

北海道内の降水量に関する異常値特性

Characteristics of singular rainfall amount in Hokkaido

北見工業大学大学院 ○学生員 杉山一郎 (Ichiro Sugiyama)
 北見工業大学工学部 フェロー 佐渡公明 (Kimiteru Sado)
 北見工業大学工学部 正会員 中尾隆志 (Takashi Nakao)

1. はじめに

日本に接近する台風による災害、特にそれがもたらす湿った風が秋雨前線を刺激し、大雨となる場合が近年増加している。さらに、土砂災害及び堤防の決壊、他方では水不足といった気象災害が全国各地で甚大な影響を与えている。こういった気象災害に対する防災としては、各地域における近年の降雨の傾向を知ることが基本と言えよう。気象水文量に関する記録的更新が続出している昨今では、過去においては異常値とされていた値が連続して観測されれば、もはや異常値とは言えない。

そこで本研究の目的は、北海道内 22 気象官署の統計開始年から 2005 年までの年降水量、年最大日雨量及び年最大無降雨連続日数の異常値特性を調べ、過去においては異常値と判定される値が、今後観測される降水量によりどの様に変化するかを調べることにある。

2. 母集団が標準正規分布の場合

2.1 異常値棄却検定法^{1),2)}

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。ここで、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)値 y_ε を異常値、その生起確率 ε を異常率とする。いま Y を正規分布に従う確率変数とし、 N 個の標本値、 $y_1, y_2, \dots, y_N$ とは無関係な標本 y_ε 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)。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。

$$F = \frac{N-1}{N+1} \frac{(y_\varepsilon - \bar{y})^2}{S_y^2} \quad \dots (1)$$

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 ε を指定したときの $F_{N-1}^1(2\varepsilon)$ 分布を採用すれば、 y_ε の棄却限界が次式で計算される。

$$\left. \begin{matrix} y_{eu} \\ y_{el} \end{matrix} \right\} = \bar{y} \pm S_y \sqrt{\frac{N+1}{N-1} F_{N-1}^1(2\varepsilon)} \quad \dots (2)$$

y_ε がこの限界値に対して、 $y_\varepsilon \geq y_{eu}$, $y_{el} \geq y_\varepsilon$ の範囲であれば、 y_ε が異常値として判定される。上式を用いて検定する際、有意水準 β_0 に対する棄却限界 ε_0 の値は次式で求められる。

$$\varepsilon_0 = 1 - (1 - \beta_0)^{1/(N+1)} \quad \dots (3)$$

式 (2) から ε を求め、これと ε_0 から、次の条件により異常値の判定を行なう。

$$\begin{cases} \varepsilon \leq \varepsilon_0 \text{ のとき} \cdots H_0 \text{ は棄却され、} y_\varepsilon \text{ は異常値である。} \\ \varepsilon > \varepsilon_0 \text{ のとき} \cdots H_0 \text{ は採択され、} y_\varepsilon \text{ は異常値でない。} \end{cases}$$

2.2 標準正規分布に従う標本の作成と異常値特性

(1) 乱数の検定

本研究では、 $N(0,1)$ に従う模擬データの作成には、乗算合同法による一様乱数と Box-Muller 法により発生する正規乱数³⁾を利用している。よって、これら擬似乱数が一様か否か、あるいは正規分布に従うかどうかの検定をする必要があり、ここでは次に示すカイ 2 乗検定⁴⁾により行なう。

まず、 M 回の試行で級 $C_1, C_2, \dots, C_k$ の出現する観測度数をそれぞれ $x_1, x_2, \dots, x_k$ とする。各級の比率 $p_1, p_2, \dots, p_k$ について、帰無仮説 H_0 のもとで、

$$\text{帰無仮説 } H_0: p_1 = p_{10}, p_2 = p_{20}, \dots, p_k = p_{k0}$$

$$\chi^2 = \sum_{i=1}^k \frac{(x_i - M \times p_{i0})^2}{M \times p_{i0}} \quad \dots (4)$$

で定義される統計量 χ^2 が、 M がある程度大きければ、自由度 $k-1$ の χ^2 分布で近似される。有意水準 α に対して、

$$\chi^2 > \chi_{k-1}^2(\alpha) \quad \dots (5)$$

であれば、帰無仮説 H_0 が有意水準 α で棄却される。本研究では $\alpha=5\%$ としてカイ 2 乗検定を行ない、 H_0 が棄却された場合は、また新たに擬似乱数を発生させ、必要な乱数標本数を得ている。

