

日本の降雨の N 年間最悪値の統計

唐沢好男

本レポートは、日本の降雨強度の長期変化特性をまとめた技術レポート YK-086 [1] の姉妹編である。電波伝搬特性に影響を与える気象現象の研究分野は電波気象と呼ばれ、電波の減衰を引き起こす降雨はその代表的なものである。レポート [1] では、我が国の降水量に着目し、年間最大値データ（極値データ）の統計的性質を線形や非線形関数による回帰分析を行い、長期変化傾向を明らかにしている。データは、気象庁が 100 年を超える長期間の気象データをホームページで公開しており、その中の、年毎の日・1 時間・10 分間の最大降水量を用いた。本レポートでは同データを用いて、10 年、100 年、あるいは 1000 年に一回起きるような強雨の最悪値（極値の中の極値）の特性を調べる。想定外の出来事に対して想定外としないための心構え（備え）に役立つ情報になる。この解析には、極値統計学が力を発揮する。

本レポートは YK-083 初版(Jan. 2024)に若干の記述追加を行った改訂版である。

1. はじめに

姉妹編レポート（YK-086 [1]）では、我が国の強雨現象の長期的な変化を調べる目的で、気象庁が公開している気象データベース [2] から、各種降水量（日降水量・1 時間降水量・10 分間降水量）の年間最大値を用いた統計解析を行った。日本全国 47 地点、最大 130 年（ただし、降水量によって地点数と年数は異なる）のデータを地点ごとの平均値で正規化した正規化降水量について、95%信頼区間推定を行い、日本全体では、最大降水量が 100 年スケールで見ても十数%程度増加している傾向を読み取ることができた。さらに、非線形関数近似と赤池情報量基準（AIC）による最適オーダ選択によって、その増加が、1990 年頃より顕著になってきていることを見出した。しかし、最大降水量の年毎のばらつきが 40% 程度と大きく、100 年間全体で見たとき、その変動の特徴は定常不規則過程の範疇にあるとみなしても良さそうにも思える。

そこで、本レポートでは、レポート [1] で用いたと同じ正規化降水量を用い、全データが確率的に均一と仮定し、そのデータを用いて、10 年、100 年、さらには、1000 年に一回の強雨量を求める。災害は忘れたころにやってくると言われるが、想定外の出来事を想定外としないための心構え（備え）に役立つ情報になる。この種の解析では、極値統計の理論（Gumbel 分布）がデータ解析結果の解釈に力を発揮することになる。果たして、何が現れてくるであろうか。尚、本レポートの内容は論文 [3] の後半部にまとめており、当該部分の和文解説である。

2. 解析のための道具：極値統計

(1) 最大値の分布

ある確率分布（確率密度関数 f 、累積分布関数 F ）を持つ母集団があり、その中から独立に n 個の標本を得たとき、小さいほうから i 番目の値の確率分布を求める学問が順序統計である。この確率分布（確率密度関数 $f_{(i)}$ 、累積分布関数 $F_{(i)}$ ）は次式となる。

$$F_{(i)}(x) = \sum_{k=i}^n \binom{n}{k} F^k(x) \{1-F(x)\}^{n-k} \quad (1a)$$

$$f_{(i)}(x) = \frac{dF_{(i)}(x)}{dx} = n \binom{n-1}{i-1} F^{i-1}(x) \{1-F(x)\}^{n-i} f(x) \quad (1b)$$

$$\binom{n}{k} \equiv \frac{n!}{k!(n-k)!} \quad (0 \leq k \leq n)$$

最大値の分布は $i=n$ であるので、その確率分布（確率密度関数 f_{\max} 、累積分布関数 F_{\max} ）は次式となる。

$$F_{\max}(x) = F_{(n)}(x) = F^n(x) \quad (2a)$$

$$f_{\max}(x) = f_{(n)}(x) = nF^{n-1}(x)f(x) \quad (2b)$$

最大値の期待値（平均値） $\langle x_{\max} \rangle$ は次式で求められる。

$$\langle x_{\max} \rangle = \int_{-\infty}^{\infty} x f_{\max}(x) dx \quad (3)$$

