〔論　文〕　　415：503（再現期間；再現期待値；極値統計）

極端な豪雨の再現期間推定精度に関する検討

藤　部　文　昭*

要　旨

日降水量・2日間降水量の累年1位値の再現期間について，極値統計手法による推定精度をモンテカルロ・シミュレーションで検討した．統計手法として，よく使われる年極値解析に加えて閾値解析を取り上げ，両者の比較を行った．極値分布関数としては，それぞれ一般化極値分布と一般化パレート分布を採用し，これらに従う母集団から多数の擬似データを作って，データの確率変動の影響を調べた．得られた結果によると，再現期間の推定精度は形状パラメーター κ の求め方に大きく依存し，κ の真値（母集団の値）が地点によってあまり異ならなければ，κ として地点ごとに求めた値をそのまま使うよりも，多数地点の平均値を使う方が良い精度が得られる．一方，年極値解析と閾値解析の精度には大きな差がない．これらの結果を参考にし，気象官署・アメダス・区内観測のデータによる累年1位値の再現期間を求め，計算手法やデータソースによる結果の違いを概観した．

1．はじめに

記録的な大雨が起きると，それが何年に一度ぐらいの事象なのかを問われることがある．極値統計解析による再現期間がその答を与えるが，その計算値が数千年以上になる場合もあり，時としてその妥当性を巡って議論になる（例えば白木ほか 2001）．

極値統計の手法は多岐にわたるが，どの方法を使うにせよ，年極値の現れ方が確率過程に支配され，それに伴うランダムな変動（確率変動）がデータに含まれるため，再現期間の計算値に何がしかの推定誤差が生ずる．極値統計解析の精度については，データに基づく評価（Efron 1979；寶 2006；本城・篠田 2008）や，擬似データを使ったモンテカルロ・シミュレーション（Hosking and Wallis 1997；三浦・水野 2005）が行われているが，手法やデータの多様さに比べて精度についての情報は多くない．石原・仲江川（2008）は国内の日降水量の100年確率値の精度を検討しているが，極端な降水量の再現期間に関する評価結果は少ないのが実情である．

極値統計の手法として最もよく使われているのは，毎年の最大値（年極値）に適当な極値分布関数を当てはめる方法である．以下これを「年極値解析」と言う．これに対し近年は，全データの中からある基準値を超える値を選び出し，これに極値分布関数を適用する「閾値解析」も使われている（Brabson and Palutikof 2000；Coles 2001；Pandey *et al*. 2001, 2004；西岡・寶 2002；本城・篠田 2008；Della-Marta *et al*. 2009）．大雨が年に2回以上あった場合，年極値解析はその中で最大のものだけを扱うのに対し，閾値解析は基準値を超える観測値をすべて対象にするので，一見合理的である．しかし，年極値解析と閾値解析の得失については，一部の文献で議論されているものの（Brabson and Palutikof 2000；西岡・寶 2002など），まだよく検討されているとは言えない．そこで本研究では，降水の累年1位値の再現期間推定精度について，年極値解析と閾値解析を比べながら，モンテカルロ・シミュレーションを使って検討した．そして，その結果を国内の実データに適用してみた．

年極値解析では，一般化極値分布（generalized

* 気象研究所予報研究部．

—2009年1月8日受領—
—2010年5月10日受理—

extreme value distribution＝GEV；詳細は2.1節参照）が標準的な極値分布関数として広く使われている（Coles 2001；本城・篠田 2008）．従来はしばしばGumbel分布が使われたが，これは数学的にはGEVの形状パラメーター κ が0である場合に相当する．GEVの理論的根拠になるのは，同じ確率分布に従う互いに独立な多くのデータ（同一分布・独立・多数のデータ）の極値が理論上GEVに従うという性質である．日々の観測値を同一分布・独立・多数のデータと見なすことができれば，その年極値はGEVに従うはずである．一方，閾値解析では一般化パレート分布（generalized Pareto distribution；GPD）が使われている．GPDはGEVの漸近形であり，GEVによる年極値解析とGPDによる閾値解析とは対の関係にある（Coles 2001；本城・篠田 2008）．以後，これら2つの分布や，それに基づく解析を合わせて「GEV/GPD」「GEV/GPD解析」と表記する．本研究でもこれらの先例に倣い，GEV/GPDに基づいて解析を進めた．

なお，実データは「同一分布・独立・多数」の条件を必ずしも満たさず，従って年極値がGEVに従うという確実な保証はない．そこで，GEVのほかにいくつかの分布関数を用意し，その中からデータへの「適合度」が良いものを選ぶ方法が提案されている（Hosking and Wallis 1997；外山・水野 2002；三浦・水野 2005；小林 2006など）．この立場からすれば，GEV/GPDの適用の可否については議論もあろうが，本稿ではこの問題には立ち入らず，分布関数の検討は将来の課題の1つとして残しておきたい．ただ，アメダスの日降水量の年極値を対象にして数種類の確率分布関数の適合度を評価した外山・水野（2002）の解析によると，GEVに適合する地点が一番多いことが示されており，GEV/GPDを使うことには一応の妥当性があると考えられる．

2．解析方法

2.1　GEV/GPD解析の概要

GEVの累積分布関数は

$$
\begin{aligned}
F(x)&=\exp\left[-\left\{1-\frac{\kappa(x-\beta)}{\alpha}\right\}^{1/\kappa}\right] && (\kappa\neq 0)\\
F(x)&=\exp\left\{-\exp\left(-\frac{x-\beta}{\alpha}\right)\right\} && (\kappa=0)
\end{aligned}
\tag{1}
$$

である（Coles 2001）．x は同一分布・独立・多数のデータにおける一定期間（例えば年）ごとの最大値であり，F は値が x 以下である確率を表す．F と再現期間 T とは $T(x)=1/(1-F(x))$ の関係がある．α，β，κ はそれぞれ「尺度パラメーター」「位置パラメーター」「形状パラメーター」と呼ばれ，$F(x)$ のグラフの幅と x 軸上の位置，および形を決めるものである．$\kappa=0$ の場合はGumbel分布になる．(1)の逆関数は

$$
\begin{aligned}
x(F)&=\beta+\frac{\alpha}{\kappa}\{1-(-\log F)^{\kappa}\} && (\kappa\neq 0)\\
x(F)&=\beta-\alpha\log(-\log F) && (\kappa=0)
\end{aligned}
\tag{2}
$$

である．この x は再現確率 F，あるいは再現期間 $T=1/(1-F)$ に対する再現期待値を表す．

一方，年最大値が一般化極値分布に従うとき，十分大きい閾値 u を超える値はGPDに従う（Coles 2001；本城・篠田 2008）．GPDの累積分布関数は

$$
\begin{aligned}
F(x)&=1-\left\{1-\frac{\kappa(x-u)}{\alpha^*}\right\}^{1/\kappa} && (\kappa\neq 0)\\
F(x)&=1-\exp\left(-\frac{x-u}{\alpha^*}\right) && (\kappa=0)
\end{aligned}
\tag{3}
$$

であり，$\kappa=0$ の場合は指数分布である．式(3)は式(1)で x が十分大きいときの漸近形である．再現期間 T と F の関係は $T(x)=1/M(1-F(x))$ であり，M は u を超える観測値の年間回数である．M を与えたとき，α^* と u は

$$\alpha^*=\alpha M^{\kappa} \tag{4}$$

$$u=\beta-\frac{\alpha}{\kappa}(M^{\kappa}-1) \tag{5}$$

の関係になる（西岡・寶 2002）．式(3)の逆関数は

$$
\begin{aligned}
x(F)&=u+\frac{\alpha^*}{\kappa}\{1-(1-F)^{\kappa}\} && (\kappa\neq 0)\\
x(F)&=u-\alpha^*\log(1-F) && (\kappa=0)
\end{aligned}
\tag{6}
$$

である．

N 年間のデータがあれば，解析対象になるデータ数はGEVによる年極値解析では N 個，GPDによる閾値解析では NM 個である．後者の場合，M を増すほど使えるデータは増えるが，M が大きすぎると「閾値 u が十分大きい」という前提がくずれ，GPD

