日本学術会議 URSI-F No.632
7 March., 2019

極値統計データにみる日本の降雨特性

～地球温暖化の影響、有りや無しや～

唐沢　好男

発表の概要

観測データを客観的に分析するための基礎知識として、私報の［1]で回帰分析と信頼区間を、[2]で最大値の性質を扱う極値統計学を取り上げ、その要点をまとめている。本発表では、この手法を用いて、無線通信にも関係深い物理量である降雨量に着目し、年間極値データ（年間最大時間降水量など）の統計的性質を調べる。データは、気象庁が100年スケールの長期統計データとしてホームページで公開している年毎の日・時間・10分間の最大降水量（年間の最大値を与える降水現象は降雨によるので、以下、降雨量と呼ぶ）[3]を用いる。

近年、地球規模での温暖化現象や、都会でのヒートアイランド現象などにより、長期傾向としての気温の増加が報告されている。では、この気温の増加が、降雨現象に影響を与えているであろうか？本発表の前半では、気温の長期増加傾向が大きい大都市での降雨量極値データの長期的変化を調べる。この問題については、気象庁としての見解もまとめられているが[4]が、統計手法の応用問題として、また、それを、筆者のような気象分野の門外漢が体感を得る意味において解析を行った。本発表の後半では、日本国内において比較的多雨地域（ITU-Rの降雨エリア区分のMが目安の）にある１８のエリアの長期データを正規化してまとめて一つの確率分布とみなし、極値統計学から導かれる諸性質と比べて、降雨特性の特徴を明らかにする。また、その結果に基づき、今後、100年～1000年スケールでの最悪値予測に関する考察を行う。なお、この本文については、[5]においてYK-021として近日公開する。

参考文献

[1] 唐沢好男, “回帰分析と信頼区間：ばらつきの大きい少数データから誤った推論をしないための,” Tech. Rep. YK-019（私報）, 2019.01., http://www.radio3.ee.uec.ac.jp/ronbun/YK-019_Kukan_Suitei.pdf

[2] 唐沢好男, “極値統計学へのいざない：想定外の出来事を想定外としないために,” Tech. Rep. YK-020（私報）, 2019.02., http://www.radio3.ee.uec.ac.jp/ronbun/YK-020_KyokuchiToukei.pdf

[3] 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

[4] 気象庁, “地球温暖化情報 第９巻,” 2017.03, https://www.data.jma.go.jp/cpdinfo/GWP/Vol9/pdf/all.pdf

[5] 唐沢好男, “極値統計データに見る日本の降雨特性、http://www.radio3.ee.uec.ac.jp/report.htm より Rep. YK-21として公開予定,