確率分布が $0 \sim 1$ 間で一様分布の場合、 $F_{\max}(\langle x_{\max} \rangle) = n/(n+1)$ であることに倣って、任意の分布の場合でも

$$\langle x_{\max} \rangle \approx F_{\max}^{-1} \left(\frac{n}{n+1} \right) \quad (4)$$

と近似できる。累積分布関数の逆関数が閉形式で得られている分布系ならば、簡易な算定が可能である。

(2) 極値分布（Gumbel 分布）

(2)式で表される最大値の分布は、標本数 n が十分大きいとき、3種類の形の分布のどれかに収束することが導かれている（極値統計学（例えば[4]））。裾（ x の大きい値の部分）の広がり指数関数的な分布であれば、その極値分布は Gumbel 分布と呼ばれる形になる（Gumbel の邦訳はガンベルとグンベルの両方がある）。電波伝搬に現れる大部分の分布、例えば、正規分布、指数分布、ガンマ分布、対数正規分布、ワイブル分布などは Gumbel 分布に収束する。Gumbel 分布（確率密度関数 f_G 、累積分布関数 F_G ）は次式となる。（一般的

導出は難しいが、例えば、(2)式で指数分布として n の無限大極限をとると Gumbel 分布になることが容易に確かめられる)。

$$F_G(z) = \exp\{-\exp(-z)\} \quad (-\infty < z < \infty) \quad (5a)$$

$$f_G(z) = \exp\{-z - \exp(-z)\} \quad (5b)$$

$$z = \frac{x_{\max} - b_n}{a_n}$$

図 1 に Gumbel 分布の形を示す。式中のパラメータ a_n , b_n は吸引係数と呼ばれ、元の確率分布のパラメータとの関係式が導かれている ([4] の表 2.4 に種々の分布に対する関係式が示されている)。1 時間降水量や 10 分間降水量の分布はガンマ分布に近似できることが調べられており [5]、その最大値の分布 (極値分布) が Gumbel 分布に収束するであろうことはよいのであるが、本レポートの 4 節での解析では、最大値が属するもとの分布に立ち返ることはしないので、これ以上の説明は割愛する (ガンマ分布と対数正規分布についての n に対する収束の度合いは [6] の図 11.8 に示している)。

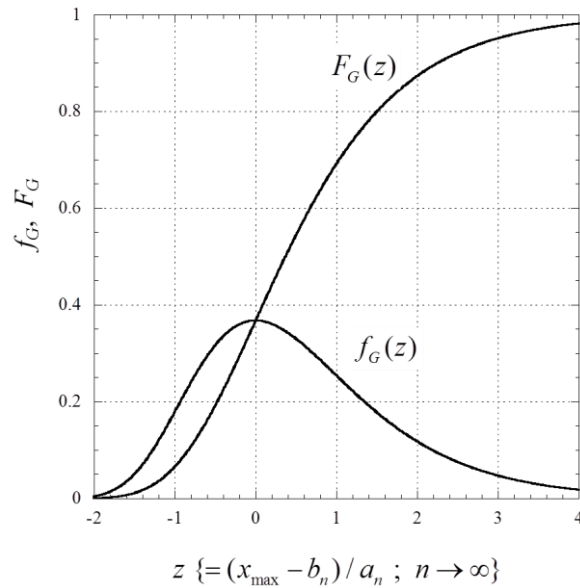


図 1 Gumbel 分布

Gumbel 分布する確率変数の期待値 (平均) m_G と分散 V_G は次式である。

$$m_G = \langle x_{\max} \rangle = b_N + \gamma a_N \quad (\gamma: \text{オイラーの定数}=0.5772\cdots) \quad (6a)$$

$$V_G = \left\langle (x_{\max} - m_G)^2 \right\rangle = \frac{\pi^2 a_n^2}{6} \quad (6b)$$