(2) 異常値の変動特性

前節で示した、式 (2) を次のように変形する。

$$F_{N-1}^1(2\varepsilon) = \lambda^2 \frac{N+1}{N-1} \quad \dots (6)$$

$$y_\varepsilon = \bar{y} + \lambda \times S_y \quad \dots (7)$$

ここで、標本平均値 $\bar{y}=0$ 、標本分散 $S_y^2=1$ とした標準正規分布 $N(0,1)$ を考えたとき、標本数 N に対する異常値 $y_\varepsilon=\lambda$ の変動特性を調べる。

まず、式 (3) から $\beta_0=5\%$ とした場合の棄却限界 ε_0 を求め、式 (6) の ε に代入する。これにより、異常値に対応する λ が逆算して求められる。

上記の方法を用いて、 $N(0,1)$ に従う標本を $N=3 \sim 200$ 個まで変化させた場合における異常値の変動特性を図-1 に示す。図-1 で表わされる曲線上を含む右側が異常値となる領域、曲線から左側が正常値となる領域を意味している。

図-1 より、 $N=33$ 個のとき、 $\lambda=3.31$ で最小となり、 y_ε が最も異常値として判定されやすい状態であることがわ

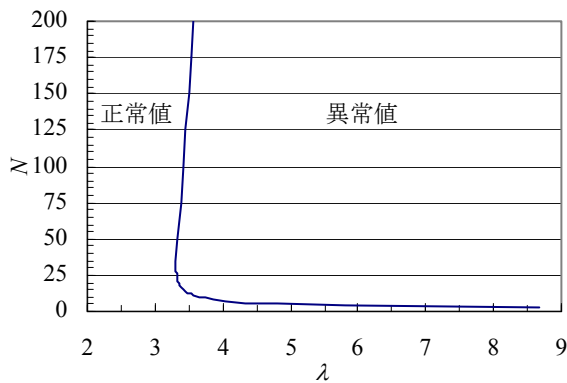


図-1 $N(0,1)$ に従う標本の個数の変化に対する異常値の変動特性 ($N>3$ 個)

かる。また、 $N<33$ 個では、 N が減少するにつれ λ が大幅に増加し、 $N=3$ 個のとき $\lambda=8.69$ で最大となっている。逆に、 $N>33$ 個では、 N が増加するにつれ λ が緩やかに増加している。つまり、 y_{ε} に対して、

- $\left\{ \begin{array}{l} N<33 \text{ 個では、} N \text{ に大きく影響される。} \\ N=33 \text{ 個では、異常値として最も判定されやすい。} \\ N>33 \text{ 個では、} N \text{ の影響をあまり受けない。} \end{array} \right.$

ことが言える。

次に、 $N(0,1)$ に従う $N=200$ 個 3000 組の正規乱数を発生させること (Monte Carlo 法) により、各標本集団の第 1 位 y_1 とその集団の異常値となる限界値 $y_{\varepsilon 0}$ との関係調べ、図-2 にそれを示す。 $y_{\varepsilon 0}$ は、図-1 で求められた $N=200$ のときの $\lambda=3.55$ を式 (7) に代入して求めており、同様に擬似乱数に対して前項で示したカイ 2 乗検定を行なっている。さらに、 $y_1 \geq y_{\varepsilon 0}$ となった場合、新たに擬似乱数を発生させ、必ず 3000 組の異常値を含まないデータが作成される様にしている。

図-2 より、 y_1 は最小 1.78～最大 3.89 の範囲 (平均 2.68) ではほぼ正規分布しているのに対して、 $y_{\varepsilon 0}$ は最小 2.95～最大 4.12 の範囲 (平均 3.52) で y_1 の範囲よりも狭い。特に $y_{\varepsilon 0}$ は平均値付近に集中しており、図-1 で求められた λ に対応していることがわかる。また、最小 2 乗法による 1 次回帰直線を見ると、相関係数 r は 0.41 で悪いが、傾きが正であることから、 y_1 が大きくなれば、 $y_{\varepsilon 0}$ も大きくなること。さらに、 y_1 が小さいとき、 $y_1=y_{\varepsilon 0}$ の破線からの距離が遠くなり、 y_1 が異常値から離れ、逆に y_1 が大きいとき、破線からの距離が近くなり、 y_1 が異常値に近くなることが示されている。

