極端な天気現象の長期変化の検出

山元龍三郎（京都大学名誉教授）

要旨

気象庁の極値順位データの統計解析により、極端な天気の長期変化を検出する手法を述べ、その解析結果の一部を例示する。

1. はしがき

筆者らが（Iwashima and Yamamoto,1993）、極端な天気の長期変化の検出の問題に取り組んだきっかけは、世界的な気象災害の増加であった。その原因探求のためには、人間社会の脆弱性の深刻化の他に、集中豪雨など極端な天気現象それ自体の激化の確認が必要だと認識したからである。

極値統計は、工学分野では1950年代から研究や応用が活発に行われ始めたが（例えば、Gumbel,1958）、日本の気象界では災害に関連して1960年代に始めて注目されるようになった（例えば、高橋浩一郎、1961）。それでも、極値の再現期間の議論が中心であり、極端現象の長期変化の研究は見当たらなかった。その理由は、極値が現象の発生確率密度分布における Outlier であり、その変化について信頼できる結果を生み出すことは、既存の統計手法では容易でなかったためだと思われる。

筆者らは、新しい統計手法の開発に精力を注がざるをえなかったが、その際に特に注意したのは、統計結果の信頼性である。これに対する筆者の関心を芽生えさせたのは、Sussex paradox と筆者が命名した英国での矛盾した観測データであった。イギリス海峡に面した Worthing という都市で、再現期間が1000年だと算定されていた大雨が、1年足らずの間に立て続けに2回も発生した(Potts,1982)。これは一見して明らかな paradox であるが、50 年程度の時系列データから算定した再現期間の誤差のなせる結果である。このような問題は、発現が稀な Outlier の統計では特に深刻であるので、筆者は区間推定にこだわることとした。

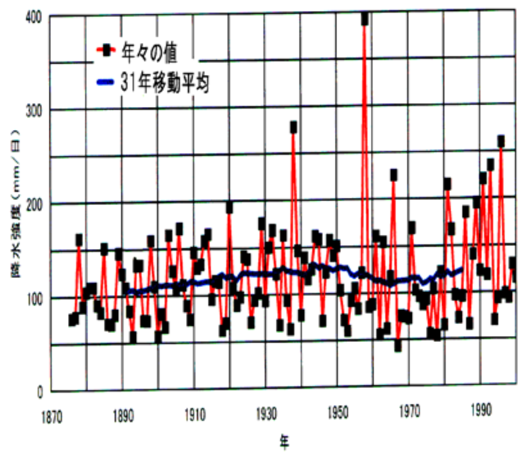

図-1　東京における日降水量の年最大値の推移

日本におけるルーチン気象観測の時系列データの期間は、高々130 年である。図-1 の例に見られるように、100 年間に不規則に僅か数回だけ発現する極端現象の長期トレンドを、1 ヶ所の時系列データから十分な信頼度で算定することは、極めて困難である。そのために、筆者らは、個々の観測点ではなく、複数観測点を含む観測網データ全体を解析対象とせざるをえないと判断した(Iwashima & Yamamoto, 1993)。

この論文では、気象庁の極値順位データセット(RANK)から、集中豪雨などの極端な天気現象の長期変化を検出する手法を述べ、若干の検出結果を示す。

２．　極値の確率分布

極値データの身近な例は、日降水量の年最大値や気温の年最高値である。これらの発現頻度は、Gumbel 分布と呼ばれる分布に従うことが知られている（Gumbel,1958；Ang & Tang,1988;Wilks,1995)。極値 x に対する累積分布関数 F(x) は

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

で与えられる。ここで、 G と B はそれぞれ location parameterと scale parameterであり、データ系列の平均値 aと標準偏差ｓから次式により算定できる(Wilks,1995)：

$$B = s\sqrt{6}/\pi \quad \cdots\cdots (2)$$

$$G = a - \gamma B \quad \cdots\cdots (3)$$

ここでπは円周率で、γは次の値を持つEulerの定数である：γ＝ 0.57721・・・・。

日降水量や時間降水量の年最大値などの気象観測データの極値が、十分な精度でこの分布に従うことは確認されている。

Gumbel 分布の確率密度は、Gauss 分布とは異なり、低い値の部分に比べて高い値の部分で非常に長い裾(Outliers)を持っている。このoutliersの年代による変化が、この論文の主題としている極値の長期変化である。