この発表の動機

・地球温暖化問題とヒートアイランド現象

降雨の特性に影響が現れているかどうかへの興味

・気象観測データの充実（気象庁：１００年スケール）

・気象データ入手の容易性（特に年間最大値）

（気象庁：ホームページ公開）

・時間的余裕ができ、手間暇掛かる単純作業もOK

発表の内容

１．解析の道具立て

・回帰分析（回帰直線と信頼区間）

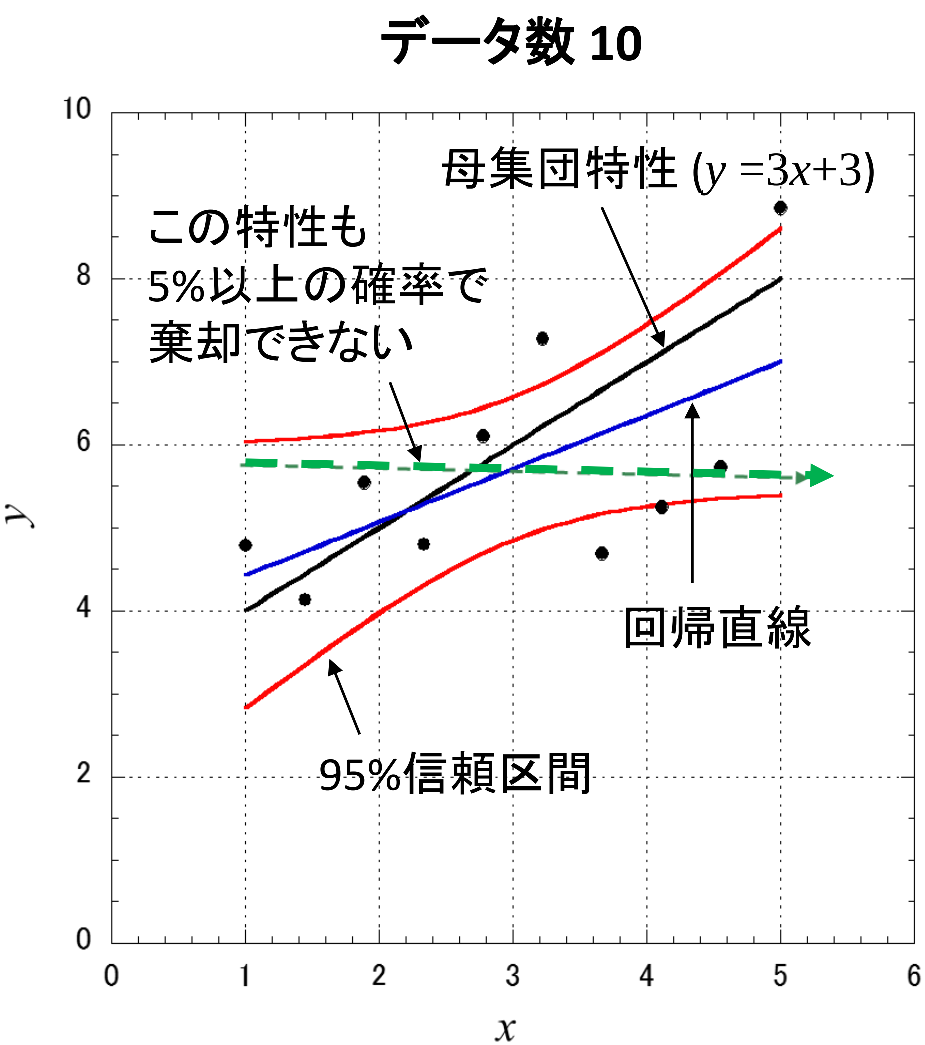
・極値統計学

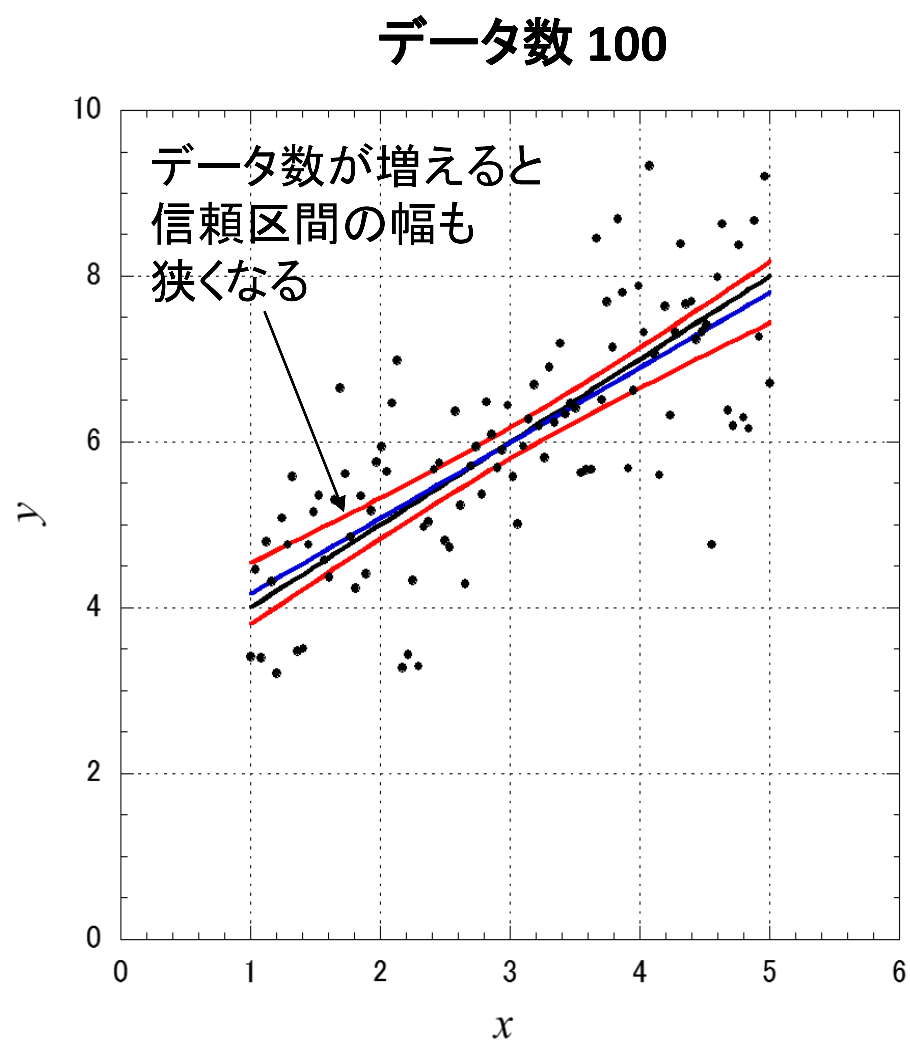
２．気象データ

３．降雨特性極値データの長期的傾向

４．極値データの統計的性質

解析の道具：　直線回帰と信頼区間





【信頼区間はデータのばらつき幅とはまったく違う *!!*】

解析の道具：　極値統計学

極値とは？

n 個の標本値の中の

最大値（または最小値）

極値の性質を扱う学問

極値統計学

今回発表では

- 日降水量
- 時間降水量
- 10分間降水量

年間最大値（降雨量）

N 年間での最大値

（*N*=1であっても、その量は既に *n*>>1の極値統計データ）

極値統計学: 最大値の確率・統計を扱う学問

最大値の確率分布は、元の量の確率分布に依存しない極値分布になる（グループメンバーの数 *n* が多い場合）

極値分布は3つのタイプに分類されるが、電波伝搬や自然現象に現れる物理量は、その大部分がグンベル（Gumbel）分布と言われるタイプになる。（正規分布、仲上m分布、仲上・ライス分布、ガンマ分布、対数正規分布の最大値はこのタイプに収斂）

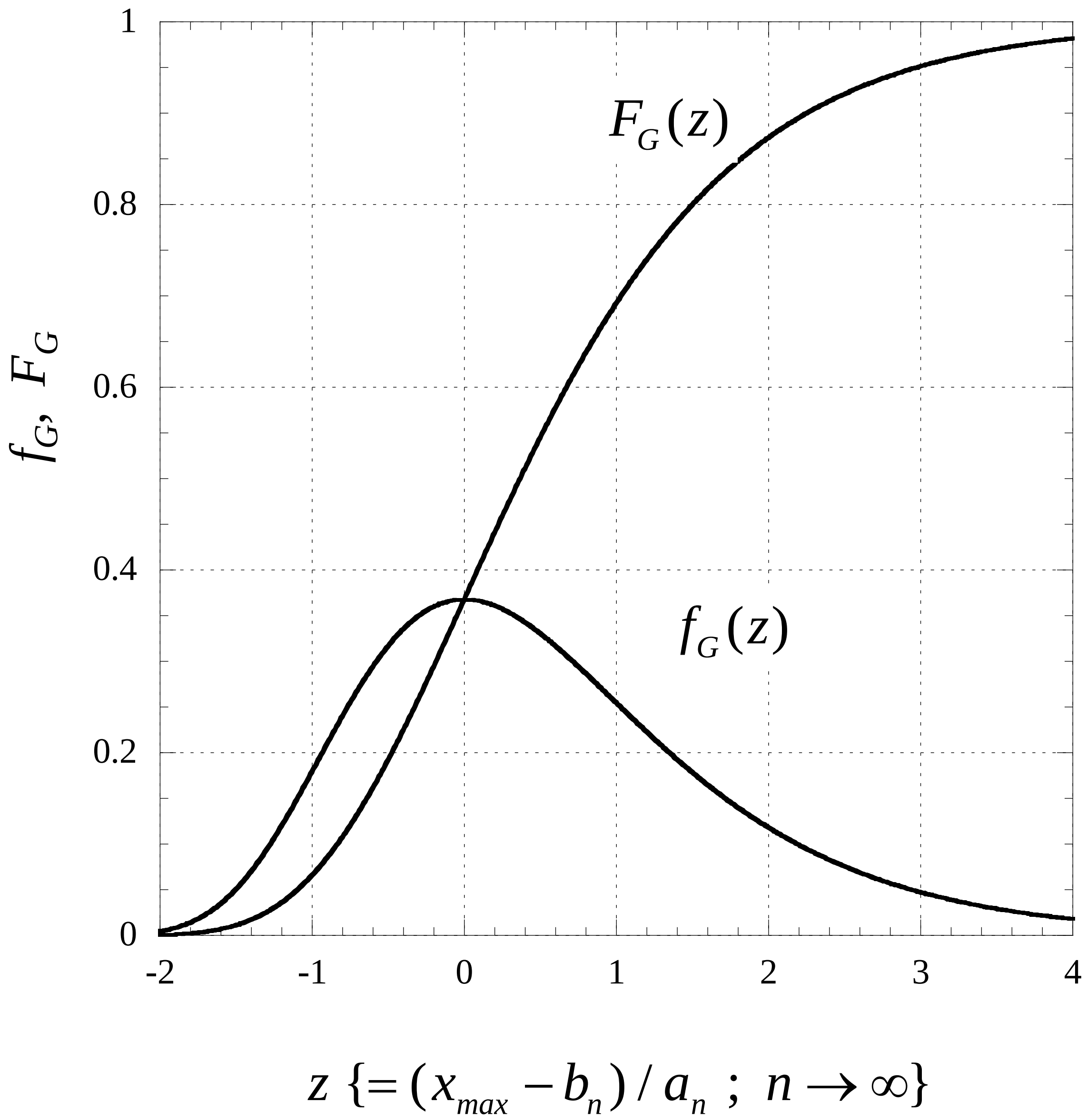
グンベル分布

$$z = \frac{x_{max} - b_n}{a_n} \quad (-\infty < z < \infty)$$

PDF $$f_G(z) = \exp\{-z - \exp(-z)\}$$

CDF $$F_G(z) = \exp\{-\exp(-z)\}$$

極値分布：グンベル（Gumbel）分布の形



メンバー数 n の中での最大値の推定

理論的には、その平均値を求めるのは難しく、最頻値(モード)を求めるのは比較的容易

最大値 x_{max} の確率分布は、順序統計の理論より

CDF $$F_n(x_{max}) = F^n(x_{max})$$

PDF $$f_n(x_{max}) = nF^{n-1}(x_{max})f(x_{max})$$

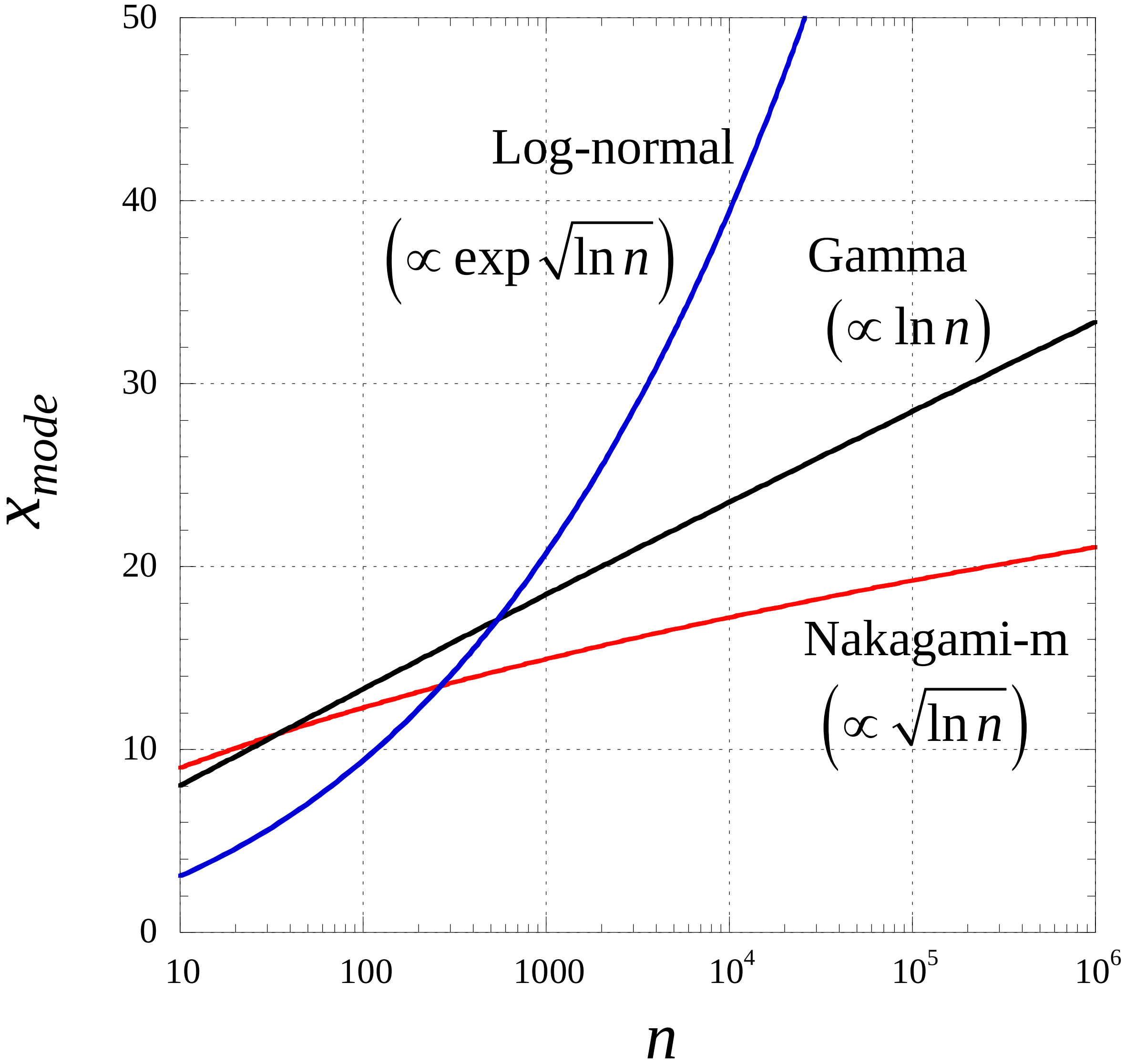
分布の最頻値(モード) x_{mode} は、PDFの微分値0のところにあり

$$(n-1)f^2(x_{mode}) + F(x_{mode})f'(x_{mode}) = 0$$

さらにいくつかの仮定や近似を経て、x_{mode} と n の関係は

$$n \approx \frac{-f'(x_{mode})}{f^2(x_{mode})} \equiv \Phi(x_{mode}) \quad \rightarrow \quad x_{mode} \approx \Phi^{-1}(n)$$