元の分布形が分からなくても極値データの平均値と分散が得られれば、吸引係数 a_n, b_n が求まり、Gumbel 分布の形が定められる。

3. 降水量データ

気象庁からは、日本各地（約 1,300 箇所）の気象データがホームページから公開されている[2]。本解析ではこのデータを使用する。使用するデータの詳細はレポート[1]の 2 節で詳しく説明しているため、ここではその概要のみを述べる。ここで扱うデータは、日降水量、1 時間降水量、10 分間降水量それぞれの年間最大値であり、以下のものである。

- ・日降水量：日本全国 45 地点、1924 年～2023 年の 100 年間（データ総数 4500 点）
- ・1 時間降水量：日本全国 45 地点、1944 年～2023 年の 80 年間（データ総数 3600 点）
- ・10 分間降水量：日本全国 47 地点、1954 年～2023 年の 70 年間（データ総数 3290 点）

各降水量は、それぞれの地点の対象期間での平均値で割って正規化し、以下の統計解析ではその正規化データを用いる。日降水量・1 時間降水量・10 分間降水量は、それぞれ、 $n=365, 8760, 52560$ に相当するが、1 年の中での降水量の生起は定常確率過程ではないので、最大値の分布について(2)式が単純に適用できるわけではない。しかし、 n の数が十分大きいという条件が満たされるならば、最大値の分布が(5)式の Gumbel 分布に近づくことが期待できる。

なお、我が国において、降水量の最大値を与えるような降水事象はその 100%が降雨であるので、本レポートで着目する降水量は降雨量と読み替えて問題ない。

4. データ解析結果

4. 1 累積分布特性

(1) 正規化日降水量年間最大値

Gumbel 分布は、最大値を得た元の標本値の数 n が十分大きい（理論的には無限大）という条件で得られる分布である。故に、本レポートで扱う三つの降水量の中では、日降水量の年間最大値が、最も近似度は低い分布になるであろうと予想される。そこで、この分布から見てゆきたい。図 2 は 4500 個の正規化最大値について、発生回数のヒストグラムとそれに合わせるように吸引係数を調整した Gumbel 分布を示している。図より、近似度が劣ると予想される日降水量においても、かなりよく一致しているが、子細に見ると、正規化降水量が大きい部分 (>1.5) で観測値の方がやや多めになっている。ここでは、10 年あるいは 100 年に一回発生するような極値中の極値の特性を調べることを目的にしている。そのためには、強い降水量の分布の裾の方で一致していないと、理論計算値と観測値に大きなずれが出てしまう。そこで、その部分を拡大して調べるために累積分布関数 F の代わりに $1-F$ で与えられる相補累積分布関数を用いる。

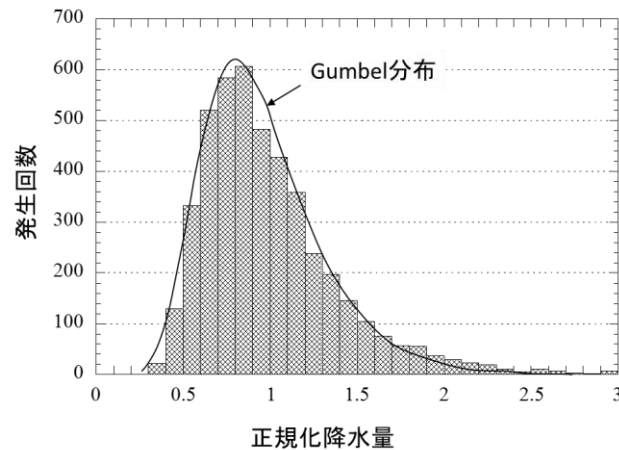


図2 年間最大正規化日降水量の発生回数ヒストグラムとそれを近似する Gumbel 分布

図3は、その相補累積分布関数を示している。図には、観測値と二つの Gumbel 分布を示している。Gumbel 分布（青線： $a_n=0.315, b_n=0.818$ ）は、観測値の平均値と分散から式(6)を用いて、吸引係数を求めている。図より、同線は正規化降水量 1.5 以下では観測値とよく一致しているが、それより大きい値では観測値より小さい値になっている。そこで、この部分の精度を上げるために相補累積確率が 0.01 と 0.001 の2点で観測値と一致するように吸引係数を定めると、黒線（ $a_n=0.368, b_n=0.705$ ）のようになる。以下の解析での理論推定においては後者の吸引係数で求めることにする。

