET maxはGumbel分布よりL-skewnessが大きく、GEVにおける形状母数が正の領域にあり、上限をもつような標本にはうまく適合させることができない。

(2) AMSとPOTのL積率比

図-4および図-5は3時間、6時間、12時間、24時間の流域平均雨量すべてのデータに対し、それぞれAMSおよびPOTについてL-skewnessとL-kurtosisの関係を示したものであり、点線でL-skewnessとL-kurtosisの重心を示した。POTのプロットの多くはGPのL-skewnessとL-kurtosisの関係に近く、一部一般ロジスティック分布(GLO)の曲線に近いものも見られる。また、GPの曲線より下方に位置するプロットも少なからず見られる。これに対し、AMSでは全体的にGEVの曲線に沿ってプロットされているが、ばらつきの範囲が広く、GPからGLOでカバーされる範囲より広い範囲に分布している。一点鎖線で表されているLN3はGEVおよびGPを横切るような形の曲線であるため、多くの資料に対してあてはめることが可能となっている。なお、石原・高瀬はひずみ係数を比較することにより、Gumbel分布がLN3の特別な場合であると見て差し支えないとしている¹²⁾が、図-3のL積率比の関係からみるとGumbel分布とLN3は少し乖離している。データのプロットの重心はPOTについては(0.298, 0.149)であり、2母数の分布ではExp(1/3, 1/6)に最も近い。AMSの重心は(0.170, 0.140)であり、2母数の分布ではGumbel分布(0.1699, 0.1504)に最も近い。したがって、上述の関係より、概ねPOTはExpやGPであてはめられ、AMSはGumbel分布やGEVであてはめられるものが多いことがわかる。注目すべき点として、POTについてはロジスティック分布や正規分布をあてはめることが適当な資料はないが、AMSについては少なからず存在するとともにL-skewnessが負のものもみられる。

4. 形状母数の特性

(1) GEV (AMS) とGP (POT) の形状母数の比較

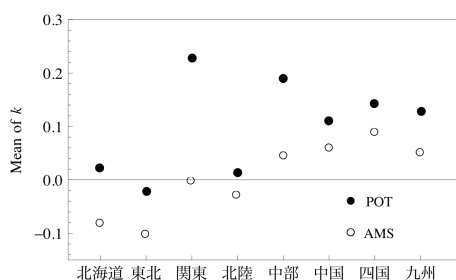


図-7 GEV (AMS) とGP (POT) の形状母数の地域間比較

図-6は非毎年最大値で整理したPOTにGPをあてはめた場合の形状母数 k とAMSにGEVをあてはめた場合の形状母数 k を比較したものである。GEV(AMS)の k の平均値0.005に比べGP(POT)の k の平均値0.097は少し大きめである。図中の破線は回帰直線であり、傾きが1.2でGEV(AMS)の k がゼロから正にかけてGP(POT)の k が前者に対して大きくなる傾向が顕著である。

(2) GEV (AMS) とGP (POT) の形状母数の地域特性

形状母数の地域的な特性をみるため、各地方におけるGEV(AMS)とGP(POT)の形状母数 k の平均値を比較したものが図-7である。この図から北から南に向かって形状母数が大きくなる傾向にあり、どの地域においてもPOTの形状母数の方がAMSの形状母数より大きい。この原因はいくつかの河川で中上位のデータが負の歪みを呈し上限値を有する分布のようなプロットとなっていてPOTの形状母数が正の大きな値になるのに対して、AMSはPOTに存在しない下位データの影響で形状母数を負の方向に偏らせるため形状母数がゼロに近い数値となることによる。図-8にこの事例として、鬼怒川の24時間雨量のプロットを示す。この図からAMSのプロットは逆S字状になっており、それを串刺しにするようにGumbel分布があてはめられ、GEVは約80mmより小さいデータの影響によりPOTにGPをあてはめたものよりGumbel分布に近くなっている(形状母数が正の小さい値となっている)。ここではPOTとして非毎年最大値を採用しているが、ある程度閾値を下げてPOTはこの図のGPの曲線上にプロットされ、GPによる確率降水量はあまり変化しない。このように形状母数が正の大きな値をとる場合は確率紙にプロットしてデータの特徴を把握することが重要である。POTの形状母数を降雨継続時間毎に整理すると、わずかではあるが降雨継続時間が長いほど形状母数が小さい傾向にあった。また、すべての降雨継続時間および地域において形状母数の平均値がAMSよりもPOTにおいて大きくなっている。この原因については次章に述べるようにAMSの下位データが少なからず影響していると考えられる。