が成り立たなくなる．最適な閾値の決め方については統計的な検定が求められるが，現実の大雨発現頻度から見て $M=3\sim4$ が上限であろう．西岡・寶（2002）は $M=4$ をGPD適用の目安にしている．本稿では $M=1$ と $M=3$ の場合を中心として検討する．

2.2　分布パラメーターの推定法

分布パラメーターの推定には以前は積率法がよく使われたが，近年になってL-momentsによる計算法が，より優れた方法として広く使われるようになった（Hosking and Wallis 1997；外山・水野 2002；三浦・水野 2005）．これは，データからL-momentsを求めて各パラメーターを算出するものであり，GEVについては

$$\kappa=7.8590c+2.9554c^2 \tag{7}$$

$$\alpha=\frac{\lambda_2\kappa}{(1-2^{-\kappa})\Gamma(1+\kappa)} \tag{8}$$

$$\beta=\lambda_1-\frac{\alpha}{\kappa}\{1-\Gamma(1+\kappa)\} \tag{9}$$

である．ただし λ_1，λ_2，λ_3 は，それぞれ1～3次のL-momentの推定値であり，Γ はガンマ関数，また

$$c=\frac{2}{\lambda_3/\lambda_2+3}-\frac{\log 2}{\log 3} \tag{10}$$

である．$\kappa=0$（Gumbel分布）の場合は

$$\alpha=\frac{\lambda_2}{\log 2} \tag{11}$$

$$\beta=\lambda_1-\alpha\gamma \tag{12}$$

である（$\gamma=0.5772\cdots$　はオイラー定数）．GPDでは

$$\kappa=(\lambda_1-u)/\lambda_2-2 \tag{13}$$

$$\alpha^*=(\lambda_1-u)(1+\kappa) \tag{14}$$

となる．

パラメーターが決まると，式(1)や(3)を使って観測値の再現期間を計算できる．

2.3　実データ

以下の3種類の日降水量資料を使った．

①1901～2007年（107年間）の全国51官署の資料

②1979～2007年（29年間）の全国のアメダス資料

③1931～1975年（45年間）の本州中部26都府県（関東・中部・近畿地方と宮城・山形・福島県）の区内観測資料

①は，品質チェックを経た100年以上にわたるデータであり，生起頻度の低い，特に著しい事象を捉えるのに適している．一方②は，期間は短いが空間分解能が高く，局地性の強い豪雨を捕捉するのに適している．③は地域が限られるが，アメダスが展開される前の豪雨の発生状況を与える高分解能データである．いずれも，日本の豪雨の大部分が暖候期に起こることを考え，5～10月を対象にした．

①は那覇などを除いて欠測がほとんどないので，51地点すべての資料を使った．②については，欠測日が3日以上ある月を欠測月として，5～10月の欠測月が期間中（29年間）に3つ以下，かつ最初と最後の5年間に2つ以下である808地点を使った．③については，5～10月それぞれの欠測月が期間中（45年間）に8つ以下，かつ最初と最後の10年間に2つ以下である308地点を使った．なお，③は1926～1978年の資料をディジタル化したものであるが（藤部ほか 2008），1930年までと1976年以降は使える地点数が少ないので，対象期間から除いた．

日降水量の日界は，気象官署では主に24時であるが，1953～63年は09時，1939年以前は一部地点で22時である．区内観測は1952年までは10時，1953年以降は09時である．アメダスは24時とした．

日降水量の極値統計に関わる問題の1つは，日界をまたいで降った降水が2日に分かれて記録されることである．このことへの対応として，本研究では日降水量のほかに2日間降水量を扱った．閾値解析に当たっては，2日間降水量を対象日が重ならないようにサンプリングした．例えば，9月6～7日にその地点の年最大値が出たときは，9月5日以前と8日以降のデータから2位以下の値を採るようにした．

3．擬似データに基づく推定誤差評価

モンテカルロ・シミュレーションによる検討内容は，年極値解析と閾値解析の再現期間推定精度の比較，データの種類（統計年数など）と精度との関係，GEV/GPD分布パラメーターの求め方と精度との関係などである．一方，実際にやってみると，形状パラメーター κ の扱いが再現期間の推定精度に大きく影響した．そこで，以下 κ の扱いに重点を置いて記述を進める．

3.1 モンテカルロ・シミュレーションの手順

GEV/GPD に従う母集団から有限期間（N 年間）のデータが得られたとき，その累年１位値の再現期間の推定精度がどの程度であるかを考える．そのため，①乱数を使って上記分布に従う N 年分の擬似データを作る，②これに上記分布を当てはめて分布パラメーターの推定値を求め，再現期間を計算する，という手順を繰り返すモンテカルロ・シミュレーションを行った．その際，2.3節で紹介した各種データを念頭に置き，$N=107$（気象官署），$N=45$（区内観測），$N=29$（アメダス）の場合を扱った．N 年分の擬似データの１セットは，現実の場面では１地点の累年データ一式に相当する．以下分かりやすさを考え，１セットの擬似データを“　”をつけて“地点”と表記する．

モンテカルロ・シミュレーションの具体的な手順は以下である．

①母集団のGEVの分布パラメーター α，β，κ の設定

以下，これら母集団の値（すなわち真の値）を α_{TRUE}，β_{TRUE}，κ_{TRUE} と書く．実データでは，母集団は地点ごとに違うはずである．しかし，α と β は式(1)に $(x-\beta)/\alpha$ の形で入っているので，α_{TRUE} や β_{TRUE} がどんな値であっても，下記②で作られる擬似データは $X=(x-\beta_{\mathrm{TRUE}})/\alpha_{\mathrm{TRUE}}$ については同じ値になり，③で計算される再現期間も同じになる．そこで，以下のモンテカルロ・シミュレーションでは全“地点”に $\alpha_{\mathrm{TRUE}}=1$，$\beta_{\mathrm{TRUE}}=0$ を与えた．

一方，κ は極値分布関数の形を決めるパラメーターであり，3.2節以降で示すように極端な値の再現期間の推定に大きく影響する．実データの場合，κ は全地点で一定であるという保証がないので，モンテカルロ・シミュレーションでは κ_{TRUE} が“地点”ごとに異なる場合を扱えるよう，κ_{TRUE} の“地点”による変動を正規乱数で与えた．以下，その標準偏差を $\Delta\kappa_{\mathrm{TRUE}}$ と表記する．κ_{TRUE} の平均値は，実データから推定された値（第９図）に近い -0.05 とした．$\Delta\kappa_{\mathrm{TRUE}}=0$ ならば，母集団は全“地点”で同じになる．$\Delta\kappa_{\mathrm{TRUE}}\neq 0$ の場合は，κ_{TRUE} の変動がある分，母集団は“地点”によって異なる．

②各“地点”の擬似データセット作成

まず，“地点”ごとに正規乱数を発生させて κ_{TRUE} を決め（$\Delta\kappa_{\mathrm{TRUE}}=0$ ならばこの操作は不要），①の GEV に対応する GPD のパラメーター α^*_{TRUE} と u_{TRUE} を(4)(5)で求める．その際，$M=20$とする．

$M=20$というのは，１年当たり20個（N 年間で20N 個）の値が取れるように閾値 u_{TRUE} を与えることを意味する．前記のように，実データでは M が大きすぎると GPD の成り立つ前提がくずれるが，シミュレーションではそういう問題はない．むしろ，M が小さすぎると値を１つも割り当てられない年が現れ，年極値解析をする上で困ったことになる．

次に，$0<\varphi<1$ の一様乱数を発生させ，これを2.1節の F と見なし，それに対応する観測値 $x(\varphi)$ を式(6)で求める．その発生年も一様乱数で与える．これを NM 回行い，得られたものをその“地点”の擬似データとする．