最大値の最頻値と $\ln n$ の関係に関する分布形状依存性



降雨量の確率分布
- 対数正規分布
- ガンマ分布
- （細矢の）M分布

極値データの統計処理からは**ガンマ分布**の特徴が現れてくると予想できる

用いる降雨量データ

気象庁のホームページで公開されているデータ
そのごく一部を利用

地点：日本の１８ヶ所（比較的降雨が多い地域：
ITU-R降雨気候区Ｍ目安で、100km程度以上離れている）

着目する降水量

(1)日降水量の年間最大値　　（平均117年間）

(2)時間降水量の年間最大値　（94年）

(3)10分間降水量の年間最大値（76年）

上記降水量は例外なく降雨時のものであり、
以下「降雨量」と呼ぶ

ホーム > 各種データ・資料 > 過去の気象データ検索 > 年ごとの値

気象庁のホームページ

年ごとの値

一覧表 グラフ

主な要素 詳細（気圧・降水量） 詳細（気温・蒸気圧・湿度） 詳細（風） 詳細（日照・雪・その他）

地点

降水量（最大）

東京 年ごとの値 主な要素

年

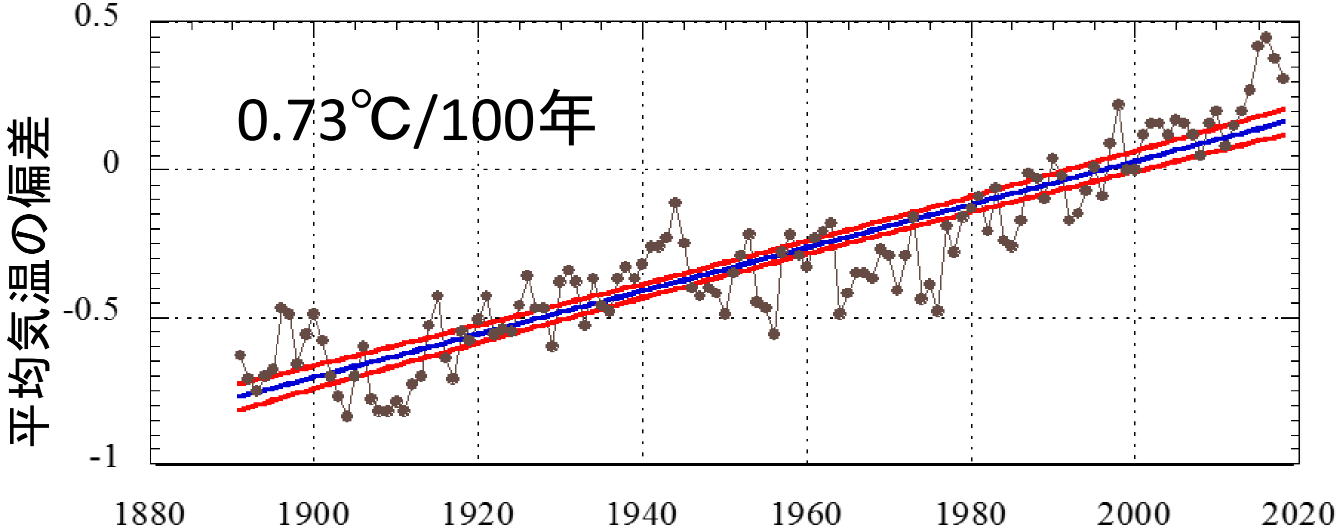
年	気圧(hPa)		降水量(mm)				気温(℃)					湿度(%)		風	
	現地	海面	合計	最大			平均			最高	最低	平均	最小	平均	最
	平均	平均		日	1時間	10分間	日平均	日最高	日最低					風速	風速
1875			1219.2]	94.0]			17.0]	21.9]	12.2]	35.1]	-3.6]				
1876		1014.7	1755.5	76.2			13.6	18.7	8.3	35.6	-9.2	78			
1877		1015.2	1317.3	79.9			14.2	19.1	8.9	34.9	-4.8	77			
1878		1015.2	1764.2	150.9			13.8	18.3	9.6	35.1	-7.6	79			
1879		1014.5	1492.7	99.8			14.6	19.4	10.0	33.9	-5.5	77			
1880		1015.0	1685.7	116.7			14.1	19.2	9.1	33.2	-6.8	76			
1881		1015.1	1444.4	86.4			13.8	18.8	8.5	34.2	-8.4	78			
1882		1015.2	1478.3	109.7			14.0	18.6	9.0	34.2	-6.3	77			
1883		1015.0	1552.6	125.6			13.3	17.8	9.1	32.8	-7.8	76			
1884		1014.9	1314.8	72.2			12.9	17.7	8.6	33.3	-7.7	76			
1885		1014.9	1531.7	149.5			13.1	17.5	9.1	31.8	-9.1	77			
1886		1014.5	1290.3	64.5	12.8]		13.9	18.7	9.9	36.6	-7.7	75			
1887		1013.6	1250.0	77.8	20.5]		13.8	18.6	9.7	32.6	-7.9	76			
1888		1013.7	1378.5	78.9	19.2]		13.5	18.5	9.3	32.8	-7.5	75			
1889		1014.0	1319.3	162.0	22.0]		13.3	18.0	9.3	33.3	-8.1	74			
1890		1013.8	1958.2	123.8	×		15.0	19.4	11.0	33.9	-5.8	77			
1891		1014.1	1220.8	109.0	23.2]		14.4	19.3	10.0	34.4	-6.7	72			
1892		1013.7	1715.1	84.5	20.5]		14.0	18.6	10.0	34.4	-6.8	73			
1893		1014.0	1161.3	56.3	28.1]		13.8	18.8	9.4	34.4	-7.2	71			
1894		1014.4	1320.8	132.8	52.9]		14.8	19.6	10.6	35.1	-5.3	73			
1895		1013.8	1397.8	133.0	22.6]		13.8	18.7	9.8	33.0	-5.7	74			
1896		1014.4	1373.9	74.0	22.2]		14.0	18.7	10.1	34.3	-5.6	74			
1897		1014.6	1497.2	73.2	34.1]		13.2	17.7	9.2	34.0	-7.0	75			

データ解析に用いた18地点、および各種年間最大降水量の平均値

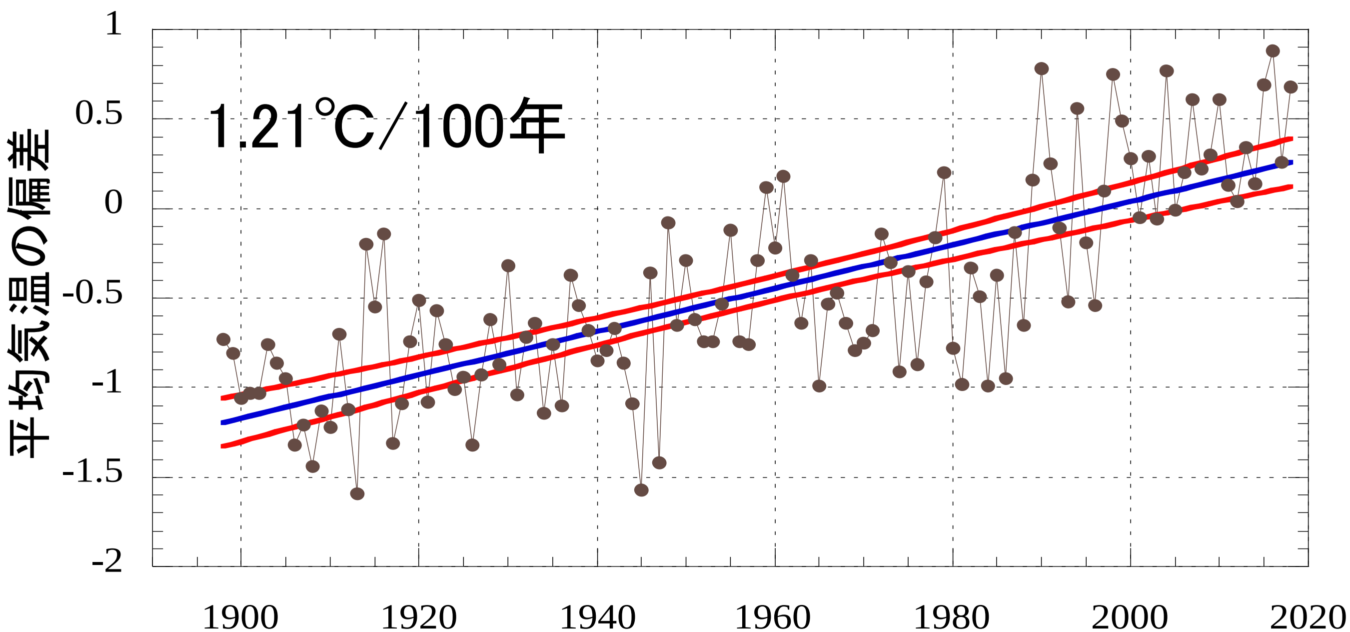
（数値の単位は mm、( )は観測年数)