なお、大河川以外の直轄河川では観測の歴史は浅く、さらに多くの中小河川ではより短いものや観測資料のないものもあり、これらの治水計画立案にあたっては地域頻度解析が有効であると考えられる。地域頻度解析は一地域内で用いる形状母数を設定するが、この際に直轄河

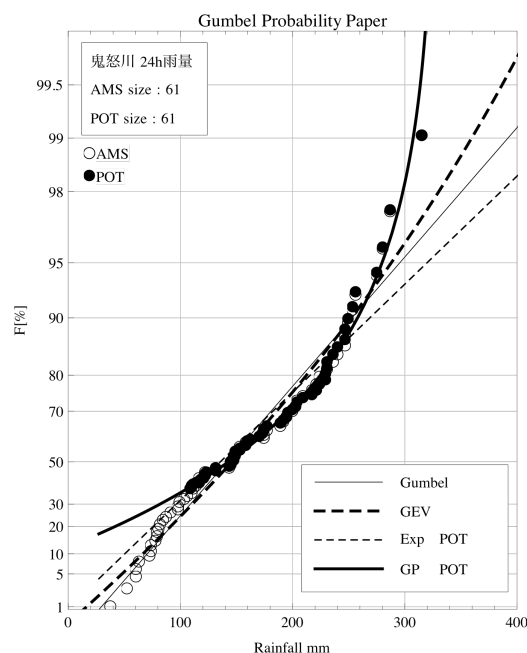


図-8 POTの形状母数が正の大きな値の事例(鬼怒川)

川流域の地点雨量の線形結合で表される流域平均雨量の形状母数情報は非常に有用な情報となる。

5. AMSとPOTにおいて形状母数が異なる原因および考えられる対応策

極値の分布は母集団から取り出す標本のサイズ n が大きくなると極値分布に近づくことが知られており、極値分布とはこの極限形式を意味する。極値分布は3つの形式に分類され、それらをまとめて一つの式で表したものがGEVである。極値が抽出されたものの母集団が従う分布が、Exp、ガンマ分布、正規分布、対数正規分布、ロジスティック分布のいずれであってもその極値はGumbel分布に従う¹³⁾。AMSは1年で区分されるよう抽出された標本の最大値からなる標本である。3章で述べたようにここで扱った資料では、概ねPOTはExpやGPであてはめられ、AMSはGumbel分布やGEVであてはめられるものが多く、式(1)の関係が適用できるものが多いことが期待できる。

しかしながら、田中・宝が指摘したように実際にはこの関係を満たす資料はそう多くない。この原因として、一つは年間のイベント数が少ないことが考えられる¹⁴⁾。極値理論はもともと1区分内のイベント数 n が十分大きい場合に極値分布に近づくのであって、これが少ない場合には極値分布に十分漸近しない可能性がある。

図-9は母集団がExpに従う条件の下で、年間 $r=360$ のイベントがあることを想定し、40年および100年分のイベントを発生させ、その資料からAMSおよびPOTを抽出したもののについてL-skewnessとL-kurtosisの関係をプロットしたものである。各条件についてそれぞれ100ケース実施した結果である。AMSはGumbel分布に十分漸近しておらず、ばらついているが、その範囲は概ねGEVに沿っている。POTも同様にExpに十分漸近していないが、そのばらつきの範囲は