この論文の目的にとって、２つの parameters G と B の値の長期変化を調べることは適当ではない。これらの parametersの値は、確率密度の大きい部分のデータに強く支配されている一方、outliersの寄与が小さいことが、その理由である。

３． Monte Carlo実験

気象庁年報には、地上気象観測結果の極値順位データが RANKファイルの名で収録されている。それは、気象庁管轄下の約160ヵ所の気象台・測候所での気温や降水強度などに関する観測開始以来の極値（第１位から第５位までまたは第10位まで）である。

Outliersデータには、通常のデータ以上に、一般に誤りのデータの介在が懸念される。RANKデータセットではデータ毎にremarksが附記されていて、入念な品質管理が施されているので、誤りのデータは少ないものと期待される。

これらのRANKデータの統計処理の結果の信頼性の判定則を設定するために、筆者は、Monte Carlo実験を行った(Yamamoto,1996)。

乱数によるMonte Carlo実験の結果を、極端な天気の長期変化の問題に適用できる前

提として、RANK データの時間的・空間的 randomness の確認が必要である。

個々の観測点での時系列データの(時間的)randomness は、統計学の初歩の教科書(例えば Hoel,1966）に記述されている run(連)検定により、容易に確認できた。

空間的 randomness については、極値順位データの観測点相互間の発現時期の非合致率で確認することとした。114 ないし 120 ヶ所での 60 年間の極値（第 1～5 位）の発現年月日の合致率は、日降水量・時間降水量及び 10 分間降水量では 0.06～0.29%であり、最高気温と最低気温については 0.66～0.83%であった。このような極めて低い合致率は、極値順位データの空間的 randomness を容認するものであり、また、極端天気現象の空間スケールが一般に小さい事実と整合している。

ここでの Monte Carlo 実験では、Gumbel 型乱数を用いた（伏見、1989）。1 地点の 1 年の極値に 1 個の乱数を対応させる。100 個の乱数から成る系列を 1 地点の 100 年時系列データとし、そのような系列の 50 組を 50 地点から成る観測網データに対応させる。そのような 50 組の乱数系列

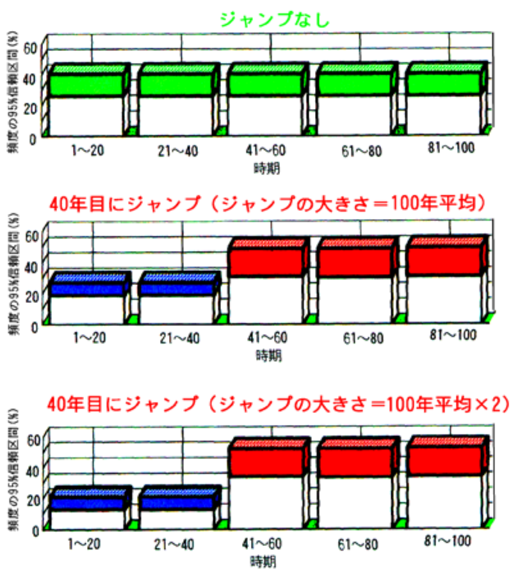

図-2 Monte Carlo 実験における 100 年最大値の発現時期の年代別頻度分布とその 95%信頼限界（42 ヶ所での第１～4 位の最大値）。上段のパネルはもとの時系列データに関するものであり、中段と下段のパネルは 40 年以降の乱数値が階段的に増加した場合の結果である。

の 1000 群を作り上げて、解析結果の信頼区間の算定に利用した。

100 年間の乱数時系列データで、100 年最大値の発現時期を20年毎に区切った5時期について求めた結果が、図-2 である。一例として、乱数の値が 40 年目に唐突に増加する場合を設定した。この場合には、40 年以降の年代で、100 年最大値の発現が明らかに頻繁である（図-2 の中段と下段のパネル）。他方、何らの修正を施さない元の系列では、100 年最大値の発現は、5 時期の全部で一様に 20%であった(図-2 の上段のパネル)。

このような Monte Carlo 実験の結果は、RANK データの観測網データから、信頼区間も考慮して、outliers である極端現象の長期変動検出が可能であることを示している。