3. 道内 22 気象官署における異常降雨の特性

3.1 降水量データ

本研究で取り扱う降水量データは、年降水量、年最大日雨量及び年最大無降雨連続日数である。これらのデータは、札幌管区気象台、網走地方気象台、気象庁 HP 電子閲覧室及び気象庁年報 CD-ROM により、各官署の統計開始年から 2005 年までを取扱う。なお、欠測等により参考値となっている場合も有効とし、日界及び観測機器等の変更による誤差もここでは考慮しない。

全 22 官署の内、100 年以上の降水データを有するのは 7 官署であり、年降水量の最長統計年数は函館の 133 年

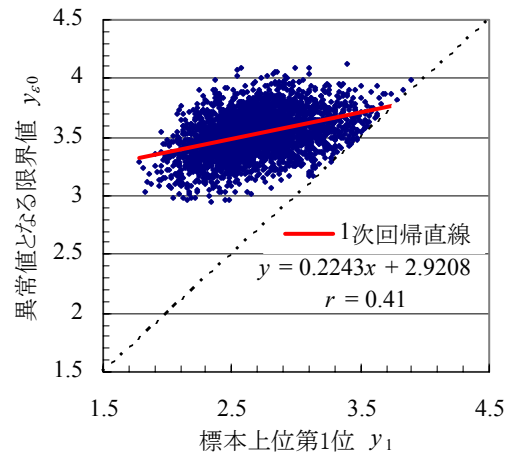


図-2 $N(0,1)$ に従う $N=200$ 個の標本 (3000 組) の第 1 位と異常値となる限界値との関係

間、年最大日雨量では札幌・函館の 123 年間、年最大無降雨連続日数については札幌の 117 年間である。逆に、最小統計年数は、各種降水データとも広尾の 48 年間である。年最大日雨量及び年最大無降雨連続日数は、雪の影響を取り除くため、5 月～11 月までを対象とした。さらに、年最大無降雨連続日数においては、“降水がわずかに観測された”とする“0.0mm”の記録は、降水有り判定している。これは、0.5mm に達しない降水が 0.0mm として記録されているためである。日数は離散量であるが、ここでは雨量と同様に扱う。

表-1 に全 22 官署における各種降水データの統計年数 (N)、上位第 1 位 (y_1)、第 2 位 (y_2) 及びそれらの差 ($y_1 - y_2$) を示す。なお年降水量の統計年数の降順を基準に並べている。また、全 22 官署中上位第 1 位に対して太字二重下線、下位第 1 位に対して単下線を引いている。本研究では、 y_1 が異常値検定の対象となる。

3.2 3 母数対数正規分布による異常値棄却検定^{1),2)}

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 μ_y 、標準偏差 σ_y の正規分布 $N(\mu_y, \sigma_y^2)$ に従う。本研究では、 a 、 μ_y 、 σ_y の 3 母数を Quantile-最尤法及び積率法の 2 種類の手法により推定する。正規化された変数 y は、前述の理論式 (6) がそのまま利用できる。すなわち、

$$\ln(x_{\varepsilon 0} - a) = \mu_y + \lambda \times \sigma_y \quad \cdots (8)$$

ここで、 x_{ε} は各気象官署における降水量データの異常値である。

まず、式 (3) から $\beta_0=5\%$ として各気象官署の統計年数に対する棄却限界 ε_0 を求め、式 (6) の ε に代入する。これにより、 ε_0 に対する 3 母数対数正規分布の異常値となる限界値が求められる。この Quantile を次式に示す。

$$x_{\varepsilon 0} = a + \exp(\mu_y + \lambda \times \sigma_y) \quad \cdots (9)$$

上記の方法により、2 種類の母数推定法により得られた 3 母数を式 (9) に当てはめ、全 22 官署における各種降水量データの異常値となる限界値の幅を図-3～5 に示す。これらの図は、 $x_{\varepsilon 0}=y_1$ を表わす破線を含む右下側が異常値を経験した領域、破線から左上側が異常値を経験