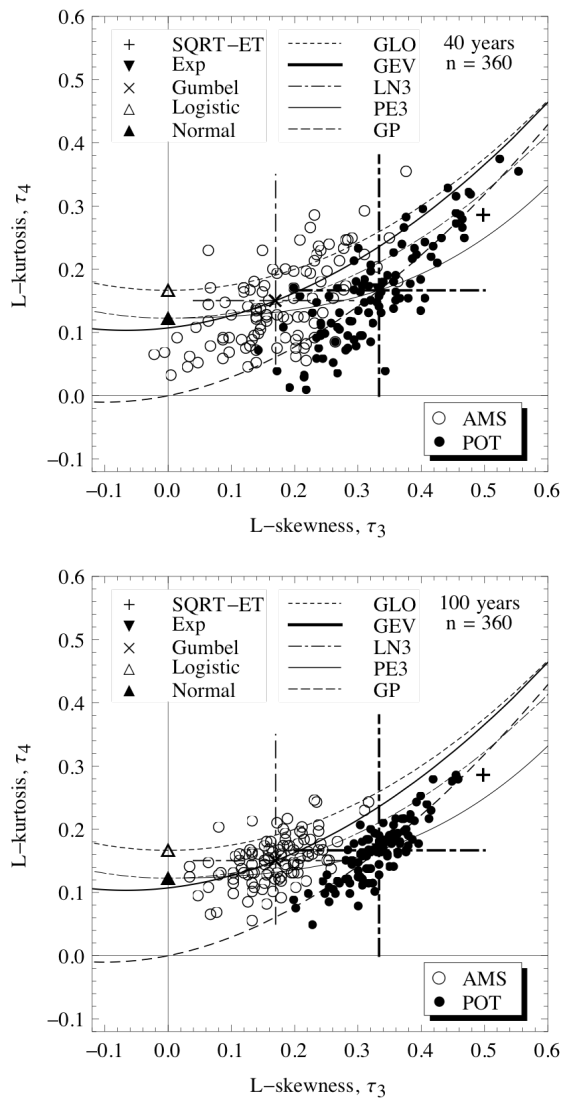


図-9 モンテカルロシミュレーションによるAMSとPOTのL-skewnessとL-kurtosisの比較(上40年, 下100年)
太い一点鎖線の十字はExp, 細い一点鎖線の十字はGumbel分布の関係を表している。

GPに沿うとともにAMSほどは広く分布していない。40年の場合、GLOからGPの範囲の外にも少なからず分布しているが、100年の場合、ばらつきが少なく、AMSはGumbel分布、POTはExpを中心に分布し、かつ、それぞれGEVおよびGPの曲線に沿ってデータが分布している。40年の場合のデータの分布は、図-4および図-5に示した分布状況と似ている。

図-10はこのシミュレーションのAMSに対してGEVをあてはめた形状母数と非毎年最大値にGPをあてはめた形状母数を求めて比較したものであり、年間のイベント発生回数10回で40年の場合と100年の場合を比較している。40年に比べ100年の方は形状母数のばらつきが小さくなっており、両形状母数の関係もかなりよく対応しているが、完全には一致しない。

以上のことから、L-skewnessとL-kurtosisの関係については、観測期間が短いことによってばらつくことは説明で

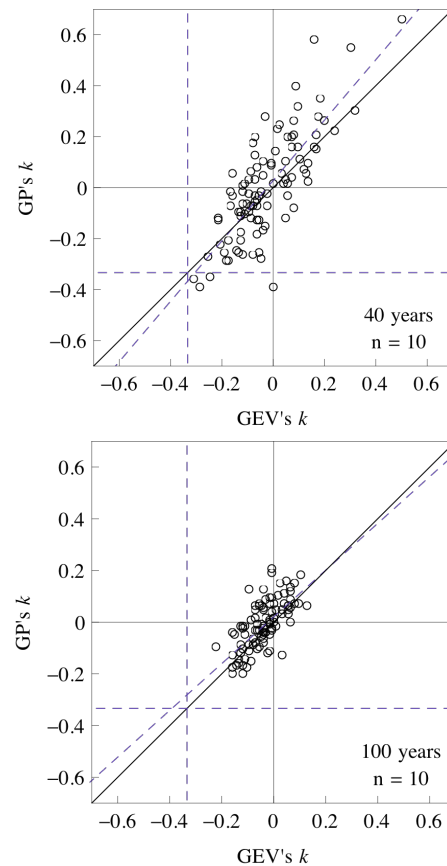


図-10 シミュレーションのGEV(AMS)とGP(POT)の形状母数の比較(上40年, 下100年)