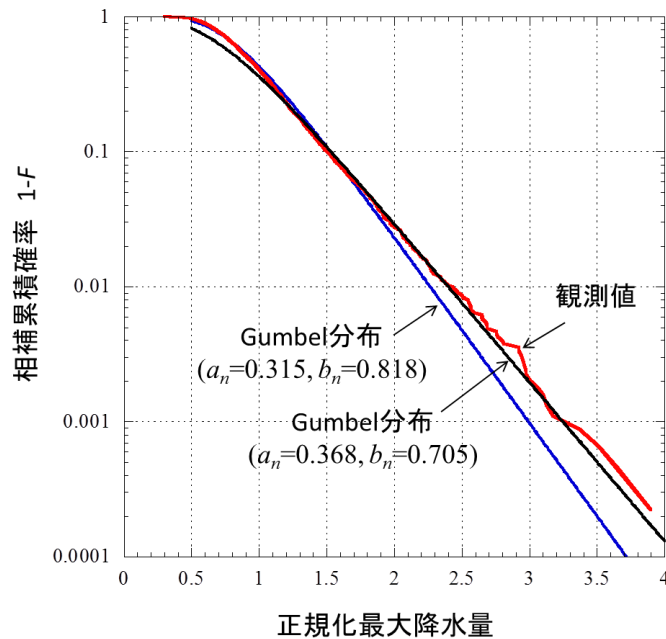


図3 正規化日最大降水量の相補累積分布特性

(2) 正規化 1 時間降水量と 10 分間降水量の年間最大値

図 4 は正規化 1 時間降水量と 10 分間降水量の年間最大値の相補累積分布をまとめて示している。両方とも吸引係数は観測値（正規化したもの）の平均値と分散から求めており、分布全体に亘って Gumbel 分布によく合っていることが分かる。（得られた吸引係数の値は図中に）。

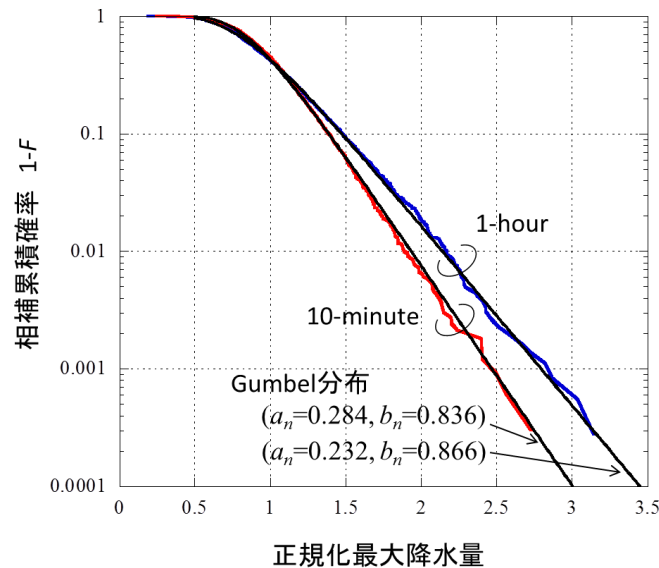


図 4 正規化時間最大降水量と 10 分間降水量の相補累積分布特性

4. 2 N 年間の最大値特性

100 年に一度というような豪雨は、想定外な出来事として当該地域に大規模災害をもたらすだろうが、それがどの程度のものか予想できていれば、それに対する心構えができ、備えを十分にすることができるであろう。そこで、ここでは、 N 年間の最大降雨について調べたい。