表-1 北海道内 22 気象官署の年降水量, 年最大日雨量, 年最大無降雨連続日数の統計量 (気象庁提供)

No.	官署名	年降水量 (mm/year)				年最大日雨量 (mm/day)				年無降雨連続日数 (day)			
		N	y_1	y_2	$y_1 - y_2$	N	y_1	y_2	$y_1 - y_2$	N	y_1	y_2	$y_1 - y_2$
1	函館	133	1745.5	1589.2	156.3	123	182.5	176.0	6.5	114	16	16	<u>0</u>
2	札幌	129	1671.5	1559.0	112.5	123	207.0	176.5	30.5	117	18	15	3
3	根室	126	1459.6	1439.4	<u>20.2</u>	122	211.5	147.5	64.0	114	<u>12</u>	<u>12</u>	<u>0</u>
4	寿都	118	1745.4	1702.6	42.8	122	206.3	173.5	32.8	114	21	18	<u>3</u>
5	旭川	117	1741.2	1556.1	185.1	114	184.2	167.0	17.2	114	19	19	<u>0</u>
6	網走	116	<u>1231.4</u>	<u>1152.4</u>	79.0	116	163.0	<u>121.5</u>	41.5	116	19	16	3
7	帯広	112	1489.0	1410.7	78.3	114	174.0	161.3	12.7	114	13	<u>12</u>	1
8	釧路	96	1703.9	1395.0	308.9	96	182.4	169.5	12.9	96	19	13	6
9	羽幌	85	1887.8	1734.8	153.0	85	140.0	137.0	3.0	85	17	16	1
10	室蘭	83	1594.5	1540.3	54.2	83	170.0	141.0	29.0	83	14	<u>12</u>	2
11	浦河	79	1579.8	1552.8	27.0	79	190.0	147.1	42.9	79	14	13	1
12	稚内	68	1753.7	1619.7	134.0	68	155.5	139.9	15.6	68	17	16	1
13	江差	65	1643.0	1562.0	81.0	65	173.5	169.5	4.0	65	21	13	8
14	雄武	63	1373.9	1253.5	120.4	64	140.0	125.7	14.3	63	16	15	1
15	小樽	63	1721.5	1547.4	174.1	63	161.0	129.5	31.5	63	20	17	3
16	北見枝幸	63	1834.0	1654.8	179.2	63	<u>134.0</u>	123.5	10.5	63	17	13	4
17	苫小牧	63	1816.5	1590.1	226.4	64	447.9	195.5	252.4	63	14	13	1
18	留萌	63	1900.6	1815.5	85.1	63	147.5	138.0	9.5	59	15	15	<u>0</u>
19	倶知安	62	2068.5	2023.0	45.5	62	180.9	134.5	46.4	62	15	14	1
20	岩見沢	59	1829.0	1575.5	253.5	59	262.0	213.8	48.2	59	18	15	3
21	紋別	50	1239.3	1154.0	85.3	50	150.5	148.0	<u>2.5</u>	50	15	<u>12</u>	3
22	広尾	<u>48</u>	2431.0	2407.0	24.0	<u>48</u>	346.0	244.3	101.7	<u>48</u>	13	<u>12</u>	1