４． 極端現象の検出結果

日降水量（56 地点）、最高気温（40 地点）および最低気温(44 地点)について、それぞれの地点での 100 年最大値、最高値および最低値の発現時期の年代別が、表-1 である。

表-1 100 年極値の発現時期の年代別頻度(%)。二重線は有意な変化の時期を示す。

年代	日降水量	最高気温	最低気温
1903-1922	15	5	36
1923-1942	13	18	33
1943-1962	25	11	12
1963-1982	23	11	19
1983-2002	24	55	0
95%信頼区間（%）	11	12	13

この表では、100 年最大値などの 20 年毎の発現頻度(%)とその 95%信頼区間が示されている。この信頼区間を凌駕している変化は 5%の危険率で有意だとして認めることができる。日降水量の 1940 年代前半での増加、最高気温の 1920 年代前半と 1980 年代前半での上昇及び最低気温の 1940 年代前半と 1980 年代前半での下降が、有意性を持っ

て検出された。

さらに、観測点所在地の人口別の長期変化を調べた（山元・奥田・伊ヶ崎、2004)。人口 50 万以上の大都会での 10 分間降水量が 1980 年代前半に有意に増加していることが検出された（表-2)。しかし、大都会での日降水量と時間降水量については、有意な増加は検出されなかった（ここでは、頻度などのデータは省略する)。

表-2　10 分間降水量の 60 年最大値の発現時期の人口別・年代別頻度(%)。二重線は有意な変化の時期を示す。

観測所の所在地の人口	50 万以上	50 万以下
1943～1962	34	39
1963～1982	21	28
1983～2002	45	33
95%信頼区間	18	7

同じ 1980 年代に、最高気温の有意な上昇と最低気温の有意な減少が、人口の大小に関係なく認められた(これらについても、ここでは、頻度などのデータは省略する)。

５．　議論

統計的有意性をもって検出された長期変化は、何らかの気候強制力の結果だと信じられるので、極端な天気の動向の予測などの問題探求のきっかけとなる。

検出された長期変化が、確率密度分布における平均値の変化によるのか、その分散の変化によるのか、それとも両者の変化によるのか、の識別は、この論文で提案した統計手法だけでは判断できない。各年の極値データを用いて、別途の解析を行う必要がある。

河川堤防などの設計基準として、今までは、極端天気の既往最大が採用される場合が多かったが、それに対する反省材料を、この研究が提示している。

６．　結語

極端な天気現象について、気象庁の極値順位の観測網デ

ータから、統計的に有意な長期変化を検出する手法を述べた。

引用文献

Ang,A.H-S.&W.H.Tang(1984) :Probability Concepts in Engineering, Planning and Design;Decision, Risk and Reliability. John Wiley &　Sons. 伊藤学　らによる翻訳(1988) :土木･建築のための確率･統計の応用 584p.　丸善。

伏見正則（1989）:乱数。UP 応用数学選書 12、東京大学出版会　164 pp.

Gumbel, E. J.,(1958):Statistics of Extremes. Columbia Univ. Press, 375PP.　　河田龍夫らによる翻訳書（1962）:極値統計学―極値の統計とその工学的応用。405PP.生産技術センター新社。

Hoel,P.G.(1966):Elementary Statistics. John Wiley. 浅井晃と村上正康による翻訳書(1970):初等統計学, 331pp。培風館

Iwashima,T.&　R.Yamamoto (1993):A statistical analysis of the extreme events. Long-term trend of heavy daily precipitation. Journ.Meteor. Soc.,Japan, Vol.71, pp.637-640.

Potts,A.S.(1982):A preliminary study of some recent heavy rainfalls in the Worthing area of Sussex.Weather,Vol.37, pp.220- 227.

高橋浩一郎(1961):気象統計（気象学講座、第８巻）、地人書館

Wilks,D.S.,1995:Statistical Method in the Atmospheric Sciences, 464p, Academic Press

Yamamoto, R.(1996):Validation of　the method of detecting the long-term trend of extremely heavy rainfall by using the Monte Carlo method. Journ. Meteor. Soc., Japan. Vol.74, pp.387-391.

山元龍三郎、奥田昌弘、伊ヶ崎英雄（２００４）:極端な気象現象の長期変動の実態―都市化の影響。日本気象学会2004 年秋季大会講演予稿集、１２９　P.