このような擬似データを10000個の“地点”について作る．

③分布パラメーターの計算

上記の擬似データについて，2.2節の方法で分布パラメーターの推定値を求めた．パラメーター推定に当たっては，κの扱い方に関して以下の３通りの方法を試みた．[1]κ に適当な定数を与え，α，β は“地点”ごとに式(8)～(12)で計算する．なお，Gumbel 分布を使うことは，この方法で全“地点”に $\kappa=0$ を与えた場合に相当する．[2]α，β，κ をすべて，“地点”ごとに式(8)～(12)で計算する．[3]いくつかの“地点”をグループ化し，上記[2]で計算した κ の値をグループ内で平均し，その平均値 $\bar{\kappa}$ をグループ内の“地点”に一律に与え，これに基づいて“地点”ごとの α と β を式(8)～(12)で計算しなおす．以下，[3]を「κ 平均法」と言う．また，α，β，κの推定値をそれぞれ $\hat{\alpha}$，$\hat{\beta}$，$\hat{\kappa}$ と表記する（$\hat{\kappa}$ は文脈によっては，[1]で与えた値や[3]の $\bar{\kappa}$ を含む意味で使う）．

κ 平均法でグループ化する“地点”の数は，$N=107$ なら $n=51$，$N=45$なら $n=308$，$N=29$なら $n=808$ とした（第４図を除く）．これは，気象官署・区内観測・アメダスのそれぞれについて，国内の全地点をひと組にして κ の平均値を使うことを想定している．実際には，κ の真値（母集団の値）が地域によって系統的に異なる可能性もあり，その場合には地域ごとに別々のグループを作るという，地域頻度解析（Hosking and Wallis 1997；外山・水野 2002；三浦・水野 2005）に類した方法も考えられよう．しかし，地域分けの方法については第５節で触れるように検討すべき余地が大きいことから，本研究ではこの問題には踏み込まないこととした．

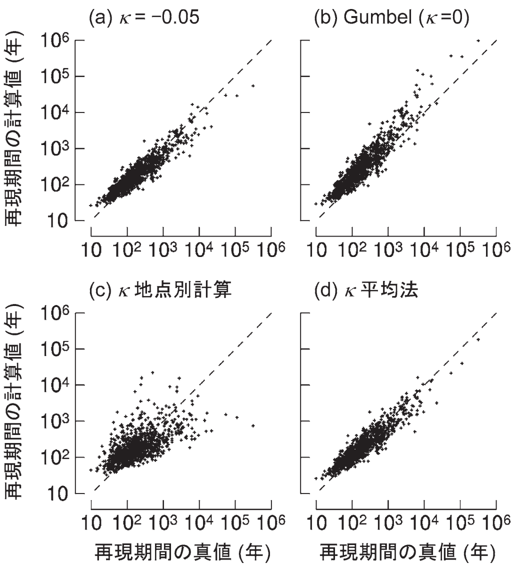
④再現期間の計算

③で求めたパラメーター $\hat{\alpha}$，$\hat{\beta}$，$\hat{\kappa}$ を使って，各

“地点”の累年1位値の再現期間や，ある再現期間に対応する再現期待値を計算した．そして，それらの値と真の値（②の φ に対応する値，すなわち $1/(1-\varphi)$．以下 T^* と表記）とを比較した．

3.2　再現期間の推定精度と形状パラメーター κ の求め方との関係

第1図は，擬似データの累年1位値の再現期間を，T^* を横軸に取り，年極値解析による推定値を縦軸に取ってプロットしたものである．$\Delta\kappa_{\mathrm{TRUE}}$ はひとまず0とし，気象官署を想定して $N=107$年とした．図の4つのグラフはそれぞれ，(a)κ を真の分布と同じ-0.05に固定した場合，(b)Gumbel 分布（$\kappa=0$）を使った場合，(c)κ を“地点”ごとに個別評価した場合，(d)κ 平均法（$n=51$）による場合の結果である．aとbは3.1節③の[1]に，cは[2]に，dは[3]に対応する．なお，dの“地点”総数は，51の倍数の中で10000に最も近い数，すなわち9996とした．その場合，$\hat{\kappa}$ は51“地点”ごとのグループ内では同じ値になるが，グループ間では値が異なる．



第1図　累年1位値の再現期間の真値と年極値解析による計算値のプロット．$\Delta\kappa_{\mathrm{TRUE}}=0$，$N=107$年とし，擬似データによる約10000回の計算結果のうち2000回分を示す．各計算方法の詳細については本文参照．

第1図によると，aでは計算値がほぼ真値のまわりに分布するが，それでも最大で数倍の推定誤差がある．再現期間の長い事例（1000年以上）では計算値が真値よりも小さくなり，短い事例では計算値が過大になる傾向がある．現実には κ の真値は分からないので，aは実用的ではないが，真の確率分布を与えてもなお，これぐらいの誤差ができることは注目すべきである．一方，bは再現期間が過大になる．これは，Gumbel 分布の $\kappa=0$ が真値 $\kappa_{\mathrm{TRUE}}=-0.05$ よりも大きいからで，一般に κ を真値よりも大きく与えると再現期間は過大になり，κ を小さく与えると過小になる．また，cは計算値のばらつきが大きい．これは，κ が未知数として加わることにより，データに含まれる確率変動に影響されやすくなり，再現期間推定の安定性が下がることを表している．これに対し，dはaと同様，計算値がほぼ真値のまわりに分布し，ばらつきはaと同程度である．このように，κ の平均値を使うことによって，推定誤差を小さくすることができる．

第1表は上記a〜dにおける $\hat{\alpha}$，$\hat{\beta}$，$\hat{\kappa}$ の平均値と90％幅（上下5％分位）を示したものである（dの $\hat{\kappa}$ の90％幅は，51“地点”ごとのグループ間の変動幅を表す）．$\hat{\alpha}$ と $\hat{\beta}$ の90％幅はどの方法を使っても同じぐらいである．$\hat{\kappa}$ については，当然ながらcの方がdよりも90％幅が大きい．また，$\hat{\kappa}$ の平均値は真値（-0.05）よりも少し大きめである．

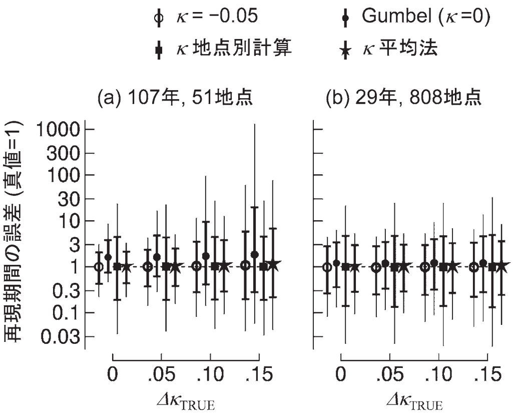
第1表の右側には，各パラメーターと累年1位値の再現期間（真値）の対数 $\log T^*$ との相関係数を示す．cでは $\log T^*$ と $\hat{\kappa}$ との相関が高く，他の3つで

第1表　GEV の分布パラメーター $\hat{\alpha}$，$\hat{\beta}$，$\hat{\kappa}$ の平均値と90％幅（上下5％分位），および各パラメーターと累年1位値の再現期間（真値）の対数 $\log T^*$ との相関係数．$N=107$年，$n=51$地点のもの．

	$\hat{\alpha}(\alpha_{\mathrm{TRUE}}=1)$		$\hat{\beta}(\beta_{\mathrm{TRUE}}=0)$		$\hat{\kappa}(\kappa_{\mathrm{TRUE}}=-0.05)$		$\log(T^*)$との相関係数		
	平均	90％幅	平均	90％幅	平均	90％幅	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\kappa}$
$\kappa=-0.05$	1.000	0.855〜1.164	0.000	−0.071〜0.064	−0.05	固定	0.55	0.15	−0.03
Gumbel	1.050	0.897〜1.221	0.009	−0.053〜0.073	0	固定	0.55	0.16	0.00
κ 個別計算	0.999	0.862〜1.145	0.001	−0.064〜0.070	−0.045	−0.168〜 0.083	0.04	−0.05	−0.63
κ 平均	1.005	0.858〜1.169	0.001	−0.060〜0.065	−0.045	−0.064〜−0.026	0.54	0.15	−0.09

