

# 水文時系列の平均値と分散のジャンプ検出 および異常値の非異常値化

DETECTION OF MEAN AND VARIANCE JUMP AND NORMALIZATION OF  
OUTLIER IN HYDROLOGICAL TIME SERIES

佐渡公明<sup>1</sup>・中尾隆志<sup>2</sup>・杉山一郎<sup>3</sup>

Kimiteru SADO, Takashi NAKAO and Ichiro SUGIYAMA

<sup>1</sup>フェロー 工博 北見工業大学教授 土木開発工学科 (〒090-8507 北見市公園町165番地)

<sup>2</sup>正会員 博(工) 北見工業大学助教 土木開発工学科 (〒090-8507 北見市公園町165番地)

<sup>3</sup>正会員 博(工) 北見工業大学非常勤研究員 土木開発工学科 (〒090-8507 北見市公園町165番地)

A non-stationary time series is composed of the trend and some jumps by the change of a mean value and a variance. There is no study example which has dealt with the power of test for a jump of variance in a hydrological time series.

In this paper, the Monte Carlo simulations are applied to compare the power of the statistical tests: Student's *t* test, *F* test, Mann-Whitney test and bootstrap test to assess the significance of jumps to affect the mean value and the variance of a time series with Normal, Logarithm Normal and General Extreme Value distributions. And also, the outliers in the annual maximum daily rainfall (MDR) and annual maximum non-rainfall days (NRD) data of each 22 meteorological observatories in Hokkaido are detected at Tomakomai and Hiroo for MDR data, and at Esashi for NRD data, respectively. Then the normalized conditions to change these three outliers into normal values are introduced by using a numerical simulation.

**Key Words :** *Jump detection, power of test, Student's t test, F test, Mann-Whitney test, bootstrap test, normalization of outlier*

## 1. はじめに

最近の地球温暖化, 気候変動が示唆されている状況のもとでは, 定常性の仮定による水文頻度解析手法には無理があり, 非定常性を考慮した解析手法の開発が必要である. 水文時系列が非定常確率過程を示すときは当然非定常頻度解析をしなければならない. 非定常変動には, 平均値のジャンプ, トレンドおよび分散のジャンプ, トレンドが含まれる. 本論文では, 確率水文学の算定に関連する平均値と分散のジャンプを扱う. 水文時系列の平均値のジャンプ検定に関しては, Z.W.Kundzewicz and A. J. Robson<sup>1)</sup>が各種の方法を整理し, また徐・竹内・石平<sup>2)</sup>, J. D. Salas<sup>3)</sup>はノンパラメトリック検定の Mann-Whitney 検定を紹介している. さらに文献<sup>4)</sup>では非正規分布または非独立な水文時系列の検定には, リサンプリング法(ブーストラップ法, 並べ替え法)が有利であることが強調されている. しかし, 分散のジャンプについては2つの正規母集団の分散の差に関するF検定(例えば文献<sup>4)</sup>)を除いては

適用例がなく, 関連してW.G.Struoczewski<sup>5)</sup>が平均値と分散の同時推定を行なっているだけである.

統計的検定手法は, 第1, 2種の誤りを伴い, 変化が弱く標本数が少ない場合の検出力が問題となる. M.

Radziejewski and Z. W. Kundzewicz<sup>6)</sup>は正規分布を示す時系列に対する平均値ジャンプの検出力と標本数の関係を調べている. 著者ら<sup>7)</sup>は正規分布と一般化極値分布に従う平均値のジャンプ検出力は, 標本数が多くジャンプの大きさが大きいときに増加することを明らかにしている. しかし, これまで分散のジャンプに関する検出力を扱った研究例はない. 次に, 水文時系列に含まれる異常値と思われる標本値を解析に含めるか含めないかにより, 確率水文学は大きく変化するので, 異常値の取扱いには充分注意しなければならない. 建設省河川砂防技術基準(案)調査編<sup>8)</sup>では, 「データの棄却検定はあくまで確率計算上の取扱いであり, これにより棄却されるデータも計画策定等に際して重要な意味を持つことがあることに留意すること.」と述べているが, 重要な意味とは何なのか不明確である. 例えば, 将来異常値が異常値でなくなるこ

ともあり得るのである。

以上を背景として、本研究の目的は、平均値のジャンプに関するt検定、Mann-Whitney検定(以下MW検定と略記)、ブートストラップt検定(BS-t検定)、ブートストラップMann-Whitney検定(平均値BS-MW検定)、および分散のジャンプに関するF検定、ブートストラップF検定(BS-F検定)、平均値BS-MWにおいて標本値を絶対偏差で置き換えたブートストラップMann-Whitney検定(分散BS-MW検定)を用いて、1) 正規分布、対数正規分布、GEV分布データの平均値と分散のジャンプの検出力をモンテカルロシミュレーションにより求めること、次に2) 北海道内22気象官署の年最大日雨量、年最大無降雨連続日数データについて、異常値を検出し、これらが将来異常値でなくなる条件を明らかにすることである。