Site	1-day	1-hour	10-minute
Sendai	100.47 (92)	32.113 (78)	12.365 (78)
Mito	106.13 (121)	36.067 (113)	14.930 (82)
Katsuura	143.70 (106)	52.517 (69)	16.936 (69)
Tokyo	119.15 (143)	42.057 (98)	16.542 (79)
Kofu	99.707 (124)	30.996 (82)	12.897 (79)
Shizuoka	186.60 (79)	52.351 (79)	16.947 (79)
Iida	103.03 (121)	30.499 (71)	11.908 (71)
Nagaya	111.37 (128)	40.477 (128)	16.084 (75)
Osaka	93.054 (135)	34.781 (115)	14.481 (77)
Wakayama	117.16 (134)	40.228 (75)	15.195 (75)
Hiroshima	107.67 (140)	35.794 (130)	14.608 (80)
Hamada	107.48 (126)	40.009 (75)	14.333 (75)
Takamatsu	96.647 (77)	31.496 (77)	13.045 (77)
Kochi	194.95 (133)	60.426 (82)	18.844 (79)
Fukuoka	122.72 (128)	39.533 (122)	16.017 (80)
Kumamoto	154.30 (128)	44.957 (128)	17.112 (82)
Kagoshima	150.16 (135)	49.488 (108)	17.840 (72)
Naha	184.70 (67)	61.131 (66)	19.564 (66)