は $\log T^*$ と $\hat{\alpha}$ および $\hat{\beta}$，特に $\hat{\alpha}$ との相関が高い．これは，データの確率変動による累年1位値の大小が分布パラメーターの計算値に影響すること，その影響は κ を個別に求めるときは主に $\hat{\kappa}$ に反映し，κ を固定すると主に $\hat{\alpha}$ に表れることを示している．第1図aやdで，再現期間の長い事例が過小評価されるのは，$\hat{\alpha}$ が過大に計算され，結果として再現期間が小さめに算定されるためであると解釈できる．

第2図aは，累年1位値の推定誤差を $\Delta\kappa_{\mathrm{TRUE}}$ の関数として示したものである．誤差は計算値と真値との比で定義し，図にはその対数の平均値と90%幅，および99%幅（上下0.5%分位）を示す．$\Delta\kappa_{\mathrm{TRUE}}$ が増すと，κ 平均法の誤差幅は大きくなる．これは“地点”ごとに κ の真値が異なることを無視して一律の κ 値を与えたことによるが，それでも，$\Delta\kappa_{\mathrm{TRUE}} \leqq 0.1$ の範囲では κ 平均法の方が κ を地点ごとに求めるよりも誤差が小さい．言い替えると，κ_{TRUE} に多少の地点間変動があっても，それを忠実に求めようとするより，κ を一律に与える方が精度のいい推定ができる．



第2図　累年1位値の再現期間の推定誤差（計算値÷真値）と，$\Delta\kappa_{\mathrm{TRUE}}$ との関係．記号は誤差の対数平均値を，上下に横棒のついた太縦線は90%幅を，細縦線は99%幅（上下0.5%分位）を表す．各 $\Delta\kappa_{\mathrm{TRUE}}$ 値についての4種類の計算結果を，重ならないよう横に並べて表示した．

第2図bは，aと同じことをアメダスを想定して $N=29$年，$n=808$地点について示したものである（κ 平均法の標本総数は，808の倍数の中で10000に最も近い9696とした）．定性的には第2図aと共通する特徴があるが，κ を地点ごとに計算する方法とそれ以外の方法との推定誤差の違いが第2図aほど大きくなく，$\Delta\kappa_{\mathrm{TRUE}}$ への依存性も第2図aに比べて弱い．これは，第2図aよりも年数が短い分，データに含まれる確率変動の影響がより大きくなり，κ の評価方法の違いによる推定誤差の違いは相対的に目立たなくなることを示している．なお，$\Delta\kappa_{\mathrm{TRUE}} \geqq 0.1$ の範囲では第2図aよりも誤差幅が小さいが，これは年数が短いため累年1位値自体が小さめだからである．第2表は $\Delta\kappa_{\mathrm{TRUE}}=0$ における $\hat{\alpha}$，$\hat{\beta}$，$\hat{\kappa}$ の平均値と誤差幅，および $\log T^*$ との相関係数を示す．各分布パラメーターのバイアスや誤差幅は第1表よりも大きく，$\log T^*$ との相関（κ を個別に求めるときは $\hat{\kappa}$，それ以外は $\hat{\alpha}$）も大きい．これらは，年数が少なくなるほどデータの確率変動の影響が増すことを反映している．

参考として，再現期間10～10000年に対応する再現期待値を第3図に示す．10年再現期待値は κ の扱いによらず，概して推定誤差が小さいが，再現期間を長くすると誤差が増し，手法間の違いが目立ってくる．ここでも κ の地点別評価は誤差が大きい．これに対し，κ 平均法は誤差もバイアスも小さく，真の分布（$\kappa=-0.05$）を使ったときとほとんど差がない．

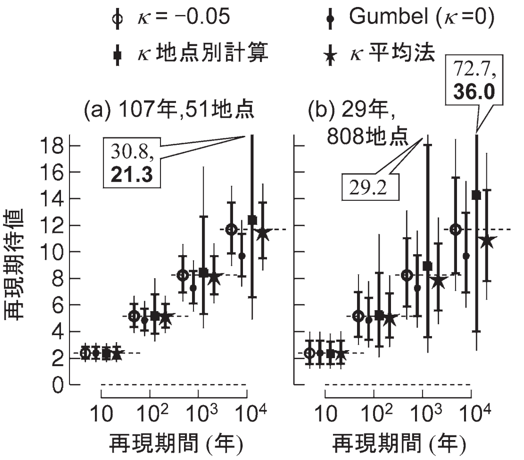
第4図は，κ 平均法による誤差幅と，その対象地点数 n との関係を示したものである．$n=1$ は κ を地点別に求める場合に他ならない．全体として，n が増すにつれて誤差は減少し，その度合いは $N=107$年の方が $N=29$年よりも，また $\Delta\kappa_{\mathrm{TRUE}}=0$ の方が $\Delta\kappa_{\mathrm{TRUE}}=0.1$

第2表　第1表と同じ，ただし $N=29$年，$n=808$地点のもの．

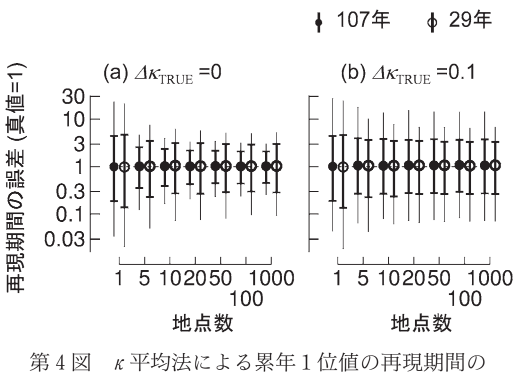
	$\hat{\alpha}$ ($\alpha_{\mathrm{TRUE}}=1$)		$\hat{\beta}$ ($\beta_{\mathrm{TRUE}}=0$)		$\hat{\kappa}$ ($\kappa_{\mathrm{TRUE}}=-0.05$)		$\log(T^*)$との相関係数		
	平均	90%幅	平均	90%幅	平均	90%幅	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\kappa}$
$\kappa=-0.05$	0.997	0.718～1.326	−0.001	−0.116～0.124	−0.05	固定	0.76	0.25	−0.02
Gumbel	1.047	0.756～1.387	0.009	−0.107～0.137	0	固定	0.77	0.29	0.00
κ 個別計算	0.994	0.729～1.289	0.003	−0.121～0.139	−0.031	−0.269～ 0.228	0.17	0.00	−0.70
κ 平均	1.017	0.735～1.355	0.003	−0.112～0.129	−0.031	−0.038～−0.026	0.76	0.26	−0.03

よりも著しい．$n=5$ や $n=10$ でも $n=1$ に比べて大幅に改善されており，地点数が少なくても，統合によって精度の向上を見込めることが示唆される．

現実には κ の真値（母集団の値）の空間変動が正規分布になるとは限らないが，仮に正規分布だとしても $\varDelta\kappa_{\mathrm{TRUE}}$ は未知であり，何らかの方法でデータから推定せざるを得ない．第5図は式(7)で計算した“地点”ごとの $\hat{\kappa}$ の標準偏差 $\delta\hat{\kappa}$ と，$\varDelta\kappa_{\mathrm{TRUE}}$ の関係を示したもので，$\varDelta\kappa_{\mathrm{TRUE}}$が増すほど $\delta\hat{\kappa}$ は増す．このことから，実データで $\delta\hat{\kappa}$ を計算すれば，これを使って $\varDelta\kappa_{\mathrm{TRUE}}$ を大まかに見積もることができよう．その適用については4.1節に記述する．



第3図　各再現期間に対応する再現期待値（対数平均値，90%幅，99%幅）．$\alpha_{\mathrm{TRUE}}=1$，$\beta_{\mathrm{TRUE}}=0$ として計算したもの．表示方法は第2図と同じ．誤差幅の上部がスケールアウトした部分は，数値を吹き出しで表示した（太字は90%幅，細字は99%幅）．