## 2. ブートストラップジャンプ検定法

### (1) 平均値のジャンプ検定

#### a) ブートストラップt検定

観測時系列 $X(=x_1, \dots, x_N)$ を、ジャンプ発生前の $n_1$ 個標本 $X_1(=x_1, \dots, x_{n_1})$ とジャンプ発生後の $n_2$ 個の標本 $X_2(=x_{n_1+1}, x_{n_1+2}, \dots, x_{n_1+n_2})$  ( $n_1+n_2=N$ ) に分ける。両標本の母集団平均値を $\mu_1, \mu_2$ とし、帰無仮説を $H_0$ 「 $\mu_1 = \mu_2$ 」とする。 $X$ の検定統計量 $t_0$ を次式で求める。

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{u_1^2}{n_1} + \frac{u_2^2}{n_2}}} \quad (1)$$

ここに、 $\bar{x}_1, \bar{x}_2$ は両標本の標本平均値、 $u_1^2, u_2^2$ は両標本の不偏分散である。ジャンプの大きさ $\delta$ は $\delta = \bar{x}_2 - \bar{x}_1$ である。ブートストラップ検定法では、帰無仮説のもとで検定統計量 $t_0$ の経験分布をブートストラップ標本を用いて近似する。従って $H_0$ を満たす母集団からの標本抽出を行うため、両標本の平均値が $X$ の標本平均値 $\bar{X}$ に等しくなるよう変換する<sup>9)</sup>。

$$\begin{aligned} y_i &= x_i - \bar{x}_1 + \bar{X}, \quad i=1, \dots, n_1 \\ y_{n_1+j} &= x_{n_1+j} - \bar{x}_2 + \bar{X}, \quad j=1, \dots, n_2 \end{aligned} \quad (2)$$

式(2)の $y_i, y_{n_1+j}$ からのブートストラップ標本をそれぞれ、 $y_i^*, i=1, \dots, n_1$ および $y_{n_1+j}^*, j=1, \dots, n_2$ とし、ブートストラップ検定統計量は次式で与えられる。

$$\hat{t}^* = \frac{\overline{y_1^* - y_2^*}}{\sqrt{\frac{u_1^{*2}}{n_1} + \frac{u_2^{*2}}{n_2}}} \quad (3)$$

ただし、 $\overline{y_1^*}, \overline{y_2^*}$ および $u_1^{*2}, u_2^{*2}$ は $y_i^*, y_{n_1+j}^*$ の標本平均値、標本不偏分散である。 $M$ 組のブートストラップ標本に対する $M$ 個の $\hat{t}^*$ を値の小さい方から順番に並びかえて、 $i$ 番目の値に対するプロットングポジション $p_i$ を次の

カナン公式から計算し、経験分布関数( $p_i \sim \hat{t}^*$ )を求める。

$$p_i = \frac{i - \alpha}{M + 1 - 2\alpha}, \quad \alpha = 0.4 \quad (4)$$

$t_0$ が $\hat{t}^*$ の $m_T \sim m_T + 1$ の範囲に入ったとき、 $t_0$ の非超過確率は $p_T = (m_T - 0.4) / (M + 0.2)$ と得られる。有意水準が $\alpha = 0.05$ の両側検定の場合、次のように判定する。

$$\left. \begin{aligned} p_i < 0.025 & \dots H_0 \text{を棄却, ハイジャンプ} \\ 0.025 < p_i < 0.975 & \dots H_0 \text{を採択, ジャンプなし} \\ 0.975 < p_i & \dots H_0 \text{を棄却, ロージャンプ} \end{aligned} \right\} (5)$$

#### b) ブートストラップMann-Whitney検定

帰無仮説は前述のブートストラップt検定と同じとする。2.(1)a)と同じように、時系列 $X(=x_1, \dots, x_N)$ をジャンプ発生前後の2つの時系列 $x_i, i=1, \dots, n_1$ と $x_{n_1+j}, j=1, \dots, n_2$ に分け、この2つの時系列を混合し昇順に並べた標本 $z_i, i=1, \dots, N$ を作成する。 $X$ の検定統計量 $T_{x_0}$ を次式の順位和で求める。

$$T_{x_0} = \sum_{i=1}^{n_1} R(x_i) \quad (6)$$

ここに、 $R(x_i)$ は $z_i$ の中の観測値 $x_i$ の順位である。式(2)の2つの時系列 $y_i, y_{n_1+j}$ からブートストラップ標本を $M$ 組発生させ、式(6)と同様のブートストラップ検定統計量を求め、経験分布関数( $p_i \sim \hat{T}_x^*$ )を作成する。 $T_{x_0}$ が $\hat{T}_x^*$ の $m_T \sim m_T + 1$ の範囲に入ったとき、 $T_{x_0}$ の非超過確率は $p_T = (m_T - 0.4) / (M + 0.2)$ と得られる。有意水準が $\alpha = 0.05$ の両側検定の場合、式(5)の $p_i$ を $p_T$ に置き代えて最終的な判定を行う。