きるが、形状母数が異なることについては、100年までの範囲では観測期間の長短の問題だけでは説明できず、記録の蓄積が100年程度まで進んでも解決されるものではない。

つぎに、実際の観測データでAMSとPOTにおいて形状母数がかかなり異なる場合の対応策について述べる。図-11は名取川水系広瀬川のAMSとPOT(非毎年最大値)をGumbel確率紙上で比較したものであり、AMSにGumbel分布とGEV, POTにExpとGPをあてはめた曲線もあわせて示している。この事例では歴代3位までの記録はAMSとPOTで同じであるが、それ以下には年間2位以下のデータがあるため、POTの方が少し右にシフトしている。この部分が影響してPOTにあてはめたExpとGPはAMSにあてはめたGumbel分布より大きい確率水文量を推定する。しかし、GEVは50mm以下のデータがそれより上位のデータと異なる傾向にあるため、形状母数が負に偏り、これが外挿範囲の確率水文量に影響し、大きい値を推定することになる。POTの最小値は54mmであり、AMSの下位17個のデータはこれ以下の値である。この場合、POTは閾値を下げてもデータはExpおよびGPの曲線上にほぼ乗っている。一般に洪水流量が小さい年は何年も続かないので、AMSについて複数年で最大値を抽出した標本を考慮することにより、1区分内でのイベント数のある程度確保するとともにAMSについて下位に同じようなデータが並ぶこと

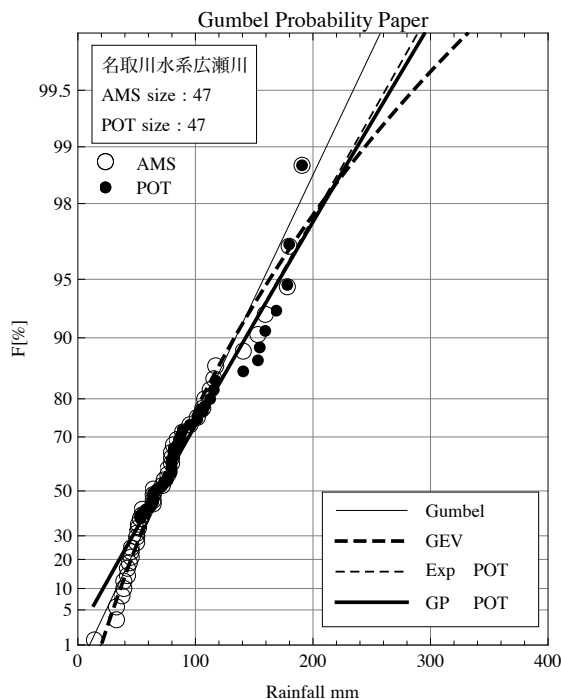


図-11 名取川水系広瀬川広瀬橋地点上流12時間雨量のAMSとPOTの比較

を回避できる¹⁵⁾。このようにして求めた確率水量はPOTによるものに近くなる。

それ以外の原因として、標本の均一性が保たれていない可能性も考えられる。これはPOTで見ると、小さい降雨がそれより上位のイベントとは異なる傾向を持つ場合などに相当し、この影響を除くにはPOT解析の閾値を適切に選定する必要があり、Tanaka and Takaraが示した方法⁷⁾により対応できる。