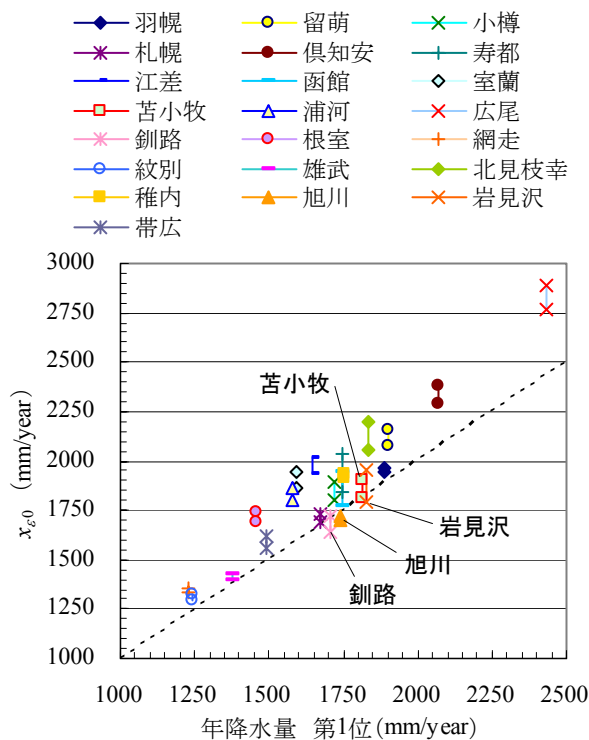


図-3 年降水量に対する道内 22 気象官署の異常降雨となる限界値の分布

していない領域となる。また、2 種類の母数推定法による異常値となる限界値の誤差幅が、垂直線によって表わされている。よって、垂直線と破線が交差しがあれば、どちらか一方の母数推定法により判定された異常値を経験していることになる。つまり、破線から右下側に離れるほど、過去に異常降雨あるいは異常渇水を経験した官署であり、逆に左上側に離れるほど、過去に異常な水文

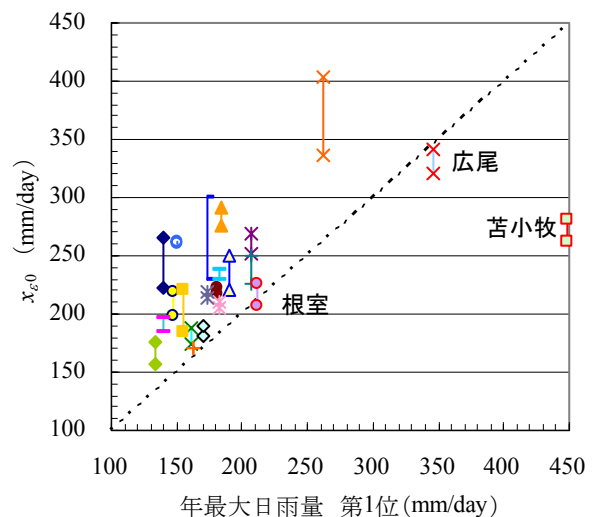


図-4 年最大日雨量に対する道内 22 気象官署の異常降雨となる限界値の分布

気象を経験していない官署とも言える。

図-3 の年降水量では、旭川のみ 2 種類の母数推定法により異常値と判定され、岩見沢、釧路及び苫小牧は片方のみで異常値と判定されていることがわかる。同様に図-4 の年最大日雨量では、苫小牧及び広尾が 2 種類、根室は片方のみで異常値と判定されている。図-5 の年最大無降雨連続日数では、江差のみ 2 種類で異常値と判定されている。

図-3 の年降水量では、破線から右に大きく外れた異常値となる顕著な値は出現していない。しかし、図-4 の年最大日雨量では、苫小牧の第 1 位が非常に大きく外れており、他の 2 官署は破線に近く、次節で触れるが、将来においては異常値と判定されなくなる可能性が高い。同様に図-5 の年最大無降雨連続日数においては、江差が右

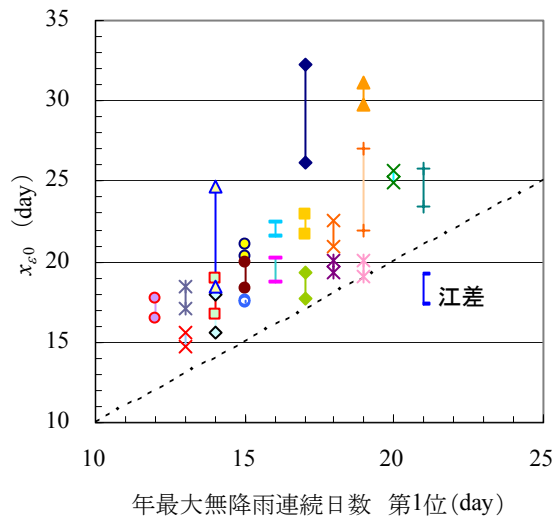


図-5 年最大無降雨連続日数に対する道内 22 気象官署の異常無降雨連続日数となる限界値の分布

側に大きく外れているが、雨量と違って日数であり、値そのものの大きさが雨量と比較して小さいことから、やはり将来においては、異常値と判定されなくなる可能性がある。

3.3 異常値と判定された値が異常値でなくなる条件

第1章で述べた様に、異常値に近い値が連続して観測されれば、その値は異常値とは言えなくなる。前節までにおいて、過去の経験から異常値と判定された値の特徴を見ると、その一つには、表-1 で示されるように y_1 、 y_2 の値の大きいことが挙げられる。