### (2) 分散のジャンプ検定

#### a) ブートストラップF検定

2.(1)と同様に、観測時系列 $X(=x_1, \dots, x_N)$ をジャンプ発生前後の2つの時系列 $x_i, i=1, \dots, n_1$ と $x_{n_1+j}, j=1, \dots, n_2$ に分ける( $n_1+n_2=N$ )。両標本の正規母集団分散を $\sigma_1^2, \sigma_2^2$ とし、帰無仮説を $H_0$ 「 $\sigma_1^2 = \sigma_2^2$ 」とする。 $X$ の検定統計量 $F_0$ を次式で求める。

$$F_0 = \frac{u_1^2}{u_2^2} \quad (7)$$

ここに $u_1^2, u_2^2$ は両標本の不偏分散である。 $H_0$ を満たす母集団からの標本抽出を行うため、両標本 $y_i, y_{n_1+j}$ の分散 $V(y_i), V(y_{n_1+j})$ が $X$ の分散 $V(X)$ に等しくなるように次の変換を行う。

$$\left. \begin{aligned} y_i &= \left( \frac{V(X)}{V(x_i)} \right)^{\frac{1}{2}} x_i, \quad i=1, \dots, n_1 \\ y_{n_1+j} &= \left( \frac{V(X)}{V(x_{n_1+j})} \right)^{\frac{1}{2}} x_{n_1+j}, \quad j=1, \dots, n_2 \end{aligned} \right\} (8)$$

式(8)からのブートストラップ標本を $y_i^*, y_{n_1+j}^*$ とし、ブートストラップ検定統計量を $\hat{F}^* = u_1^{*2} / u_2^{*2}$  ( $u_1^{*2}, u_2^{*2}$ は $y_i^*, y_{n_1+j}^*$ の標本不偏分散)とし、 $F_0$ の非超過確率 $p_F$ は、2.(1)と同様にカナン公式から得られ、 $\alpha = 0.05$ の両側検定

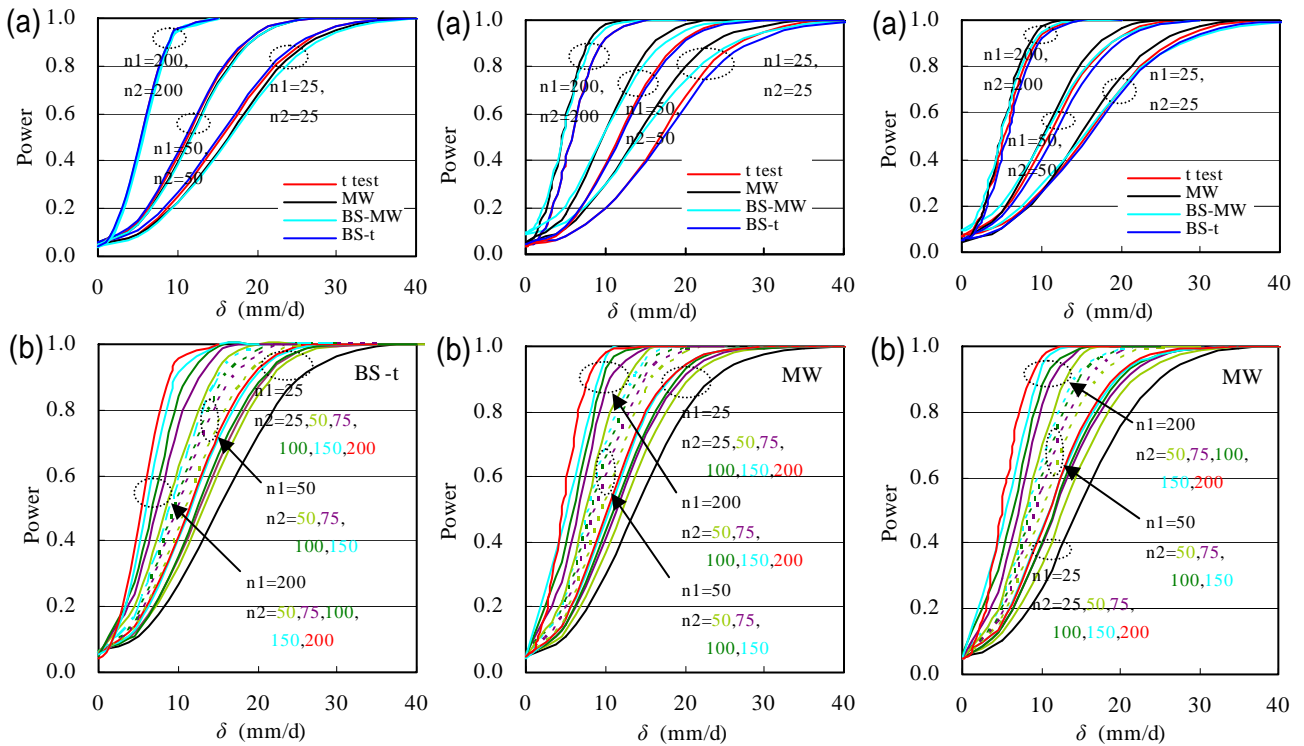


図-1 (a)  $N(73, 28^2)$  に  $N(73+, \sigma, 28^2)$  が  
 続く正規分布時系列に対する  
 平均値ジャンプの検出力  
 (b) (a) で最も検出力の高い  
 BS-t 検定の検出力