地球温暖化問題：それは確実にある



世界（地上＋海上）

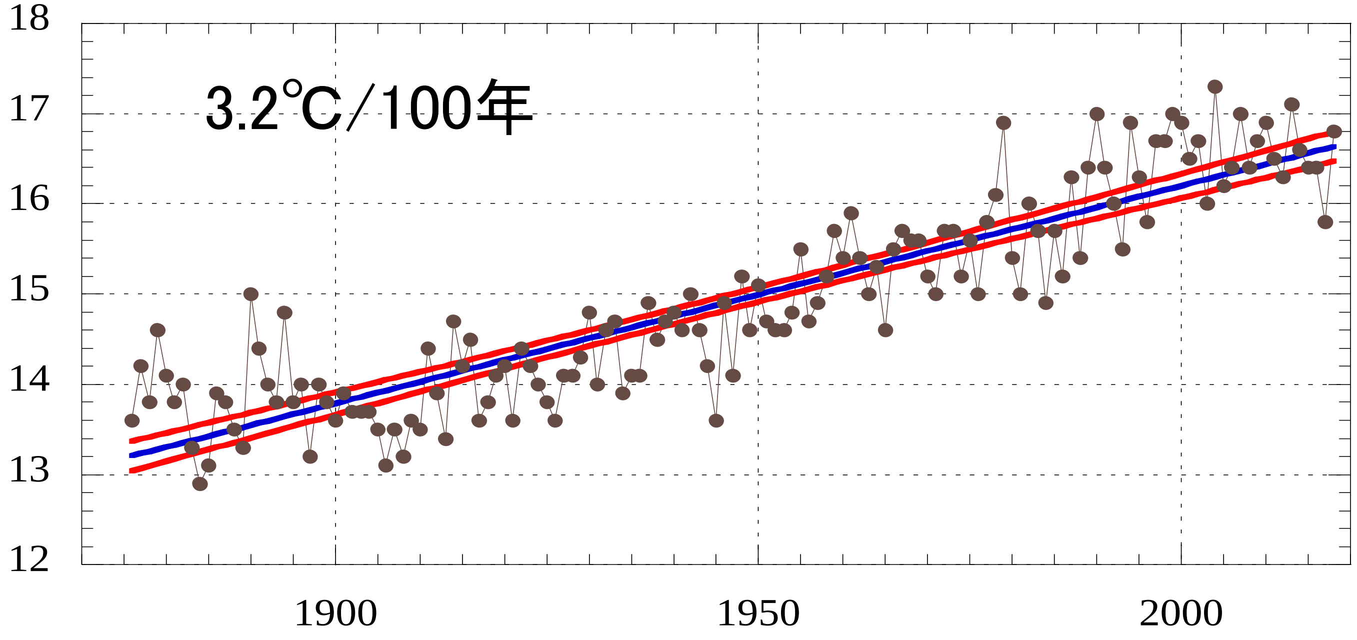


日本（都市化による影響が比較的小さい15地点の平均）

（数値データは気象庁ホームページより：以下同じ）

地球温暖化問題＋ヒートアイランド現象

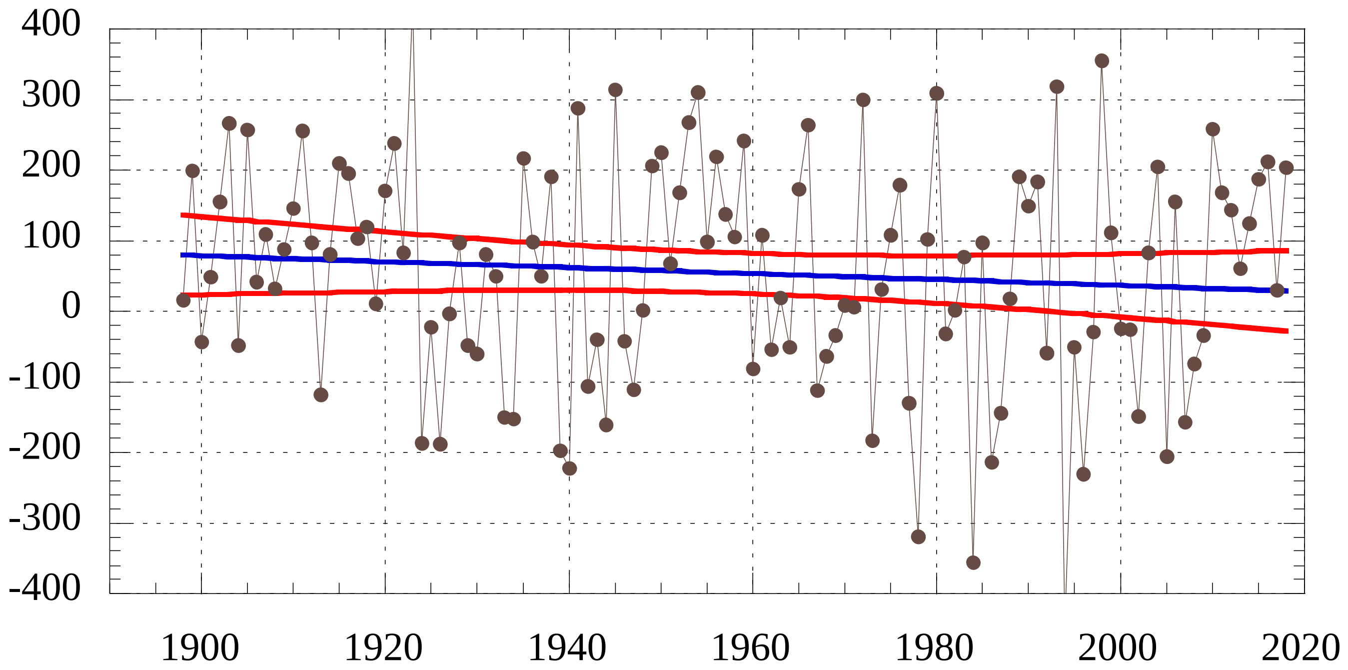
東京（大手町）



東京（大手町にある東京管区気象台エリア）は都市化率（92.9%）、気温の上昇率ともに、日本の都市の中で最大であり、ヒートアイランド現象が最も強く現れているエリア

降雨量には地球温暖化問題が見えているか？

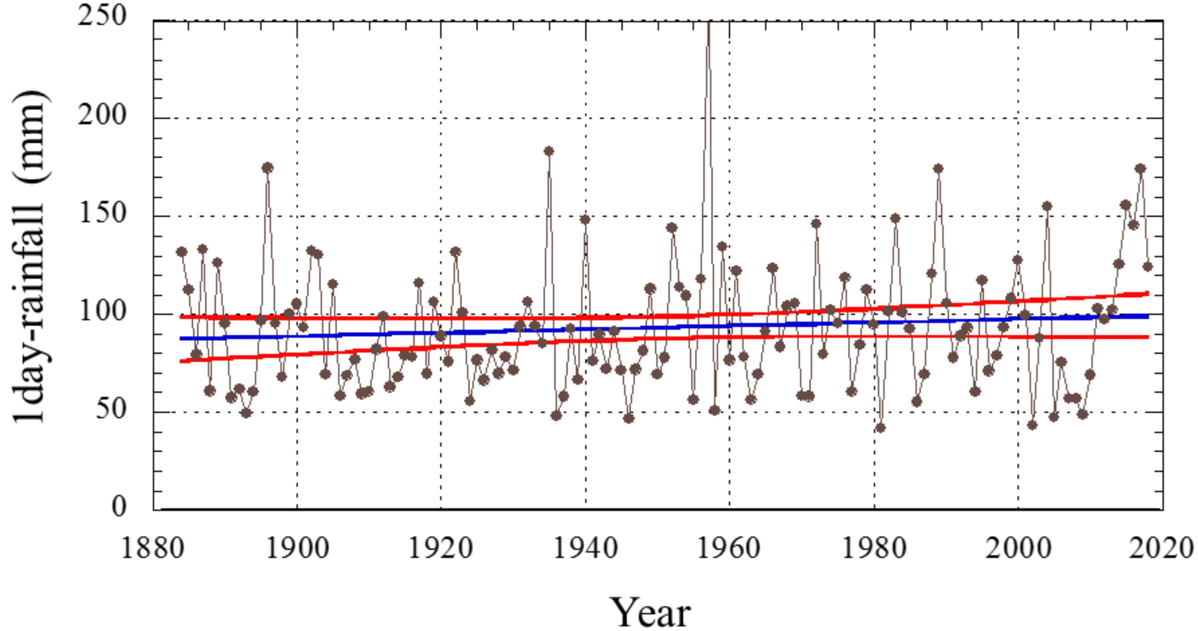
日本（51地点）の年間降水量偏差の年変化と回帰特性



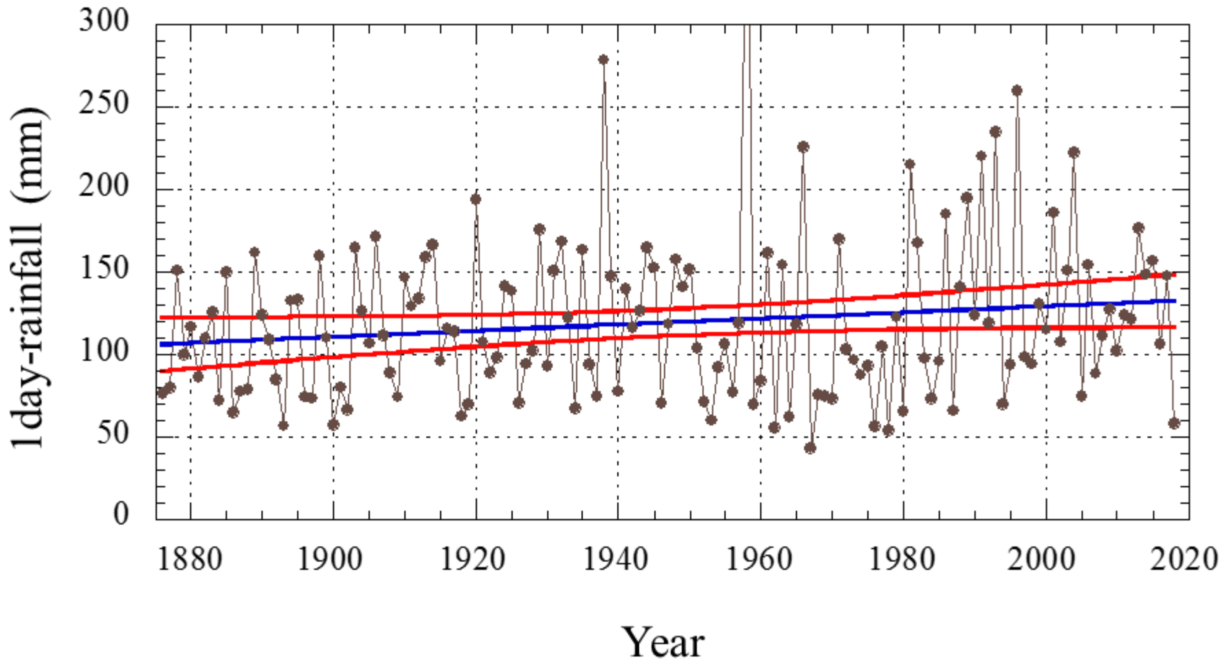
（数値データは、気象庁ホームページより）

大都市での降雨量には地球温暖化問題が見えているか？

日降雨量年間最大値

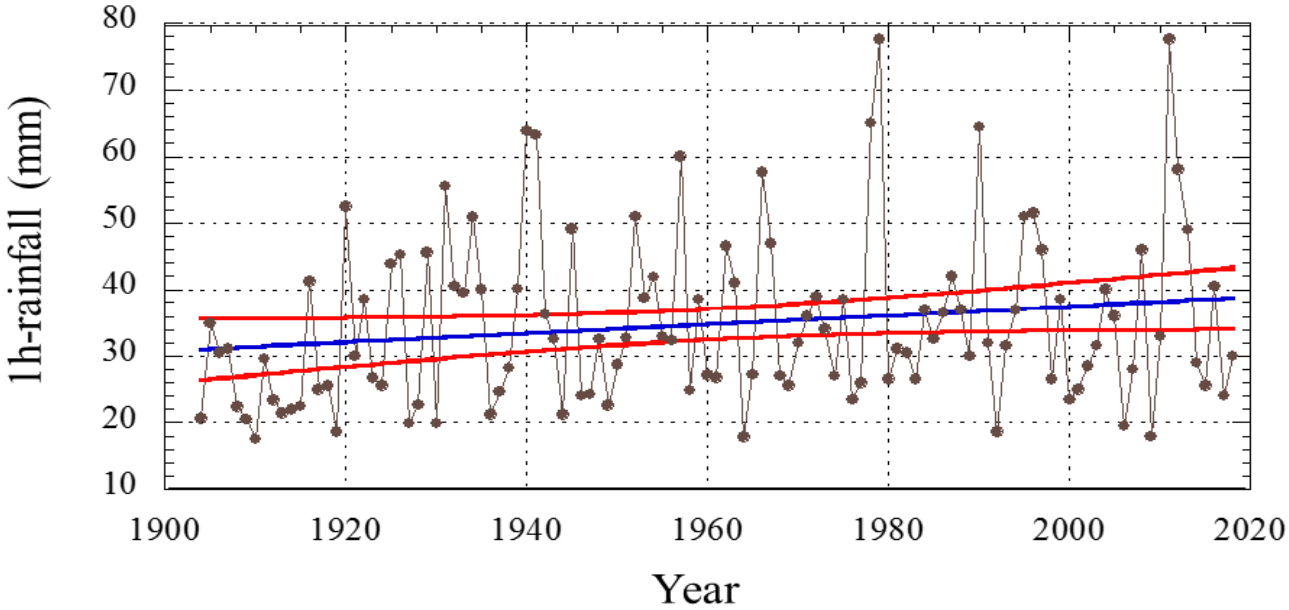