第4図　κ 平均法による累年1位値の再現期間の推定誤差（対数平均値，90%幅，99%幅）と，その対象地点数 n との関係．地点数の各値に対する2種類（$N=107$ 年と29年）の計算結果を，重ならないよう横に並べて表示した．

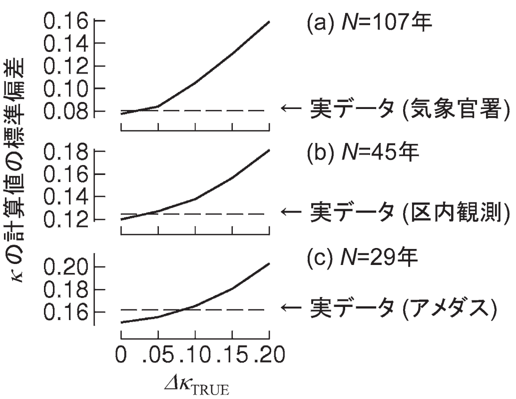
3.3　閾値解析と年極値解析の比較

GPDによる閾値解析においても，κ の与え方と推定誤差との関係は上記と同様である．以下は κ 平均法に基づき，閾値解析と年極値解析の結果を比較する．

第6図は閾値解析による累年1位値の再現期間の誤差幅を，M の関数として示したものである．各グラフの右端に，年極値解析による値を併せて示す．$\varDelta\kappa_{\mathrm{TRUE}}=0$ の場合は，M が増すにつれて誤差幅は小さくなる．しかし，$\varDelta\kappa_{\mathrm{TRUE}}=0.1$ においては M が増すと誤差はむしろ大きくなる．年極値解析による誤差幅は，閾値解析で $M=1$ としたときの誤差幅と同程度か，やや小さめである．

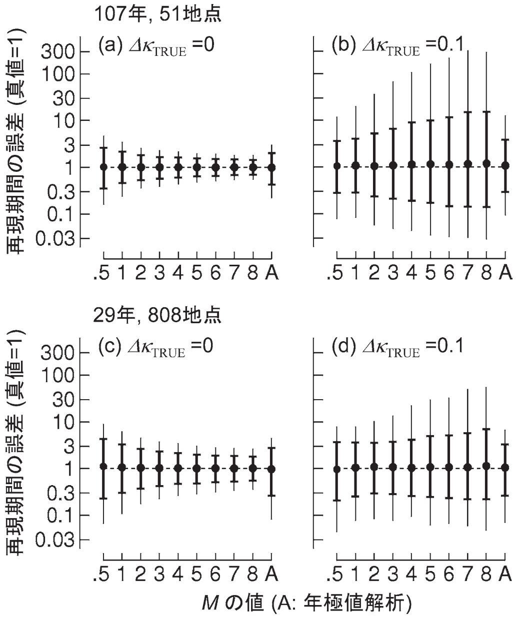
$\hat{\alpha}$ と $\log T^*$ の相関は，$\varDelta\kappa_{\mathrm{TRUE}}=0$ の場合，$M=1$ なら0.42，$M=3$ なら0.29であり，年極値解析による値（0.54；第1表）よりも小さい．これは，閾値解析が全期間中から上位事例を採用するため，年極値解析ほど累年1位値の影響を受けないからであり，大きな再現期間の過小評価傾向も年極値解析に比べて抑制される．この性質は M が増すほどよく現れるが，$\varDelta\kappa_{\mathrm{TRUE}}$ が増すにつれて，再現期間の大きい事例が過大評価されやすくなる（図は省略）．このように，閾値解析と年極値解析の優劣はデータの性格に依存する．

第7図は，年極値解析による累年1位値の推定誤差

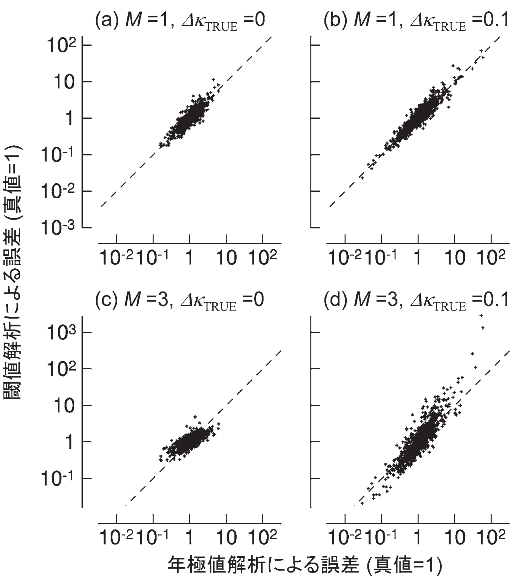


第5図　擬似データで得られる κ の標準偏差と $\varDelta\kappa_{\mathrm{TRUE}}$ との関係（実線），および対応する実データ（2日間降水量）から求めた $\hat{\kappa}$ の標準偏差（破線）．年極値解析によるもの．

と，閾値解析（$M=1$，$M=3$）による誤差とを比べたものである．左段は$\Delta\kappa_{\mathrm{TRUE}}=0$，右段は$\Delta\kappa_{\mathrm{TRUE}}=0.1$の場合である．誤差の間には高い相関があり，その対数同士の相関係数は$\Delta\kappa_{\mathrm{TRUE}}=0$の場合は0.88，$\Delta\kappa_{\mathrm{TRUE}}=0.1$の場合は0.95である．これは，年極値解析で再現期間が過大あるいは過小に評価されるときは，閾値解析を使ってもそうなる傾向が強いことを意味する．



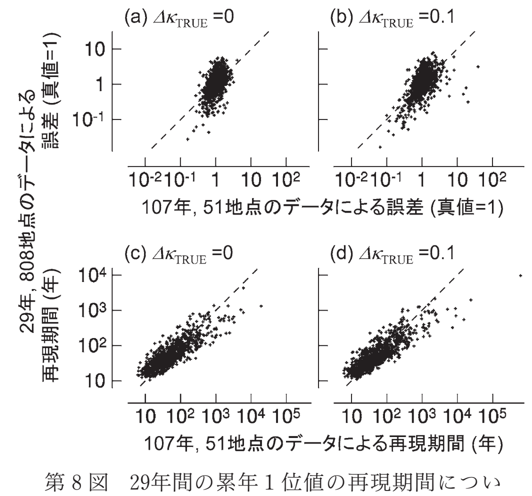
第6図　閾値解析による累年1位値の再現期間の推定誤差と，Mの関係．右端の“A”は年極値解析の結果を示す．



第7図　年極値解析による累年1位値の推定誤差と，閾値解析による誤差のプロット．約10000回の計算結果のうち2000回分を示す．

3.4　期間の長さの影響：気象官署とアメダスの比較

第8図は，29年間の808地点のデータが得られ，かつ，そのうちの51地点では107年間のデータが得られるという状況を与え，それら51地点の29年間の累年1位値を，$N=29$年，$n=808$地点のデータからκ平均法で求めたときと，$N=107$年，$n=51$地点のデータから求めたときについて比較したものである．これは，アメダスの累年1位値の再現期間を，107年間の気象官署データで計算したらどうなるかという問題に相当する．上段は再現期間の誤差同士の関係を示したものである．$\Delta\kappa_{\mathrm{TRUE}}=0$の場合は両者の対数の間に0.57，$\Delta\kappa_{\mathrm{TRUE}}=0.1$の場合は0.60の相関があるが，データの分布域は$y=x$の線よりも立っており，29年間のデータ（アメダス）による計算値は，107年間のデータ（気象官署）による値に比べて誤差が正負いずれにしても大きめになる．下段は再現期間そのものの比較である．再現期間の大きい事例の多くは$y=x$の線よりも下側にあり，統計期間が短いと，極端な事例の再現期間が小さめに算定される傾向（第1図）がよ