図-2 (a)  $\mu=73+$  ,  $\sigma=28$  ,  $\gamma=1.4$  の3母数  
 対数正規分布時系列に対する  
 平均値ジャンプの検出力  
 (b) (a) で最も検出力の高い  
 MW 検定の検出力

図-3 (a)  $\mu=73+$  ,  $\sigma=28$  ,  $\gamma=1.4$  のGEV  
 分布時系列に対する平均値  
 ジャンプの検出力  
 (b) (a) で最も検出力の高い  
 MW 検定の検出力

の場合、式(5)の  $p_i$  を  $p_{i+1}$  に置き換えて最終判定を行う。

### b) ブートストラップMann-Whitney検定

2. (1)b) の平均値ジャンプに対するBS-MW検定において、平均値ジャンプ発生前の時系列  $x_i$  を絶対偏差  $|x_i - \bar{x}_1|$  と置き換え、発生後時系列  $x_{n1+j}$  を絶対偏差  $|x_{n1+j} - \bar{x}_2|$  と置き換える。その後は、2. (1)b) と同じ検定方法を行うことにより、絶対偏差のジャンプ検定を行うことができる。

## 3. ジャンプの検出力

### (1) 平均値のジャンプ検出力

平均値ジャンプ発生前後の2つの時系列を  $x_i$  と  $x_{n1+j}$  として、 $N(73, 28^2)$  に従う正規乱数を発生させる。その後、 $x_{n1+j}$  にジャンプの大きさ  $\delta$  (0, 5, 10, ..., 100mm/d) を加えることにより、帰無仮説  $H_0$  「 $\mu_1 = \mu_2$ 」 が偽りとなる模擬時系列  $x_i, x_{n1+j}$  を3000回発生させる。t検定、MW検定、BS-t検定、BS-MW検定を行い、検出力 (=  $N_{rej} / 3000$ ,  $N_{rej}$ : 棄却回数) を比較する。ブートストラップ標本は  $M = 3000$  組発生させる。図-1(a) は  $\alpha = 0.05$  の両側検定の結果である。ジャンプ前後の標本の大きさ  $n1, n2$  と  $\delta$  が大きいほど検出力が増加し、ジャンプを検出しやすことが分かる。4者の比較では、BS-t検定、t検定が検出力が大きい、 $n1, n2$  の増加と共にその差はなくなる。これは、

パラメトリック検定のt検定、BS-t検定を利用できる状況の中で、あえてMW検定、BS-MW検定のノンパラメトリック検定を使用しても検出力の面で不利であることを表わしている。図-1(b) は最も検出力の大きいBS-t検定について、 $n1, n2$  の影響を表している。図-2(a), 3(a) は図-1(a) の正規分布を平均値  $\mu=73$ mm/d、標準偏差  $\sigma=28$ mm/d、ひずみ係数  $\gamma=1.4$  の3母数対数正規分布、GEV分布に変えた場合の検出力の比較を示している。図-1(a) と逆にノンパラメトリック検定のMW検定、BS-MW検定がt検定、BS-t検定より検出力が大きい、この場合も  $n1, n2$  が増加すると差は小さくなる。図-2(b), 3(b) は最も検出力の大きいMW検定について、 $n1, n2$  の影響を表している。ここで用いた平均値73mm/d、標準偏差28mm/d、ひずみ係数1.4は、北海道内22気象官署の年最大日雨量データについて、各官署ごとに開所年~2006まで求めた上記3つの統計量に対する22官署の平均値である。図-1~3の縦軸の切片0.05は有意水準を表わす。図-1~3はまた、ある平均値ジャンプをある確率(=縦軸の検出力)で検出するのに必要な標本の大きさ  $n1, n2$  を表していると見ることもできる。

### (2) 分散のジャンプ検出力

乗算合同法による区間[0,1]の一樣乱数<sup>10)</sup>を用いて、分散ジャンプ発生前後の2つの模擬時系列  $x_i, x_{n1+j}$  として、

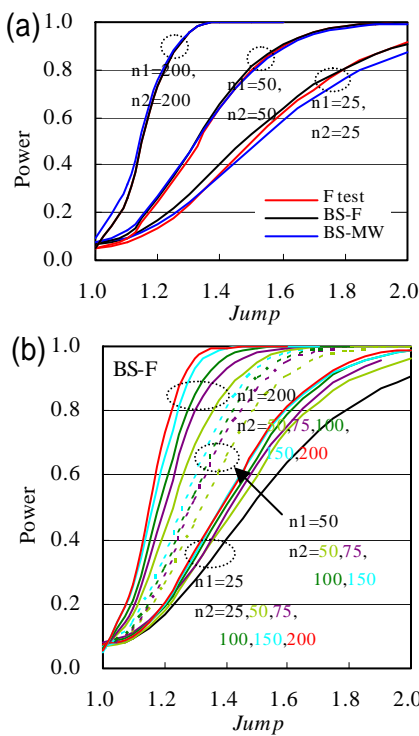


図-4 (a)  $N(73, 28^2)$  に  $N(73, (28 \times \text{Jump})^2)$  が続く正規分布時系列に対する分散ジャンプの検出力  
(b) (a) で最も検出力の高い BS-F 検定の検出力