大阪



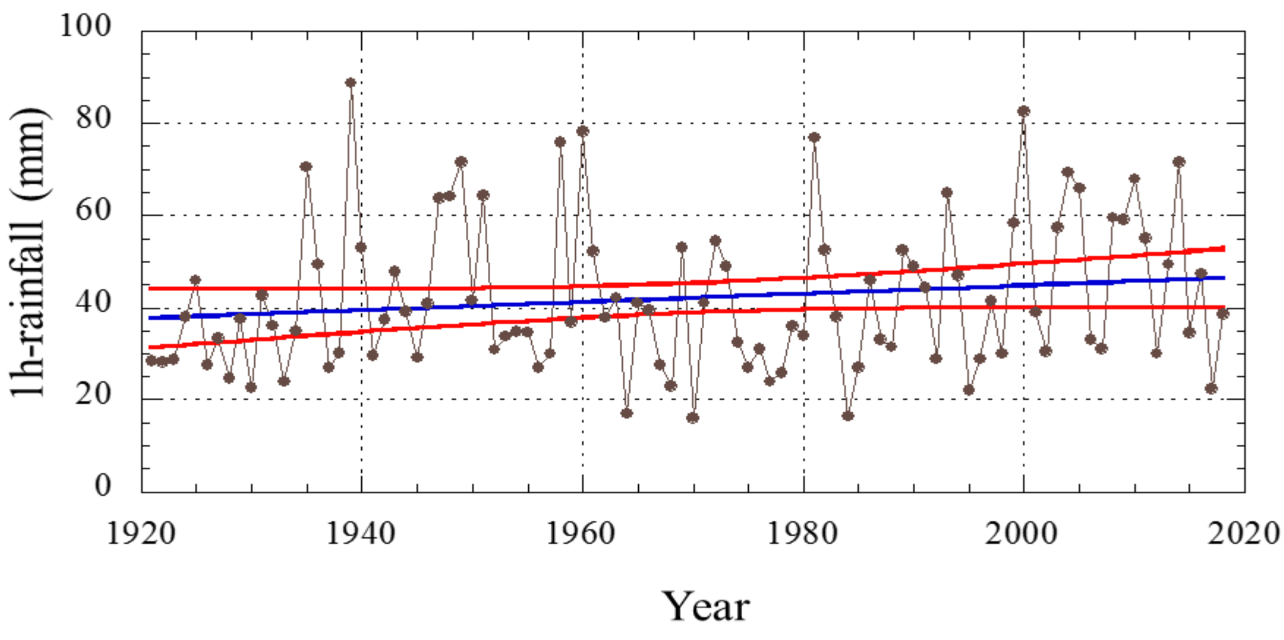
東京

大都市での降雨量には地球温暖化問題が見えているか？



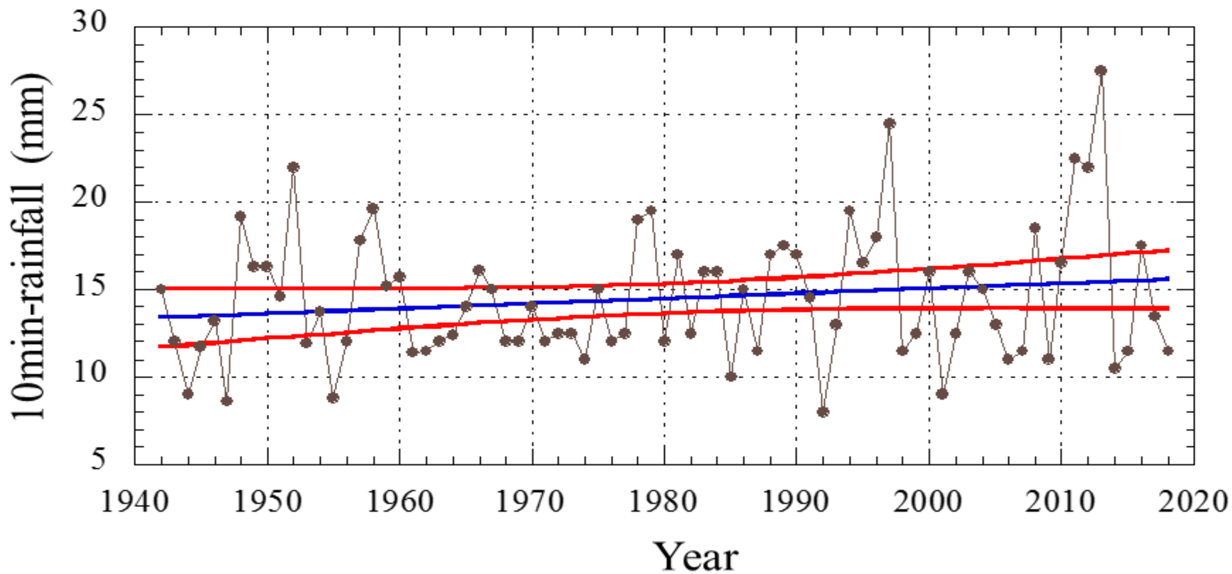
時間降雨量
年間最大値

大阪



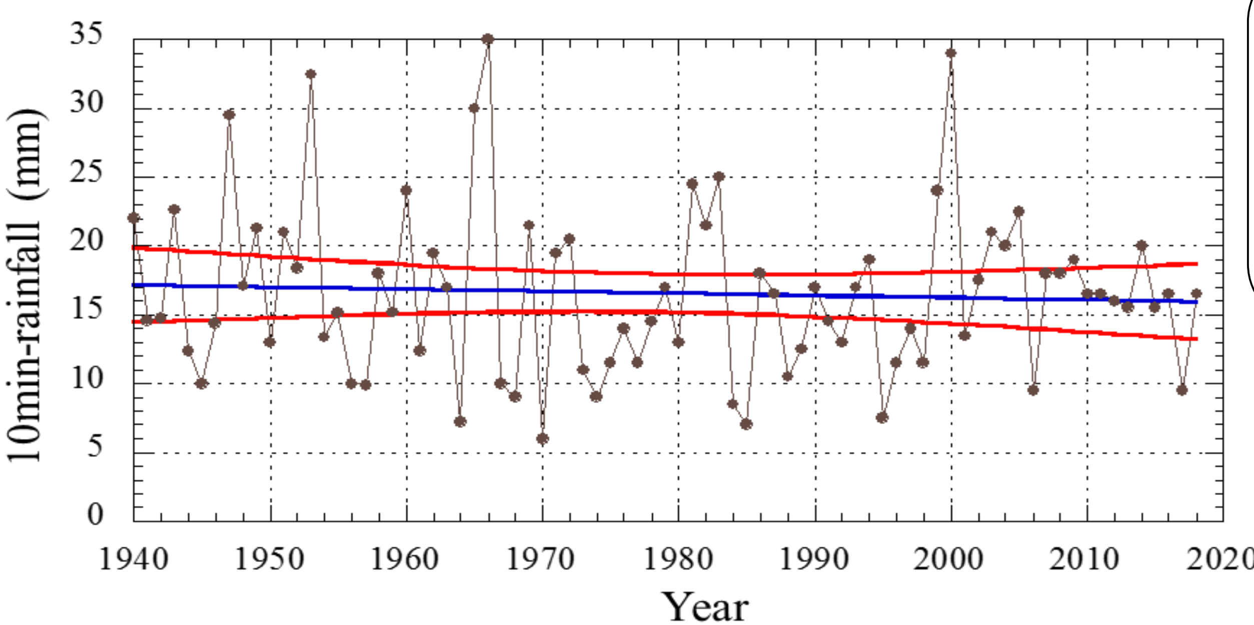
東京

大都市での降雨量には地球温暖化問題が見えているか？



10分間降雨量
年間最大値

大阪



東京

結論
区間推定の考え方からは、
地球温暖化の影響は
降雨量の長期変化には
まだ、見えていない

今後もこの定常状態が続くと仮定して
1000年に一度の最大値を推定してみる

一地点でのデータは、最大100年のオーダ

→100km以上離れたエリアでの気象の独立性を仮定して、
地点数を増やしデータ数を増やす（表に示した18地点）。
約2000点のオーダ

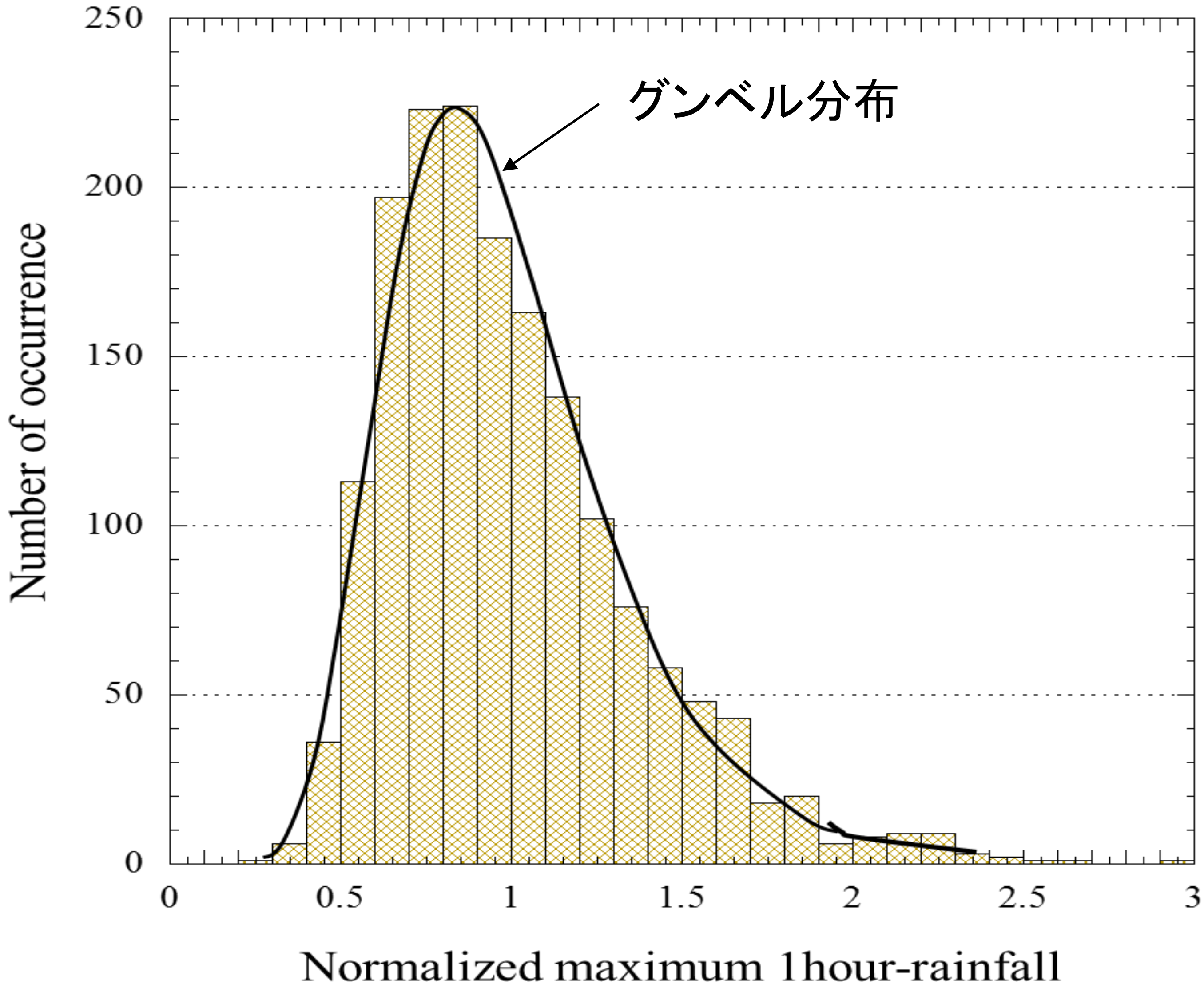
→平均値が異なる各地点データを、
地点ごと、データの平均値で割って規格化した量を用いる