第8図　29年間の累年1位値の再現期間について，29年間808地点のデータから求めた値と，51地点107年間のデータから求めた値をプロットしたもの．

り目立つようになる．これは第2表で見たように，統計期間が短いほど分布パラメーター推定値（特に$\hat{\alpha}$）のデータへの依存性が強まることを反映している．

4．実データへの適用

この節では誌面の制約上，専ら2日間降水量についての結果を示すが，日降水量についても基本的に下記と同様の結果になる．

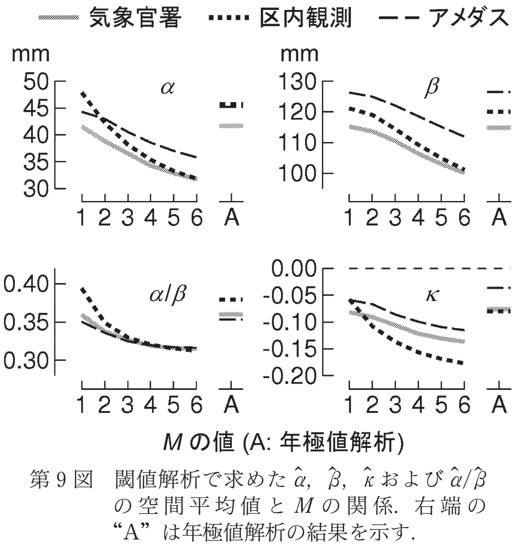
4.1　パラメーターの検討

4.1.1　$\Delta\kappa_{\mathrm{TRUE}}$に関する評価

第5図の横破線は，実データから求めた$\hat{\kappa}$の標準偏差$\delta\hat{\kappa}$を描き込んだものである．$\delta\hat{\kappa}$は気象官署と区内観測については前節のモンテカルロ・シミュレーションで$\Delta\kappa_{\mathrm{TRUE}}=0$としたときの値に近く，アメダスについても$\Delta\kappa_{\mathrm{TRUE}}<0.1$の範囲に入っている．これらは，実データにおける$\kappa$の地点間変動が小さいことを示唆し，全域に一律の$\kappa$値を与えることの妥当性をひとまず支持している．以下，このことに基づき，κ平均法によるGEV/GPD解析の実データへの適用結果を示す．なお，実データがGEVに従うこと自体が仮定であるが，その妥当性に関する検討は今後の課題としたい．

4.1.2　閾値解析におけるMの設定

第9図は気象官署・区内観測・アメダスそれぞれの2日間降水量について，$\hat{\alpha}$，$\hat{\beta}$，$\hat{\kappa}$および$\hat{\alpha}/\hat{\beta}$の空間平均値を示したものである．閾値解析の場合，Mを与えるとデータからuが決まり，これとGPDのパラメーター$\hat{\alpha}^*$および$\hat{\kappa}$を式(4)(5)に当てはめることにより，対応するGEVのパラメーター$\hat{\alpha}$と$\hat{\beta}$が求まる．$\hat{\alpha}$，$\hat{\beta}$，$\hat{\kappa}$はMによらないはずであるが，第9図では各パラメーターともMの増加につれて減少し，かつ，年極値解析で求めた値からずれていっている．これらの変化は統計的に有意であり，例えば区内観測とアメダスにおける$M=1$と$M=3$の$\hat{\kappa}$の差は地点ごとの値が独立ならば1％の水準で有意，気象官署についても$M=1$と$M=4$以上の$\hat{\kappa}$の差は有意である．これらのことは，データがGPDに厳密には従わないことを示唆する．その具体的な理由は分からないが，Mが増すにつれてより弱い降水事例が統計対象に含まれるようになり，GPDからのずれが大きくなる可能性が考えられよう．GEV/GPDと現実の極値分布との異同や，GEV/GPDを適用することの妥当性については，今後の課題として考えたい．以下の解析ではMを3までにとどめ，$M=1$と$M=3$の結果をGEV解析のものと比較しながら示す．



第9図　閾値解析で求めた$\hat{\alpha}$，$\hat{\beta}$，$\hat{\kappa}$および$\hat{\alpha}/\hat{\beta}$の空間平均値とMの関係．右端の“A”は年極値解析の結果を示す．

4.1.3　信頼幅の見積もり

再現期間の計算結果の信頼幅を見積もるためには，各地点の$\hat{\alpha}$，$\hat{\beta}$，$\hat{\kappa}$の計算値を真値と仮定して前節のようなモンテカルロ・シミュレーションを行い，計算結果の変動を調べる方法がある（三浦・水野 2005）．今回もこの方法を使って，累年1位値の再現期間の信頼幅を見積もった．手順は前節と同じであるが，α_{TRUE}，β_{TRUE}，κ_{TRUE}が$\hat{\alpha}$，$\hat{\beta}$，$\hat{\kappa}$に置き換わる．ただし，$\hat{\kappa}$自体が推定値であり，何がしかの推定誤差が見込まれる．前節のモンテカルロ・シミュレーションでは，各試行で計算されるκの値（前節で$\hat{\kappa}$と表記したものだが，本節で扱う実データからの推定値と区別するため，ここでは$\hat{\kappa}_{\mathrm{MS}}$と書く）と$\kappa_{\mathrm{TRUE}}$との差が，その推定誤差に他ならない．そこで，本節の$\hat{\kappa}$にもこれと同程度の推定誤差があることを想定し，κ_{TRUE}を

$$\hat{\kappa}'=\hat{\kappa}+(\hat{\kappa}_{\mathrm{MS}}-\kappa_{\mathrm{TRUE}}) \tag{15}$$

に置き換えた計算も行ってみた（$\hat{\kappa}_{\mathrm{MS}}$は51“地点”のグループごとに異なり，それに応じて$\hat{\kappa}'$も変わる）．こうして得られる信頼幅と，$\hat{\kappa}$を固定して計算した信頼幅との差は，気象官署（$N=107$）については数％である．例えば，東海豪雨における名古屋の2日間降水量（566.5mm）の再現期間を気象官署データで評価すると，$\hat{\kappa}$を固定して計算した90％幅は1134～8915

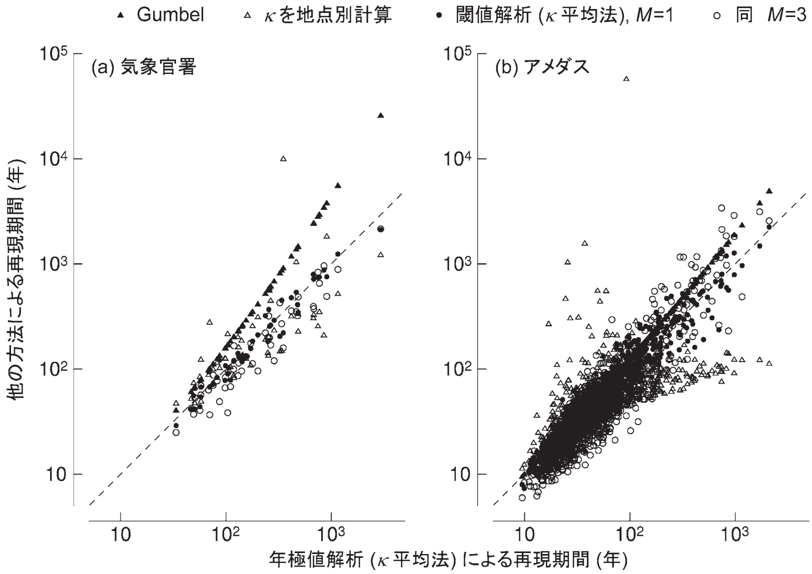
年，$\hat{\kappa}'$から求めたものは1165～9083年であり（第3表），ほとんど差がない．しかし，区内観測（N＝45）やアメダス（N＝29）では数割以上の差ができる．例えば，那須豪雨における那須の降水量（757mm）をアメダスデータで評価すると，再現期間の90%幅は$\hat{\kappa}$を固定すれば792～6845年であるが，$\hat{\kappa}'$を使うと438～27640年になる（第5表）．