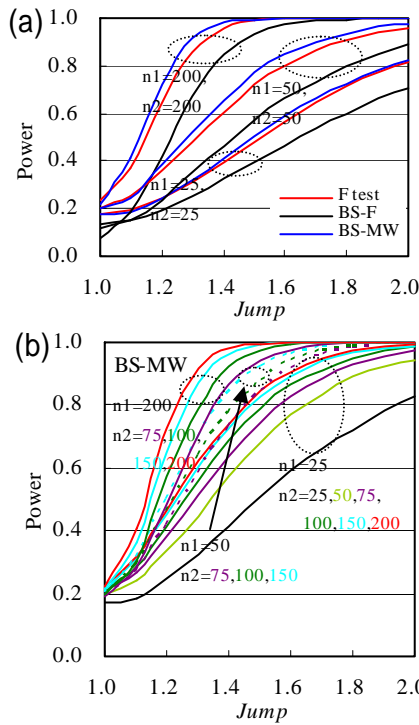


図-5 (a)  $\mu=73, \sigma^2=28 \times \text{Jump}, \sigma=1.4$  の3母数対数正規分布時系列に対する分散ジャンプの検出力  
(b) (a) で最も検出力の高い BS-MW 検定の検出力

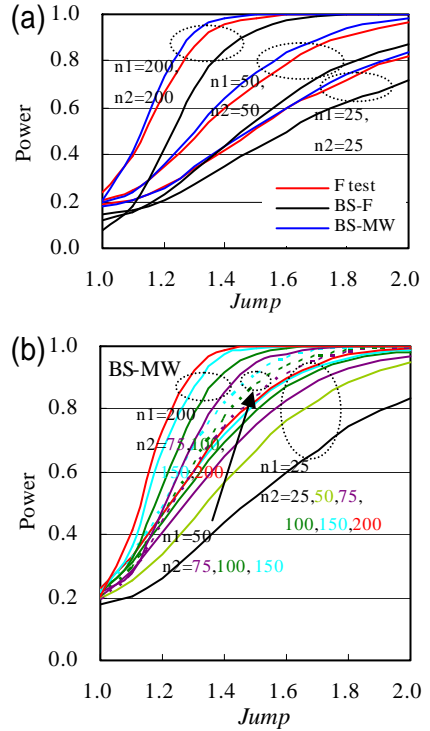


図-6 (a)  $\mu=73, \sigma=28 \times \text{Jump}, \sigma=1.4$  の GEV 分布時系列に対する分散ジャンプの検出力  
(b) (a) で最も検出力の高い BS-MW 検定の検出力

それぞれ  $N(73, 28^2)$ ,  $N(73, (28 \times \text{Jump})^2)$  を 3,000 回発生させる。すなわちジャンプの大きさ  $\text{Jump}(1.0, 1.1, \dots, 2.0)$  は 2 つの時系列における標準偏差の比を表す。  $\text{Jump}=1.0$  を除いて帰無仮説  $H_0$  「 $\sigma_1^2 = \sigma_2^2$ 」が偽りである。ブートストラップ標本は  $M=3000$  組、 $\alpha=0.05$  とした。図-4(a), (b) に結果を示す。標本の大きさ  $n_1, n_2$  と  $\text{Jump}$  が大きいほど検出力が増加している。3 者の比較では、BS-F 検定、F 検定、BS-MW 検定の順に検出力が大きくなり、図-1(a) と同様に正規分布時系列ではパラメトリック検定がノンパラメトリック検定よりジャンプを検出しやすいことが分かる。図-5(a), 6(a) は図-4(a) の正規分布を平均値 73mm/d、ひずみ係数 1.4、標準偏差がジャンプ発生前後で 28mm/d、 $(28 \times \text{Jump})$ mm/d に変えた場合の検出力の比較を示している。対数正規分布、GEV 分布の分散ジャンプにおいても、平均値ジャンプと同様に、ノンパラメトリック検定の BS-MW 検定がノンパラメトリック検定の F 検定、BS-F 検定より検出力が大きい。