観測値については以下のような処理を行う。各地点ごとの正規化年間最大降水量について、 N 年単位で分割してそれぞれの最大値を求める。日降水量であれば 100 年間データなので、割り切れる数値として $N=5, 10, 20, 50, 100$ を選ぶ。例えば、 $N=50$ の場合は、1 地点で二つの、全国で合わせて 90 個の最大値が得られる。この平均値を N 年に一度の値とする。1 時間降水量では 80 年分なので、 $N=5, 10, 20, 40, 80$ と、10 分間降水量では 70 年分なので $N=5, 10, 14, 35, 70$ とした。

Gumbel 分布での N 年最大値の平均値は、まず、(2)式で最大値の分布 f_{\max} を求める。この時、元の分布の f と F を Gumbel 分布にすればよいのであるが、(5b)は変数 z に対する式なので、確率密度関数 f_G にヤコビアン ($1/a_n$) を掛ける。この分布を用いて、(3)式により、平均値を求める。 N 年最大値の平均値を x_N と置くと、具体的な計算式は次式になる。

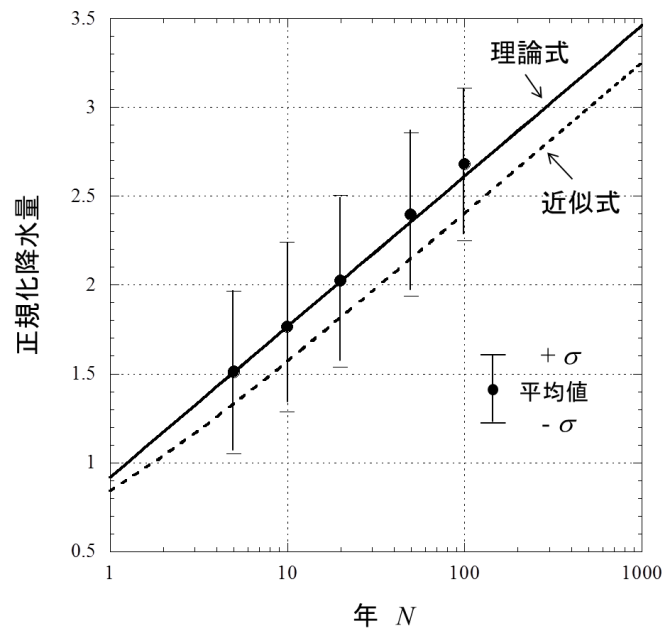
$$x_N = \frac{N}{a_n} \int_0^\infty x \{F_G(z(x))\}^{N-1} f_G(z(x)) dx \quad (7)$$

この式はこれ以上簡易にはなりそうもなく、数値積分で解く。一方、近似式ではあるが平均値を与える式として(4)式でも求められる。Gumbel 分布の累積分布関数 F_G は、逆関数が閉形式で求まるので、次式で計算できる。

$$x_N \approx F_G^{-1}\left(\frac{N}{N+1}\right) = -a_n \log\left\{-\log\left(\frac{N}{N+1}\right)\right\} + b_N \quad (8)$$

