

極値降水データの「外れ値」の統計的評価 Statistical Evaluation of Outliers in Extreme Rainfall Data

○近森秀高*・工藤亮治*・三宅佑季†

CHIKAMORI Hidetaka*, KUDO Ryoji*, MIYAKE Yuki†

1. はじめに 近年、台風や前線などの影響で夏季に発生する豪雨災害では、従来の観測値を大きく上回る規模の雨量が観測される事例が少なくない。想定を上回る規模の観測値は、過去のデータに適応した確率分布からは大きくかけ離れた「外れ値」であることが多い。年最大値などの極値の標本では、発生頻度が低い「外れ値」は、極端現象の定量的評価のために貴重な情報となるはずであるが、従来の統計解析法の手順に従って「外れ値」を含むデータに確率分布を適応しても、その適合性に問題があるために、ほとんどの場合に「外れ値」を異常値と同様に扱わざるを得ない。本研究では、このような「外れ値」を持つと見られる年最大日雨量データに対して、一般化極値分布 (Generalized Extreme Distribution, GEV) を適応し、パラメータの推定法による「外れ値」への適合性の相違を調べた。また、各地点の雨量データを対象に、bootstrap法によって得られるリサンプリング標本にそれぞれ一般化極値分布を適応してパラメータと確率日雨量の分布を調べ、「外れ値」の存在が確率分布のパラメータと確率水文量の推定値の分布に及ぼす影響を調べた。

2. 解析対象資料 近年、大規模な豪雨を記録した雨量観測点を10カ所選定し、1988～2017年の40年間に観測された年最大日雨量のデータを用いて解析を行った。解析対象とした観測点の位置の概略を **図1** に示す。

3. 一般化極値分布の適応法と「外れ値」への適合性 解析対象の雨量観測点における年最大日雨量データにそれぞれ一般化極値分布を適応した。一般化極値分布の確率分布関数は以下のものである。

$$F(x) = \begin{cases} \exp \left\{ - \left[1 - \frac{k(x-c)}{a} \right]^{1/k} \right\} & (k \neq 0) \\ \exp \left\{ - \exp \left(- \frac{x-c}{a} \right) \right\} & (k = 0) \end{cases}$$



図1 解析対象雨量観測点
Objective observatories

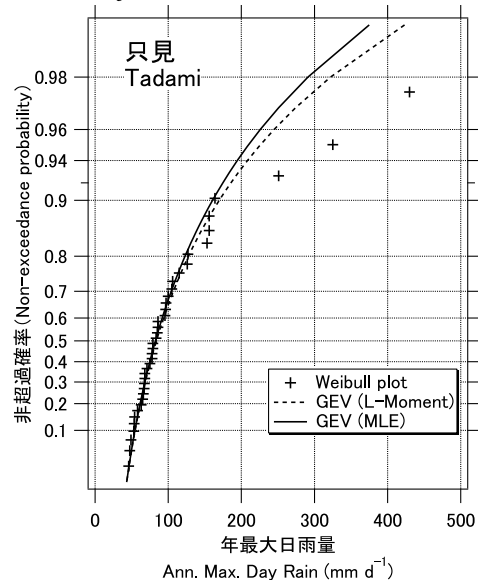


図2 年最大日雨量への一般化極値分布の適応例
An example of application of GEV to annual maximum daily rainfall

* 岡山大学大学院環境生命科学研究所, Graduate School of Environmental and Life Science, Okayama University

† 香川県庁, Kagawa Prefectural Government

キーワード: 水文統計, 降雨特性, 外れ値

ここに、 a , c , k はパラメータである。適応には、L積率 (L-Moment) 法および最尤法 (MLE) を用いた。只見地点における適応結果を Weibull プロットによる非超過確率の推定値とともに図2に例示する。ほとんどの地点で、年最大日雨量の上位値は、適応した一般化極値分布を示す実線および破線に対して右側に外れてプロットされる「外れ値」となった。ただし、川井地点では、2～4位の年最大値は分布曲線よりも右側にプロットされたが、第1位の値は分布曲線近くにプロットされた。確率分布の適応法による「外れ値」への適合性は地点によって異なった。

4. 確率分布のパラメータおよび確率日雨量の bootstrap 分布 ここでは、リサンプリング手法の一つである bootstrap 法を用い、bootstrap 標本から得られる一般化極値分布のパラメータと確率日雨量の分布について調べた。