#### 4. 異常値が異常値でなくなるための条件

##### (1) 3母数対数正規分布による異常値棄却検定

気象庁では、「ある場所（地域）で過去30年に一回程度発生する現象」を異常気象として定義している。こ

では、ある特定の分布に従う標本の中に、本来極めて稀にしか起こらないはずの大きな（小さな）出現値を異常値、その生起確率を異常率とする。以下に角屋の異常値棄却検定法<sup>11)</sup>を要約する。いま  $Y$  を正規分布に従う確率変数とし、 $M$  個の標本値、 $y_1, y_2, \dots, y_N$  とは無関係な標本  $y_\varepsilon$  を上述した同一母集団より得られた 1 個の標本と仮定する（帰無仮説： $H_0$  「 $y$  は異常値でない」）。このとき、次式に示す統計量  $F$  は、自由度 1、 $N-1$  の  $F$  分布に従う。

$$F = \left( \frac{N-1}{N+1} \right) \frac{(y_\varepsilon - \bar{y})^2}{s_y^2} \quad (9)$$

ここで、 $\bar{y}$ ：標本平均値、 $s_y^2$ ：標本分散である。 $F$  の値として片側異常率  $\varepsilon$  を指定したときの  $F_{N-1}^1(2\varepsilon)$  分布を採用すれば、 $y_\varepsilon$  の棄却限界が次式で計算される。

$$\left. \begin{matrix} y_{\varepsilon u} \\ y_{\varepsilon l} \end{matrix} \right\} = \bar{y} \pm s_y \sqrt{\frac{N+1}{N-1} F_{N-1}^1(2\varepsilon)} \quad (10)$$

$y_\varepsilon$  がこの限界値に対して、 $y_\varepsilon > y_{\varepsilon u}$ 、 $y_\varepsilon < y_{\varepsilon l}$  の範囲であれば、 $y_\varepsilon$  が異常値として判定される。上式を用いて検定する際、危険率  $\beta_0$  に対する棄却限界  $\varepsilon_0$  の値は次式で求められる。

$$\varepsilon_0 = 1 - (1 - \beta_0)^{1/(N+1)} \quad (11)$$

式 (10) から  $\varepsilon$  を求め、これと  $\varepsilon_0$  から、次の条件によ

り異常値の判定を行なう。

$$\left. \begin{array}{l} \varepsilon \leq \varepsilon_0 \text{ のとき} \cdot H_0 \text{ は棄却, } y_e \text{ は異常値である} \\ \varepsilon > \varepsilon_0 \text{ のとき} \cdot H_0 \text{ は採択, } y_e \text{ は異常値でない} \end{array} \right\} (12)$$

降水量原系列 $x$ が下限値 $a$ を持つ対数正規分布に従う場合、 $y = \ln(x - a)$ で表わされ、 $y$ は平均値 $\mu_y$ 、標準偏差 $\sigma_y$ の正規分布 $N(\mu_y, \sigma_y^2)$ に従う。本研究では、 $a, \mu_y, \sigma_y$ の3母数をクォンタイル-最尤法および積率法の2種類の手法<sup>12)</sup>により推定する。異常値か否かを検定しようとする標本値を $x_e$ とし式(10)において $y_{eu} = \ln(x_e - a), \bar{y} = \mu_y, s_y = \sigma_y$ として $\varepsilon$ を求める。次に $\beta_0$ の値と標本の大きさを式(11)に代入し $\varepsilon_0$ を求め、式(12)より $x_e$ が異常値か否かを判定する。

## (2) 北海道内22気象官署への適用

ある特定の母集団確率分布に従う観測時系列 $x_j, j=1, \dots, N$ の中に異常値と判定された標本値 $x^*$ が含まれているとする。この場合、将来 $x^*$ が異常値から当該母集団分布に復帰するためには、将来の観測値 $x_{N+i} (i=1)$ にどのような条件が必要になるかを考える。

$i=1$  のとき、将来の標本値 $x_{N+1}$ に対し、下記条件を満足する $x_{a1}, x_{b1}$ を求める。

$x_{N+1} < x_{a1} \dots \dots x^*$ は異常値のまま  
 $x_{a1} \quad x_{N+1} < x_{b1} \dots \dots x^*$ は異常値でなくなる、  
 $x_{N+1}$ は同じ母集団分布に従う  
 $x_{b1} \quad x_{N+1} \dots \dots x^*$ は異常値でなくなる、  
 $x_{N+1}$ は新たに異常値となる  
 $i=2$  のとき、 $N+1$  で異常値のまま(  $x_{N+1} < x_{a1}$  ),  
 $x_{N+1} = x_{a1} - \delta x$ の降雨が発生したとして下記条件を満足する  
 $x_{a2}, x_{b2}$ を求める。

$x_{N+2} < x_{a2} \dots \dots x^*$ は異常値のまま  
 $x_{a2} \quad x_{N+2} < x_{b2} \dots \dots x^*$ は異常値でなくなる、  
 $x_{N+2}$ は同じ母集団分布に従う  
 $x_{b2} \quad x_{N+2} \dots \dots x^*$ は異常値でなくなる、  
 $x_{N+2}$ は新たに異常値となる