より正しくは，$\hat{\alpha}$や$\hat{\beta}$の推定誤差も考慮すべきであろう．しかし，第1図や第8図のところで論じたように，$\hat{\alpha}$や$\hat{\beta}$はデータへの依存性があるため推定誤差の評価が難しい．これは今後の問題点として残しておきたい．

4.2 結果

第10図aは各気象官署の2日間降水量の累年1位値について，年極値解析（κ平均法）による再現期間と他の手法による再現期間を比べたものである．全体として手法同士の相関は高いが，Gumbel分布による再現期間は過大である．また，κを地点別に計算した結果はばらつきが大きく，前節の結果（第1，2図）と符合する．第10図bはアメダスについての結果であり，aと同様の特徴が見られる．

第3表は気象官署について再現期間の上位10件を示したものである．年極値解析によるものを主とし，他の方法による計算値を併記する．前節でも見てきたように，再現期間の90%幅は数倍，99%幅は1桁以上の範囲にわたる．これに比べ，年極値解析と閾値解析の差や，閾値解析のMによる差は小さい．なお，第3表の10件の中に寿都と高知が2回ずつ出てくるのは統計上不自然ではない．51地点から重複を許して10回の抽出をすると，15%の確率で2地点が2回以上選ばれる．第4，5表はそれぞれ区内観測とアメダスについ



第10図 年極値解析（κ平均法）による累年1位値の再現期間と，他の手法による再現期間の比較．

第3表 気象官署における2日間降水量の再現期間の上位10件．

	年/月/日	2日間降水量(mm)	再現期間（年）					
			年極値解析			閾値解析		アメダス・区内データ
				90%幅	99%幅	M＝1	M＝3	
名古屋	2000/9/12	566.5	2937	1165～9083	675～17121	2123	2159	981
寿都	1962/8/ 3	289.6	1154	496～3275	301～ 5881	1244	887	
前橋	1947/9/15	391.6	905	397～2506	245～ 4439	756	492	287
高知	1976/9/12	901.5	850	375～2339	232～ 4128	876	961	
水戸	1938/6/30	440.9	770	343～2097	213～ 3679	750	656	313
旭川	1981/8/ 5	293.0	747	334～2031	208～ 3556	745	834	431
高知	1998/9/25	874.0	690	311～1860	194～ 3243	713	792	735
浜松	1910/8/10	542.0	681	307～1831	192～ 3193	720	376	
松山	1943/7/23	380.3	675	305～1814	190～ 3161	796	395	
寿都	1961/7/25	264.5	616	280～1642	176～ 2848	668	507	

てのものである．気象官署に比べ，90％幅がさらに大きくなっている．

第11図は，区内観測・アメダスによる累年1位値のうち，107年間51地点の気象官署資料に含まれるものについて，それぞれのデータソースによる再現期間を比べたものである．概して，区内・アメダスによる再現期間と気象官署資料による値との間には高い相関（それぞれ対数同士で0.80，0.88）があるが，再現期間の大きい事例については，区内・アメダスによる値のほうが気象官署による値よりも小さい傾向がある（第3表も参照）．これはモンテカルロ・シミュレーションの結果（第8図）と符合する．なお，区内観測と気象官署資料とは日界の違いのため日降水量や2日間降水量が一致するとは限らないので，第11図や第3表に示した気象官署と区内観測の比較は，あくまでも大まかな関係を見るためのものである．

第4表　区内観測における2日間降水量の再現期間の上位10件．1つの事例で再現期間が10位以内に入る地点が2つ以上あるときは，再現期間が最大のものだけを採録し，その他の地点は注で示す．

	年/月/日	2日間降水量(mm)	再現期間（年）	
				90％幅
桐生（群馬）	1947/ 9/15[1]	370.0	775	245～3975
小国（山形）	1967/ 8/28[2]	565.0	555	184～2694
古川（宮城）	1948/ 9/16[3]	354.5	548	182～2655
久々野（岐阜）	1958/ 7/25	507.0	521	174～2500
水府（茨城）	1938/ 6/29[4]	423.5	470	159～2216
勝浦（千葉）	1971/ 9/ 7	558.0	317	113～1399
下里（和歌山）	1939/10/16	702.0	295	106～1287
中町（兵庫）	1938/ 8/ 2	360.0	293	105～1273
山形（山形）	1938/ 8/31	265.3	278	101～1197
長岡（新潟）	1961/ 8/ 5	292.0	263	96～1126

1）水上373.6mm／561年，沼田286.8mm／487年,前橋387.4mm／287年．いずれも群馬．
2）村上（新潟）338.0mm／355年．
3）築館（宮城）344.0mm／540年．
4）日立504.8mm／448年，柿岡467.2mm／345年，水戸464.1mm／313年，真壁452.6mm／265年．いずれも茨城．

第5表　アメダスにおける2日間降水量の再現期間の上位10件．1つの事例で再現期間が10位以内に入る地点が2つ以上あるときは，再現期間が最大のものだけを採録し，その他の地点は注で示す．

	年/月/日	2日間降水量(mm)	再現期間（年）	
				90％幅
那須（栃木）	1998/8/28	757	2107	438～27640
岩見沢（北海道）	1981/8/ 5[1]	398	1709	368～21190
久米島（沖縄）	2001/9/12	787	1162	268～12980
名古屋（愛知）	2000/9/12[2]	567	981	233～10458
厳原（長崎）	1985/6/24	591	863	209～ 8892
栃尾（新潟）	2004/7/13	427	822	201～ 8369
石狩沼田（北海道）	1988/8/26	422	742	184～ 7334
高知（高知）	1998/9/25[3]	874	735	183～ 7246
三石（北海道）	1981/7/ 6	322	699	175～ 6796
小谷（長野）	1995/7/12	389	670	169～ 6432

1）美唄389mm／978年，新篠津316mm／828年．いずれも北海道．
2）東海（愛知）589mm／748年．
3）南国（高知）874mm／727年．

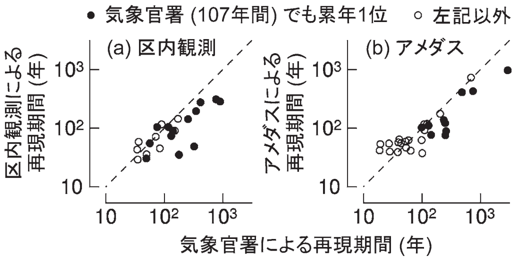
5．まとめと補足

日降水量・2日間降水量の累年1位値の再現期間について，年極値解析と閾値解析を比較しながら，その推定精度を検討した．極値分布関数としては，それぞれ一般化極値分布（GEV）と一般化パレート分布（GPD）を採用し，これらに従う母集団から多数の擬似データを作ってモンテカルロ・シミュレーションを行った結果，以下のことが示された．

①形状パラメーターκの真値（母集団の値）が地点によって大きく変動しなければ，地点ごとに算出されたκ値をそのまま使うよりも，その平均値を使う方（κ平均法）が精度が良い．

②年極値解析と閾値解析の精度に大きな差はない．

③分布パラメーターの推定値はデータの確率変動に影響される．この結果，再現期間の長い事例については，再現期間が真値よりも短めに算定される傾向がある．この傾向はデータの期間が短いほど顕著である．



第11図　区内観測・アメダスによる2日間降水量の累年1位値のうち，107年間51地点の気象官署資料に含まれているものについての，再現期間の比較．●は107年間の気象官署資料でも1位のもの，○はそうでないもの．

将来の発展としては，GEV/GPD 以外の分布への拡張が挙げられる．4.1.2節で示したように，実データが GEV/GPD から外れていることが示唆されており，これについての詳しい分析が待たれる．分布関数の選び方としては，冒頭で触れたように複数の分布関数の「適合度」に基づく方法があるが，今後はその過程でデータに含まれる確率変動がどう影響するかについての検討を進めていくことが望まれよう．上記①は，解析の自由度が大きすぎるとデータに含まれる確率変動に引きずられ，かえって結果の安定性が損なわれることを示しており，データへの適合性と統計的安定性の両面を考慮した解析の必要性を示唆している．