(a) 100年確率日雨量の分布 各 bootstrap 標本からそれぞれ 100年確率日雨量を推定し、その分布を調べた結果、五十里、高知、朝倉の3地点では、L積率法、最尤法のいずれを用いた場合でも2個以上のピークを持った分布が得られた。特に、朝倉地点では、L積率法により適応した一般化極値分布から求められた100年確率日雨量の分布に3個のピークが見られた。只見および宮古島では最尤法、津和野ではL積率法を用いた場合に2個のピークが見られた。これらの地点を除く5地点では、いずれの適応法を用いた場合もピークは1個となった。図3に、朝倉地点における100年確率日雨量の bootstrap 分布を例示する。また、bootstrap 標本における第1位の年最大値の有無が確率日雨量の推定に及ぼす影響を調べた結果、図4に示すように、bootstrap 標本における第1位の有無によって、100年確率日雨量の分布が大幅に異なることが分かった。

(b) パラメータの分布 朝倉地点における一般化極値分布のパラメータ k の分布を図5に示す。この図から、 k の分布が100年確率日雨量と同様に2個のピークを持つことが分かる。100年確率日雨量と同様に、bootstrap 標本における第1位の年最大値の有無が k の分布に及ぼす影響を調べた結果、図6に示すように、標本における第1位の有無によって、 k の分布が大幅に異なることが分かった。なお、他のパラメータの分布には、 k のような明確な相違は見られず、確率日雨量の分布の相違には、主に k の分布の相違が影響しているものと思われる。

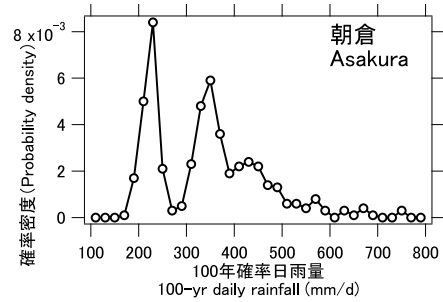


図3 100年確率日雨量のBootstrap推定値 (朝倉)
Distribution of Bootstrap estimates of 100-yr daily rainfall at Asakura

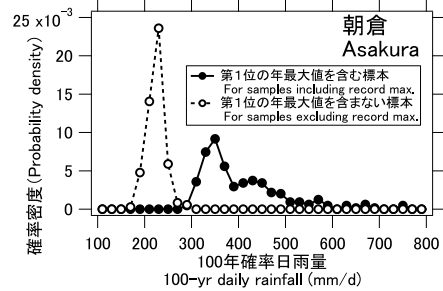


図4 標本における第1位の有無が100年確率日雨量のBootstrap推定値に与える影響 (朝倉)
Effects of existence of the largest annual maximum value in bootstrap samples on estimates of 100-yr daily rainfall at Asakura

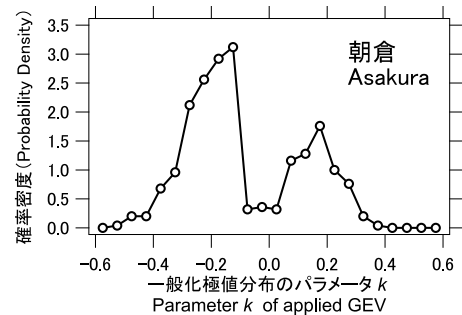


図5 一般化極値分布のパラメータ k のBootstrap推定値 (朝倉)
Distribution of bootstrap estimates of the parameter k of GEV at Asakura

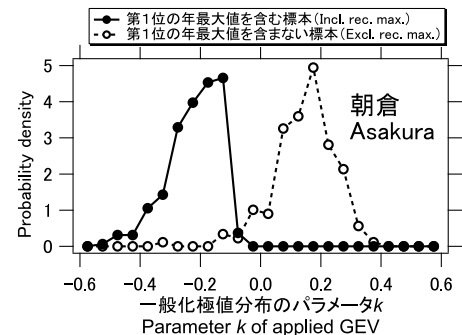


図6 標本における第1位の有無がパラメータ k のBootstrap推定値に与える影響 (朝倉)
Effects of existence of the largest annual maximum value in bootstrap samples on estimates of parameter k of GEV