6. 結論

わが国の一級河川の3時間、6時間、12時間および24時間の流域平均雨量資料についてAMSとPOTの両面から特性を調べた結果、次のことが結論できる。

- 1) L-skewnessとL-kurtosisの関係については、AMSはGEV特にGumbel分布、POTはGP特にExpに近いものが多く、この傾向はシミュレーションでも再現される。シミュレーションは観測期間が長くなればこの傾向はよりはっきりすることを示しており、極値分布をあてはめることの妥当性を裏付けている。
- 2) AMSを用いたGEVの形状母数とPOTを用いたGPの形状母数の比較については、シミュレーションでははっきりした大小関係は見られないが、観測では後者の方が大きい傾向がある。
- 3) 同一の時系列資料から抽出したAMSとPOT（非毎年最大値）は上位データがほぼ等しいにもかかわらず上述2)の相違がみられる。AMSに問題があることはPOTと比較することにより初めて認識される。特に、Gumbel分布とGEVの確率水量の乖離が大きい場合、AMSの下位データに問題がある場合が多い。した

がって、AMS解析を基準とするのではなくPOT解析の結果もあわせて判断すべきである。

謝辞：本検討にあたって、国土交通省（旧建設省）から貴重な資料を提供いただいた。ここに記して感謝の意を表します。

参考文献

- 1) 江藤剛治, 室田 明, 米谷恒春, 木下武雄：大雨の頻度, 土木学会論文集, 第369号/II-5, pp.165-174, 1986.
- 2) Stedinger, J. R., Vogel, R. M. and Foufoula-Georgiou, E. : Frequency Analysis of Extreme Events, Chap. 18, Handbook of Hydrology, (Ed.) D. R. Maidment, McGraw-Hill, New York, pp.18.1-18.66, 1993.
- 3) 角屋 睦：雨量分布とその年最大値の分布, 京大防災研究所年報, 第4号, pp.122-131, 1961.
- 4) 星 清：洪水ピークの確率評価法, 開発土木研究所月報, No.539, 1998.
- 5) 田中茂信, 宝 馨：洪水頻度解析におけるAMSとPDSの比較, 水工学論文集, 土木学会, 第45巻, pp.205-210, 2001.
- 6) Rasmussen, P. F., Ashkar, F., Rosbjerg, D. & Bobee, B.: The POT method for flood estimation: a review. Extreme Values: Floods and Droughts, edited by K.W. Hipel, Stochastic and Statistical Methods in Hydrology and Environmental Engineering, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1, 15-26, 1994.
- 7) Tanaka, S. and Takara, K.: A Study on Threshold Selection in POT Analysis of Extreme Floods, IAHS Publ. no. 271, pp.299-304, 2002.
- 8) Hosking, J.R.M.: L-moments: Analysis and Estimation of Distributions using Linear Combinations of Order Statistics, J.Royal Statistical Soc., B, vol.52, no.2, pp.105-124, 1990.
- 9) 星 清：水文頻度解析, 開発土木研究所月報, No.540, 1998.
- 10) 宝 馨, 高埴琢馬：水文頻度解析モデルの母数推定法の比較評価, 水工学論文集, 土木学会, 第34巻, pp.7-12, 1990.
- 11) Hosking, J.R.M., and J.R. Wallis: Regional Frequency Analysis - An Approach Based on L-Moments-, Cambridge Univ. Press, p.207, 1997.
- 12) 石原藤次郎, 高瀬信忠：対数正規分布とその積率による解法, 土木学会論文集, 第47号, pp.18-23, 1957.
- 13) 志村隆彰：吸引領域と離散分布, 極値理論の工学への応用(5), 統計数理研究所共同研究リポート212, 統計数理研究所, pp.25-32, 2008.
- 14) 西岡昌秋, 宝 馨：季節性を持つ水文時系列に基づくPDS法とAMS法の比較, 京都大学防災研究所年報, 45-B2, pp.149-162, 2002.
- 15) 田中茂信：複数年で区切った最大値資料の特性, 極値理論の工学への応用, 統計数理研究所共同研究リポート169, pp.1-7, 2004.

(2009. 9. 30受付)