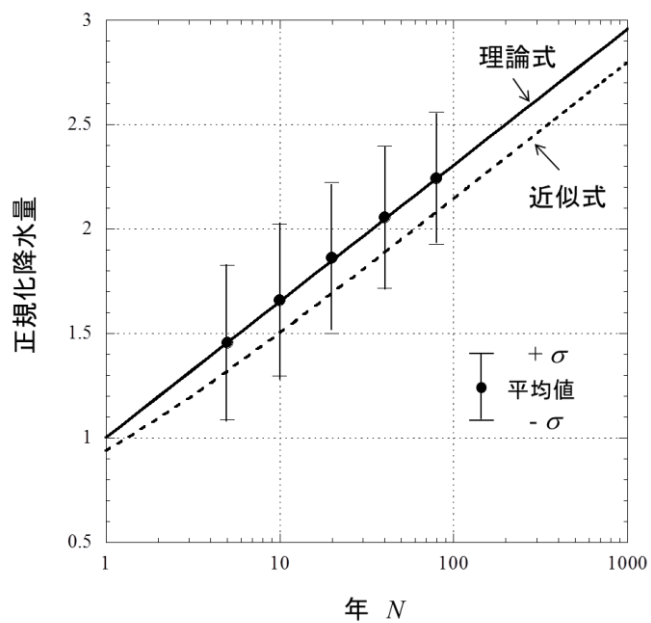
各降水量に対する吸引係数 (a_n, b_n) は先に示したものをを用いる。

図5はこの解析結果を示している。観測値から得られた最大値にはばらつきが大きいので、平均値と共に±標準偏差 (σ) も示している。どの降水量についても、観測値の平均値と(7)式での計算値は非常によく一致している。(8)式による近似推定値はやや小さめの値を示している。その場合でも標準偏差の範囲内に納まっており、簡易に推定するにはこの式でもよいであろう。(筆者メモ：(8)式の推定値は最頻値（モード）に対して良い推定になっているはずである)。

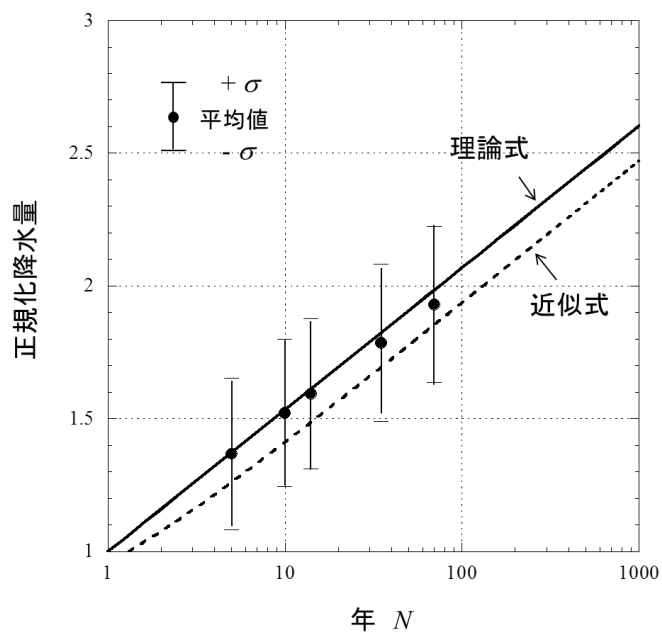


(a) 日降水量

図5 N 年間正規化最大降水量（観測値と二つの推定法との比較）



(b) 1 時間降水量



(c) 10 分間降水量

図 5 (続き)

4. 3 仮想長期間での N 年間の最大値

上記では、地点ごとに N 年間の最大値を求めその平均値を示したが、観測値が 100 年程度しかないので、それ以上長い期間のデータがない。そこで強引ではあるが、全地点のデータを一つにまとめて、日降水量・1 時間降水量・10 分間降水量で、それぞれ、4500 年、3600 年、3290 年分のデータがあったとしよう。それぞれの地点の平均値で正規化しているので、全体としてのばらつきの統計を見るのに非合理ではないと言う理屈である。地点データに相関が強いと、このやり方は間違った結果を生み出すが、今回のデータでは地点を都道府県単位に選んでいるので、無相関を期待して良いであろう。理論計算は N の値が拡張されただけの違いであり、ここでは(7)式による平均値を示す。観測値は値を一つしかもたない端の点を除き、割り切れるように適当に分割してそれぞれの平均値を示す。