*N*年間（実効的な意味）での最大値の平均値を調べる

データ数 N_{data} に対して、*N*個ずつグループ分けし、各グループの最大値の平均値を、*N*ごとに求める

（極値統計の説明で示したメンバー数*n*とこれから示す年数*N*とは意味合いが異なる。*N*=1のデータ自体が、既に多くのメンバーn_eから選ばれた最大値： $n = n_e N$ の関係）

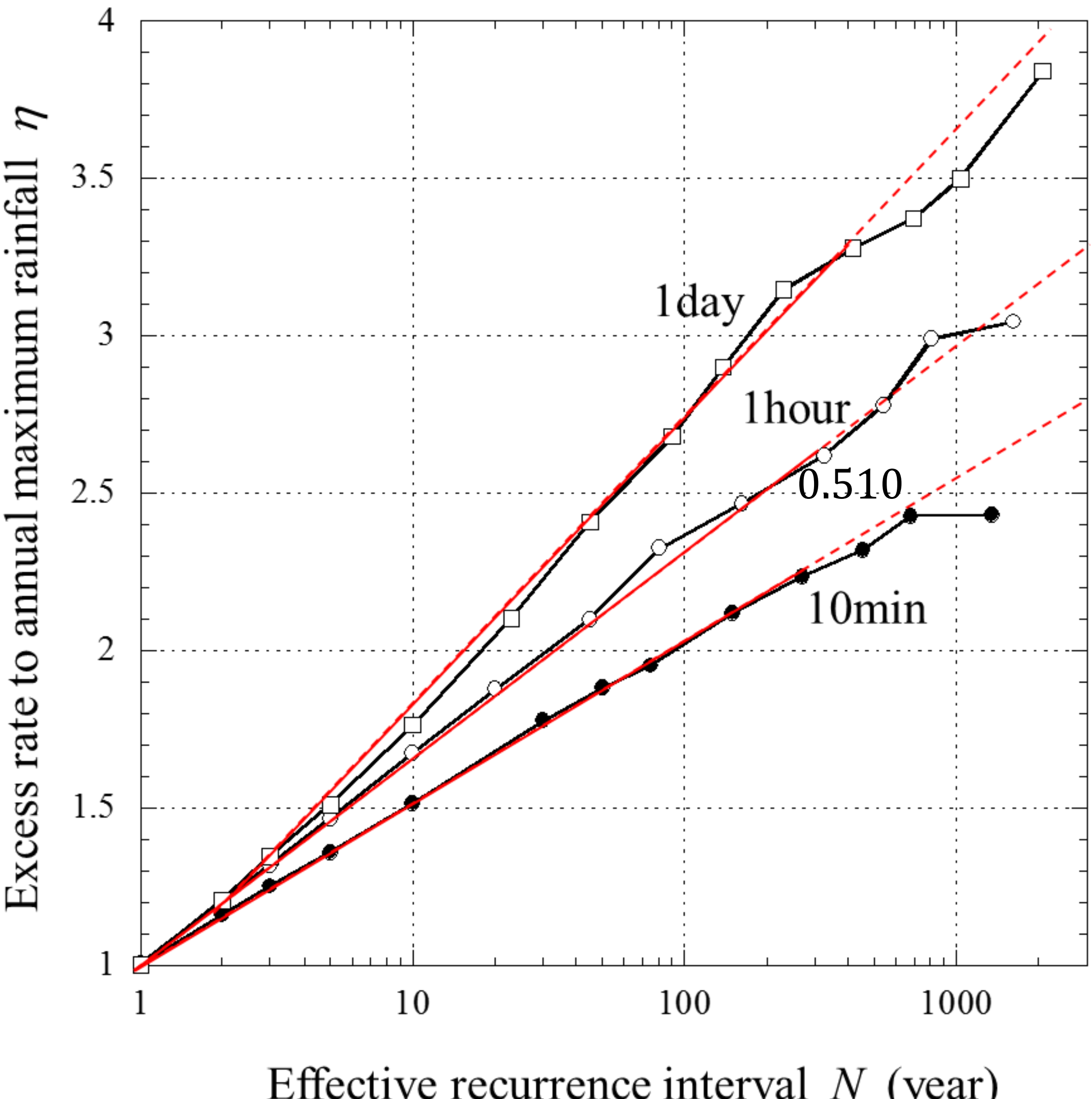
時間降雨量年間最大値（各地の平均値で正規化した）のヒストグラムと極値分布（グンベル分布）でのフィッティング