$i=3$  のとき、 $N+1, N+2$  で異常値のまま(  $x_{N+1} < x_{a1}, x_{N+2} < x_{a2}$  ),  $x_{N+1} = x_{a1} - \delta x, x_{N+2} = x_{a2} - \delta x$ の降雨発生条件のもとで、下記条件を満足する $x_{a3}, x_{b3}$ を求める。

$x_{N+3} < x_{a3} \dots \dots x^*$ は異常値のまま  
 $x_{a3} \quad x_{N+3} < x_{b3} \dots \dots x^*$ は異常値でなくなる、  
 $x_{N+3}$ は同じ母集団分布に従う  
 $x_{b3} \quad x_{N+3} \dots \dots x^*$ は異常値でなくなる、

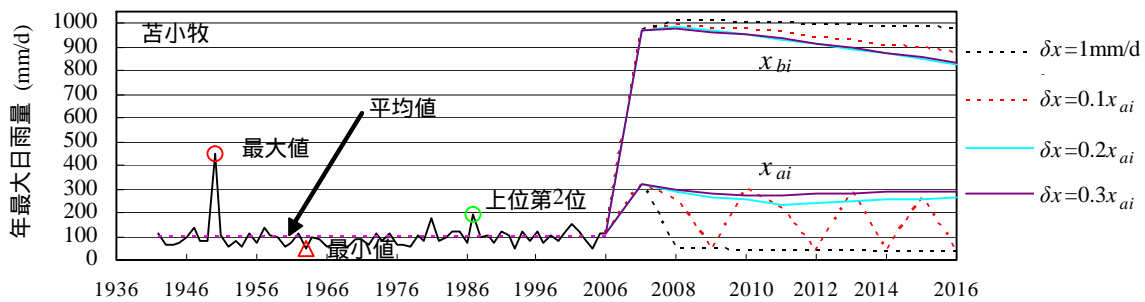


図-7 苫小牧の2006年までの年最大日雨量に関する将来異常値でなくなる $x_{ai}$ と新たな異常値下限 $x_{bi}$

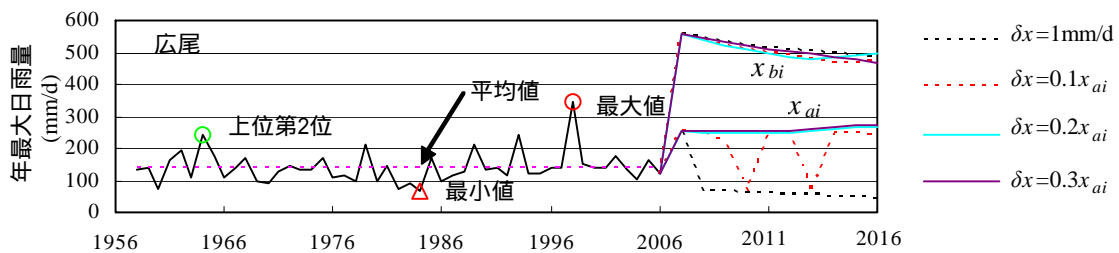


図-8 広尾の2006年までの年最大日雨量に関する将来異常値でなくなる $x_{ai}$ と新たな異常値下限 $x_{bi}$

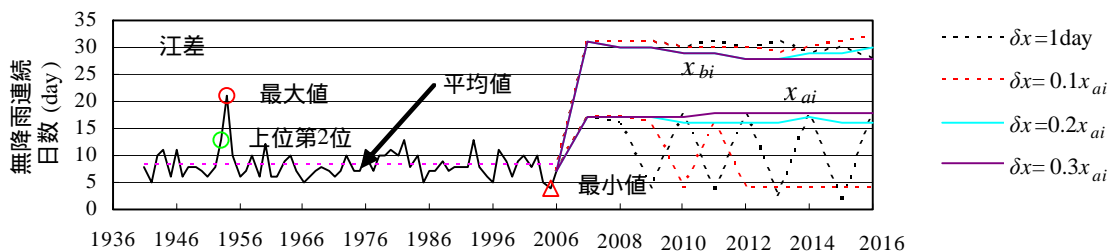


図-9 江差の2006年までの年最大無降雨連続日数に関する将来異常値でなくなる $x_{ai}$ と新たな異常値下限 $x_{bi}$

$x_{N+3}$ は新たな異常値になる

以下同様に、 $N+i$ のとき、それまでに $x_{N+1}=x_{ai}-x, \dots, x_{N+i-1}=x_{ai-1}-\delta x$ の降雨が発生したという条件のもとで、 $x^*$ が始めて異常値が異常値でなくなるのに必要な標本の最低値 $x_{ai}$ 、および新たな異常値下限となる標本最低値 $x_{bi}$ を定義する。