図 6 はこの解析結果である。同図より、このような拡張をして解析しても、極値統計 (Gumbel 分布) はそれを正しく予想していることが分かる。極値統計学、お見事、である。言うことは、悪いことも起こるべくして起きているわけであり、想定外の出来事は無いと教えてくれているのである。少なくともこの 100 年間を対象にした降水量の最大値の解析においてはである。

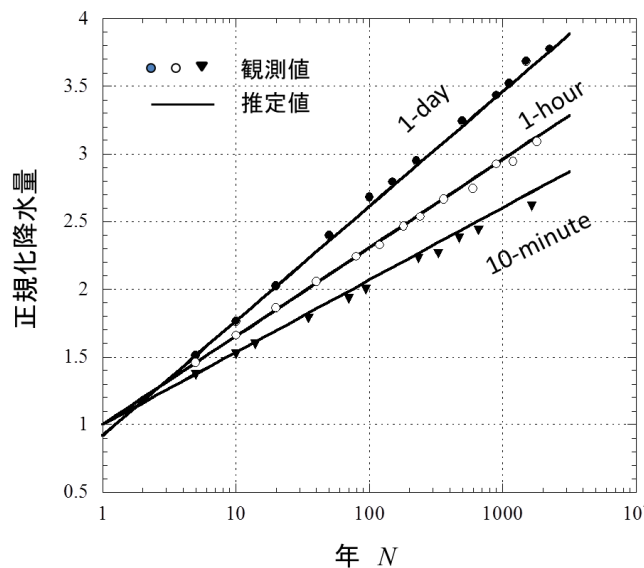


図 6 N 年間正規化最大降水量 (観測値と Gumbel 分布を適用した推定値)

東京(千代田区)の 100 年間(1924~2023)の日降水量の年間最大値の平均値は 125.7mm である。日降水量の 10 年・100 年・1000 年に対する比率は、図 6 より約 1.8, 2.7, 3.5 倍なので、それぞれ、226mm, 339mm, 440mm と推量される(図 7 に赤線で示す)。因みに、この 100 年間での東京での最大値は 371.9mm (1958 年 9 月の狩野川台風)、2 番目は

287.3mm（1938 年 7 月の集中豪雨）である。各地の各降水量の平均値 ($\langle R_{\max}(1) \rangle$; [1] の表 1 に示している) を用いれば、 N 年間の最悪値を求めることができる。

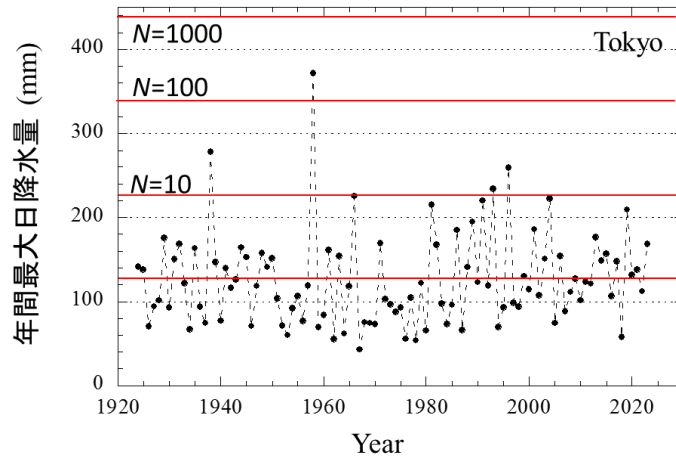


図 7 東京における年間最大日降水量の 100 年間特性と $N=10, 100, 1000$ 年に対する推定値（赤線）

5. 雑感

最近よく聞く強い雨に対する呼び方で、「これまで経験したことない大雨」、「未曾有の豪雨」、「観測史上で初めて」などがあり、想定外の出来事が何か頻繁に起きているような印象を受ける。本当にそうであろうか。