分布関数・分布パラメーターの地域依存性の評価も残されたテーマである．従来からよく使われてきた Gumbel 分布による解析や，それに基づいて多数地点の値を1つのグラフ上にプロットする SY 法（station-year method；鈴木・菊地原 1984）などは，全地点に同じ確率分布を仮定する方法であり，今回の κ 平均法もその延長である．一方，地域性を扱う方法としては，分布形の似た複数の地点をグループ化する「地域頻度解析」があるが（Hosking and Wallis 1997），グループ化の方法は検討の余地が大きい．1つの手法としてクラスター分析があるが，その場合どのグループにも属さずに残る地点があり（外山・水野 2002；三浦・水野 2005），それらについて個別に分布パラメーターを求めると，今回の解析で見たように推定精度が悪くなることが懸念される．石原（2008）は都道府県ごとに分布関数を求める方法を試みているが，都道府県という区分が最適かどうかについては議論があろうし，気象官署のように地点数の少ないデータにはそのままでは適用できない．多種多様な極値統計手法について，データの確率変動の影響を考慮しながらそれぞれの得失を評価し，極端事象の再現期間についてより精度の高い方法を見出していくことが望まれる．

極値統計解析に関わるもう1つの問題は，気候変動への対応である．極値統計は母集団の定常性を前提にしているが，現実には循環場の変動や長期的な気候変動により，母集団自身も変動する．例えば Brabson and Palutikof（2000）は，スコットランドの強風の統計に関し，10年スケール変動の影響を指摘した（そして，この問題に比べると年極値解析と閾値解析の違いは重要ではないことを示唆している）．また，数十～100年スケールで見ると，強い降水の増加傾向は世界の多くの地域に認められ（IPCC 2007），日本も例外ではない（Fujibe *et al*. 2006；Fujibe 2008）．

母集団の非定常性に関連すると思われる問題として，再現期間が数万年以上になるような特異事例の存在がある．例えば，今回の解析対象期間には含まれないが，彦根では1896年に日降水量596.9mm（9月7日），2日間降水量758.8mm（9月7～8日）という大雨が観測されており，これらの値に今回求めたパラメーターを当てはめると，再現期間はそれぞれ22万年と54万年になる（90％幅は6.8万～88万年と15万～246万年）．こうした数値は実体のあるものというより，母集団の定常性を前提にした解析の限界を暗示するものであろう．母集団の長期変動を扱う方法として，パラメーターを時代ごとにスライスして求めたり（寒川・鈴木 2008），その時間変化項を入れたりする手法（Coles 2001；Kharin and Zwiers 2005）も行われているが，短いスケールの変動まで考慮した極値統計解析のあり方については，今後検討を進めていく必要があろう．

謝　辞

研究費の一部として日本学術振興会科学研究費補助金基盤研究(B)「極端な気象現象の発生頻度とその長期変動に関する研究」(課題番号18340145) の支出を受けた．L-moments の計算には，Hosking によって公開されている Fortran77のサブルーチン(http：//lib.stat.cmu.edu/general/lmoments) を利用した．

参　考　文　献

Brabson, B. B. and J. P. Palutikof, 2000：Tests of the generalized Pareto distribution for predicting extreme wind speeds. J. Appl. Meteor., **39**, 1627-1640.

Coles, S., 2001：An Introduction to Statistical Modeling of Extreme Values. Springer, 208pp.

Della-Marta, P. M., H. Mathis, C. Frei, M. A. Liniger, J. Kleinn and C. Appenzeller, 2009：The return period of wind storms over Europe. Int. J. Climatol., **29**, 437-459.

Efron B., 1979：Bootstrap methods：Another look at the jackknife. Ann. Stat., **7**, 1-26.

Fujibe, F., 2008：Long-term changes in precipitation in Japan. J. Disaster Res., **3**, 51-60.

Fujibe, F., N. Yamazaki and K. Kobayashi, 2006：Long-term changes of heavy precipitation and dry weather in Japan（1901-2004）. J. Meteor. Soc. Japan, **84**, 1033-1046.

藤部文昭，松本　淳，小林健二，2008：区内観測による日降水量データのディジタル化と気候研究への利用における問題点．天気，**55**，283-287.

本城勇介，篠田昌弘，2008：統計的手法による作用モデルの構築．土木学会土木構造物荷重指針連合小委員会編「性能設計における土木構造物に対する作用の指針」，土木学会，A-I-15-38.

Hosking, J. R. M. and J. R. Wallis, 1997：Regional Frequency Analysis：An Approach based on L-moments. Cambridge Univ. Press, 224pp.

IPCC, 2007：Climate Change 2007：The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC［Solomon, S. *et al*.（eds.）］. Cambridge Univ. Press, 1056pp.

石原幸司，2008：地球温暖化に伴う確率降水量変化の都道府県別評価に向けて．日本気象学会講演予稿集，（94），251.

石原幸司，仲江川敏之，2008：全国51地点におけるノンパラメトリック手法を用いた確率降水量の算出．水文・水資源学会誌，**21**，459-463.

Kharin, V. V. and F. W. Zwiers, 2005：Estimating extremes in transient climate change simulations. J. Climate, **18**, 1156-1173.

小林健二，2006：確率雨量と再現期間の推定．測候時報，**73**，51-72.

三浦大輔，水野　量，2005：L-moments を用いた強風の再現期待値の推定．研究時報，**56**，97-125.

西岡昌秋，寶　　馨，2002：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，**45B2**，149-162.

Pandey, M. D., P. H. A. J. M. Van Gelder and J. K. Vrijling, 2001：The estimation of extreme quantiles of wind velocity using L-moments in the peaks-over-threshold approach. Struct. Saf., **23**, 179-192.

Pandey, M. D., P. H. A. J. M. Van Gelder and J. K. Vrijling, 2004：Dutch case studies of the estimation of extreme quantiles and associated uncertainty by bootstrap simulations. Environmetrics, **15**, 687-699.

白木正規，川本直樹，平井章仁，2001：東海豪雨：名古屋の日降水量の再現期間．日本気象学会予稿集，（80），131.

寒川典昭，鈴木將史，2008：日本列島20世紀の降水量時系列の経年的非定常性とその確率降水量の評価値に及ぼす影響．自然災害科学，**26**，355-365.

鈴木昭夫，菊地原英和，1984：異常豪雨を考慮した日降水量再現期間の推算法．天気，**31**，179-189.

寶　　馨，2006：大標本時代の水文頻度解析手法─リターンピリオドを超えるようなサイズの標本に対する極値データ解析─．京都大学防災研究所年報，**49B**，7-12.

外山奈央子，水野　量，2002：L-moments を用いた地域頻度解析による全国アメダス地点における確率雨量の推定．研究時報，**54**，55-100.

Assessment of Uncertainty in Estimating the Return Periods of Extreme Rainfalls

Fumiaki FUJIBE*

* *Meteorological Research Institute, Tsukuba 305-0052, Japan. E-mail： ffujibe@mri-jma.go.jp.*

(Received 8 January 2009；Accepted 10 May 2010)

Abstract

The accuracy of extreme value analysis in estimating the return periods of record highest rainfalls was examined with a focus on the influence of sample variability. A series of Monte Carlo simulations was applied to an annual maximum value (AMV) analysis based on the generalized extreme value distribution (GEV) and a peaks over threshold (POT) analysis based on the generalized Pareto distribution (GPD), by generating sample data corresponding to these distributions. It was found that the shape parameter κ had a key role controlling the accuracy of estimation, which was improved by using the average value of κ over stations, rather than calculating the κ value for each station, unless the true value of κ was not too variable among stations. On the other hand, no significant difference was found in the accuracy of AMV analysis and POT analysis. Based on these results, return periods of record highest rainfalls were calculated using three kinds of climatic data in Japan, with attention to differences according to methods and data sources.