北海道内22気象官署の開所年～2006までの年最大日雨量の第1位データについて、 $\beta_0 = 0.05$ として3母数対数正規分布による異常値棄却検定を行なった結果、クオインタイル-最尤法と積率法の両方で異常値と判定されたのは苫小牧と広尾であった。また、全く同様に年最大無降雨連続日数については、江差の第1位が異常値と判定された。これら3つの異常値を図-7～9に示す。図-7, 8は、苫小牧(2006年で $N=65$ )、広尾(2006年で $N=49$ )の異常値に対する将来2007年から10年間にわたる異常値が異常値でなくなる計算結果である。2007年( $i=1$ )のとき異常値でなくなるのに必要な年最大日雨量は、異常値と第2位の間にある。2008～2016年( $2 \leq i \leq 10$ )のとき $\delta x$ が増加し、それまでの年最大日雨量が少ないほど、その後の異常値解消には大きな年最大日雨量 $x_{ai}$ が必要であり、また新たな異常値下限となる $x_{bi}$ は減少傾向にあることが分かる。新たな異常値下限は当然2006年までの異常値より大きいことが示されている。図-9は、江差(2006年で $N=66$ )の年最大無降雨連続日数に対する異常値解消の計算結果である。図-7, 8と同様のことが示されている。図-7～9の $x_{ai}$ は振動している。これは年最大日雨量の $x_{N+i}$ が大きいたと、翌年の $x_{ai+1}$ は減少し、逆に $x_{N+i}$ がある閾値以下に小さくなると $x_{ai+1}$ は増加するからである。

## 5. 結論

本研究により得られた結果を以下にまとめる。

- (1) 平均値ジャンプの検出力を比較すると、正規分布時系列の場合パラメトリック検定のBS-t検定、t検定がノンパラメトリック検定のMW検定、BS-MW検定より検出力が高く、対数正規分布、GEV分布時系列の場合はこの逆になる。
- (2) 分散のジャンプ検出法として、ブートストラップ法によるBS-F検定、平均値BS-MWにおいて標本値を絶対偏差で置き換えた分散BS-MW検定を提案した。
- (3) 分散ジャンプの検出力を比較すると、正規分布時系列の場合パラメトリック検定のBS-F検定、F検定がノンパラメトリック検定のBS-MW検定より検出力が高く、対数正規分布、GEV分布時系列の場合はこの逆になる。
- (4) 北海道22気象官署における開所年～2006年までの

年最大日雨量データに関する苫小牧と広尾の異常値、および年最大無降雨連続日数データに関する江差の異常値について、2007～2016年間にわたる異常値が異常値でなくなる条件を明らかにした。

- (5) (4)に関し、2007年に異常値が異常値でなくなるための最小値は、それまでの異常値と第2位の間にあり、2008～2016年の標本実現値が小さいほどその後の異常値解消には大きな標本値が必要である。

謝辞：本研究は財団法人北海道河川防災研究センターの平成16-18年度の研究助成を受けたものです。ここに感謝の意を表します。

## 参考文献

- 1) Kundzewicz, Z. W. and Robson A. J. : Change detection in hydrological records – a review of the methodology, *Hydrol. Sci. J.*, Vol. 49, No. 1, pp. 7-19, 2004.
- 2) 徐 宗学, 竹内邦良, 石平博 : 日本の平均気温・降水量時系列におけるジャンプおよびトレンドに関する研究, 水工学論文集, 第46巻, pp.121-126, 2002.
- 3) Salas, J.D. : Handbook of Hydrology, edited by D.R. Maidment, *Ch. 19 Analysis and Modeling of Hydrologic Time Series*, pp.19.1-19.20, McGraw-Hill, 1992.
- 4) 田代嘉宏, 脇本和昌, 大崎絏一 : 応用数学要論シリーズ2 確率と統計要論, 森北出版株式会社, 1997.10.
- 5) Strupczewski, W.G. : Simultaneous estimation of trends in mean and variance, World Climate Programme Data and Monitoring, WCDMP-45 Chap.13, WMO/TD-No.1013, World Meteorological Organization, Switzerland, 2004.
- 6) Radziejewski, M. and Kundzewicz, Z. W. : Detectability of changes in hydrological records, *Hydrol. Sci. J.*, Vol. 49, No. 1, pp.39-51, 2004.
- 7) 中尾隆佐, 佐渡公明, 杉山一郎 : 年最大日雨量時系列の非定常性を示すトレンドとジャンプの検出について, 水工学論文集, 第51巻, pp.301-306, 2007.
- 8) 建設省河川局監修 : 建設省河川砂防技術基準(案)調査編, pp.56-57, 山海堂, 1976.
- 9) 汪 金芳, 田栗正章, 手塚 集, 樺島祥介, 上田修功 : 計算統計 -確率計算の新しい手法-, 統計科学のフロンティア11, 岩波書店, pp. 1-64, 2006.
- 10) 黒瀬能幸, 松島勇雄, 松尾俊彦 : Fortran90のためのサブルーチンライブラリ 数値計算, 統計計算, 3次元CGライブラリ, 森北出版株式会社, 1998.
- 11) 角屋 睦 : 異常(確率)水文学とデータの棄却検定, 農業土木研究別冊第3号, pp.23-27, 1961.
- 12) 星 清 : 洪水ピークの確率評価法について, 開発土木研究所月報 別冊, 北海道開発局土木研究所, 1998.

(2007.9.30受付)