理論的に予想されるグンベル分布型極値分布と、かなり良く一致している

日降雨量も10分間降雨量も、同様によく一致している

N 年間での最悪値（N =1を基準とする増加係数）推定



どの降雨量も直線近似

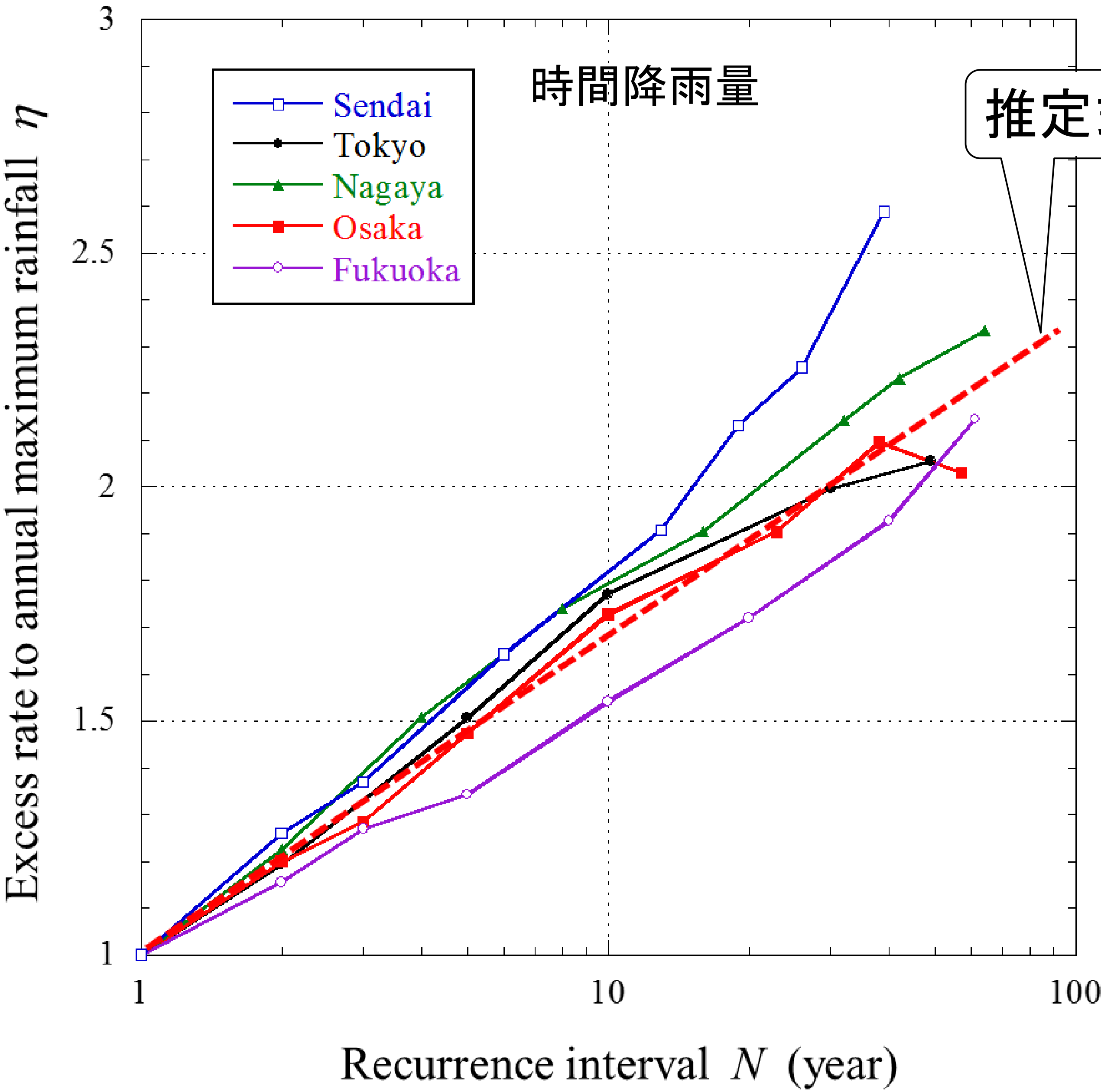
→ 降雨量はガンマ分布

$$\eta \approx \alpha \log_{10} N + 1$$

$$\alpha = \begin{cases} 0.909 & \text{（日）} \\ 0.659 & \text{（時間）} \\ 0.510 & \text{（10分間）} \end{cases}$$

【注：図の回帰直線の延長部分（点線部分）にある3点は、平均値を求めるデータ数が、3, 2, 1 と少なく、参考値】

最大値増加係数を地点ごとで見てみると(時間降雨量の例)



地域によるばらつきあり
降雨量大きい地点は
　小さめに
降雨量少ない地点は
　大きめに
→ モデル高精度化の
　課題

他の降雨量でも、
同様な傾向

【注：N=N_{data}/2までをプロット。Nが大きいところで、ηの値に反転があるのは、最大値付近のデータのグループ分けの偶然による】

推定結果の解釈

東京の時間降雨量年間最大値の平均値は、42.06mm

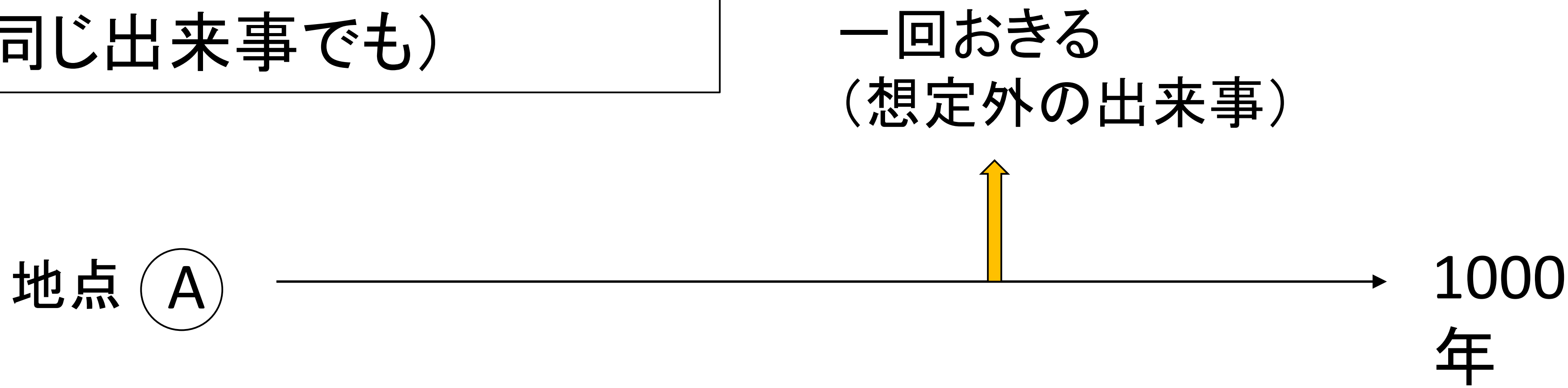
→ 10年、100年、1000年に対しては、72mm, 97mm, 126mm

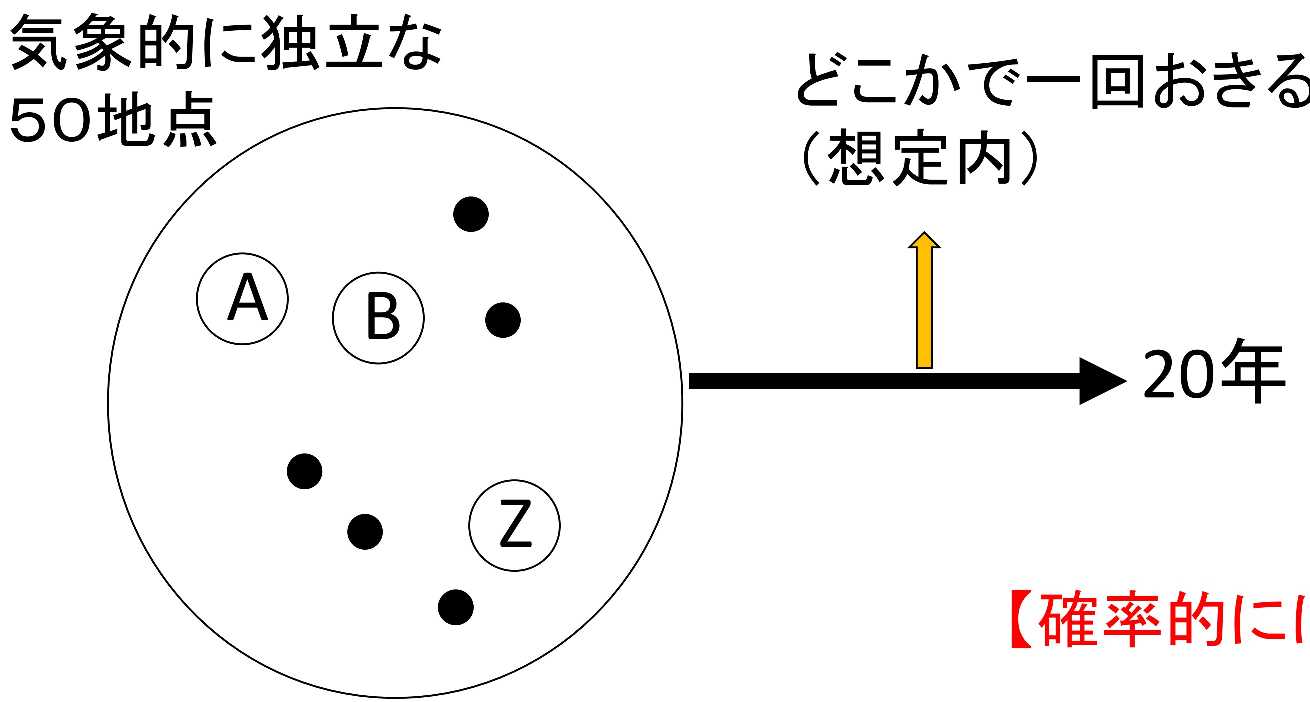
日降雨量年間最大値の平均値は、119.2mm

→ 10年、100年、1000年に対しては、226mm, 334mm, 441mm

本当だろうか？

パラレルワールドの考え方
（同じ出来事でも）





【確率的には同じこと】

まとめ（１）

気温の長期的増加傾向に見られる地球温暖化問題と
ヒートアイランド現象

降雨量の統計に対しては

・ゲリラ豪雨
・バケツをひっくり
　返したような雨
⇨ 10分間降雨量の年間最大値に現れるのでは？ ⇨ 長期傾向は見えなかった

・河川氾濫
・山崩れ
　（大災害）
⇨ 日・時間降雨量の年間最大値に現れるのでは？ ⇨ 長期傾向は見えなかった

結論

区間推定の考え方からは、地球温暖化の影響は降雨量の長期変化には、まだ見えていない

まとめ（2）

降雨の過去100年間が定常状態とみなし、18地点の極値データを地点毎に平年値で規格化し、一つの大きなデータを作った。

そのデータから*N*年間での最大値を、日・時間・10分間降雨量について推定式を定め、1000年に一度の豪雨の推定を行った。

そのモデルの推定を、パラレルワールドの考え方に結びつけ、考察を行った。

<u>今後の課題</u>

推定式には、降雨量がガンマ分布であることの性質を示しているが、極値統計学と結びつけて、モデル推定（元の分布のパラメータ推定）にまで持ってゆきたい。（ガンマ分布の二つのパラメータおよび、*n*と*N*の関係）

入手しやすいデータをいっぱい集めて、それから、ものの本質を探るビッグデータの考え方に結び付けられれば面白い