もし、そうであれば、本レポートで扱った各種降水量の年間最大値特性には、その変化が現れるであろうと考える。最大日降水量の増加は洪水を、1 時間降水量あるいは 10 分間降水量の増加であれば強雨の発生をもたらすはずである。しかし、今回の解析結果は、4 節で述べたように、凡そ 100 年間全体に亘って、極値統計理論（Gumbel 分布）が予想するようにしか現れていないと気づくであろう。そう、降雨による非常災害は、起きるべくして起こっている、という以上のことは言えない。この解析を通じて、筆者はその感を一層強くしている（注）。

一つの地点で見れば 1000 年に一回の大惨事も、気象的に独立な 50 地点（パラレルワールド）を俯瞰してみれば、エリア全体としては、20 年に一回はどこかで起きる最大値と読み替えることもできる（図 8）。テレビのニュースで知るのはこの出来事である。

注： レポート[1]で行った長期間の傾向変化の解析で、今回着目した年間最大降水量については、凡そ 100 年間で 10%程度の値の増加が有意に認められることを述べた。また、その増加が、1990 年以降に顕著であることも述べた。さらに、気象庁が強雨の降り方に関して、1 時間降雨強度が 50mm を超える発生回数がこの 50 年で 1.5 倍程度になっているデータを示している（[1]の図 3 で紹介している）。このように長期的な意味で雨の降り方も変わりつつあるのであるが、年毎のばらつきの方がはるかに大きくて、全体を定常確率過程にあるデータとしてみても、極値統計理論が予測する範囲を超えていないという意味である。なお、これからの 100 年と言った未来を見れば、降雨強度の継続的な増加によって、これまでのような定常不規則過程の扱いができなくなってくるかもしれない。これも心に留めておいてほしい。

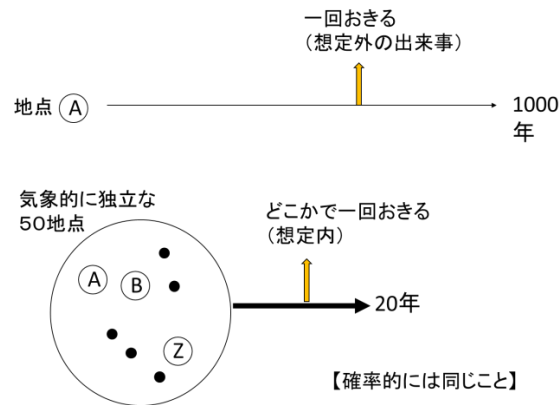


図8 想定外の出来事もパラレルワールドで見れば・・・

参考文献

- [1] 唐沢好男, “日本の極値雨量の長期変化傾向,” 技術レポート (私報) YK-086_rev2, 2025.
http://www.radio3.ee.uec.ac.jp/ronbun/YK-086_Rainfall_Statistics_5.pdf
- [2] 気象庁, 過去の気象データ検索, 気象庁ホームページ, <https://www.data.jma.go.jp/obd/stats/etrn/index.php>
- [3] Y. Karasawa, “Long-term statistical properties of extreme rainfall data in Japan,” *IEEE Trans. Geoscience and Remote Sensing*, (to be published in 2025). Early Access area on IEEE Xplore:
<https://ieeexplore.ieee.org/document/10640060>
- [4] 高橋倫也, 志村隆彰, 極値統計学, 近代科学社, 2016.
- [5] 森田和夫, 樋口伊佐夫, “降雨による電波の減衰量に関する統計的研究,” 研実報, vol. 19, no. 1, pp. 97-150, 1970.
- [6] 唐沢好男, 電波伝搬モデルの基礎 (第 11 章 極値統計), ネット公開私製本,
http://www.radio3.ee.uec.ac.jp/ronbun/Probability_Distributions_Chap_11.pdf

読者の皆様へ

本レポートは下記レポートの姉妹編です。併せて見ていただけると、降雨特性 100 年の全貌が見えてきます。

唐沢好男, “日本の極値雨量の長期変化傾向,” 技術レポート (私報) YK-086_rev2, 2025.

http://www.radio3.ee.uec.ac.jp/ronbun/YK-086_Rainfall_Statistics_5.pdf