そこで、前節において y_1 が異常値であると判定された官署に対して、2006 年に上位第2位となる降水データを観測すると仮定したとき、 y_1 が異常値でなくなる y_2^* の値を調べる。

まず、異常値を含む各官署の標本集団に対して、 $y_2 \leq y_2^* < y_1$ の範囲で、 y_2^* の値を刻み幅 0.5mm (または 1 日) で毎回変化させて与える ($N^* = N+1$)。刻み幅を 0.5mm としたのは、現在の観測雨量が 0.5mm 単位だからである。 y_2^* の値を与えて変化させる毎に、対数正規分布の 3 母数を新たに更新し、式 (3) 及び式 (8) を利用して、 $\varepsilon_0 < \varepsilon$ となる ε を求める。そのときの y_2^* が、 y_1 を異常値でないと判定させる降水データとなる。ここでは、2 種類の母数推定法の両方で $\varepsilon_0 < \varepsilon$ になった場合を y_2^* としている。上記の方法により、各種降水データにおいて異常値と判定された官署について調べた結果を表-2 に示す。

表-2 から、年降水量ではいずれの官署においても 1400 ~ 1600mm (日本全国の平年値: 約 1700mm) の雨量を観測すれば、 y_1 は異常値として判定されなくなることがわかる。年最大日雨量においては、苫小牧で 322.0mm/day の雨量を観測すれば、 y_1 が異常値として判定されなくなる。この値は非常に大きいように思われるが、2003 年 8 月の台風 10 号により、苫小牧にそれほど遠くない日高支庁日高で 308mm/day を観測している。また、江差の 18 日は、2006 年 7~8 月にかけて、降水量 0.0mm も含めて

表-2 y_1 を異常値でないと判定させる 2006 年の上位第2位の降水データ y_2^*

(A: 年降水量, B: 年最大日雨量, C: 年最大無降雨連続日数)

官署名		N^*	y_1	y_2^*	y_2
A	苫小牧	64	1816.5mm	1590.1	1590.1
	釧路	97	1703.9	1440.5	1395.0
	旭川	118	1741.2	1578.6	1556.1
	岩見沢	60	1829.0	1580.5	1575.5
B	苫小牧	65	447.9mm	322.0	195.5
	広尾	49	346.0	254.8	244.3
	根室	123	211.5	147.5	147.5
C	江差	66	21 日	18 日	13 日

カウントした場合、25 日間にも及ぶ記録的な渇水期を観測しているため、近い将来実現しうる値と言える。

4. 結論

本研究により得られた結果を以下に示す。

- (1) 標準正規分布に従う標本の個数の変化に対する異常値の変動特性を示した。特に、異常値の判定は $N < 33$ 個のとき、 N に大きく影響され、 $N = 33$ 個で、異常値として最も判定されやすく、 $N > 33$ 個では、 N の影響をあまり受けないことを明らかにした。
- (2) $N(0,1)$ に従う $N=200$ 個 3000 組の正規乱数を発生させ、各標本集団の上位第1位と異常値となる限界値との相関は悪いことを示した。
- (3) 道内 22 気象官署における年降水量、年最大日雨量、年最大無降雨連続日数に対する異常降雨と上位第1位との関係を 2 種類の母数推定方法による 3 母数対数正規分布を用いて明らかにした。
- (4) 過去に経験した降雨から異常値として判定された値が、2006 年に上位第2位となる降水データを観測すると仮定したとき、過去の第1位と第2位の平均値程度の値であれば、異常値でなくなることを示した。

以上より、異常値と判定された値は、将来観測される降水データにより、異常値として判定されなくなる可能性が低いとは言えないため、この取り扱いについては、簡単に棄却できるものではなく、十分慎重に議論する必要があるだろう。

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, 17 年度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) 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, 1961.7.
- 2) 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, 1998.7.
- 3) 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ー-数値計算, 統計計算, 3次元CGライブラリー, 森北出版株式会社, 1998.2.
- 4) 田代嘉宏, 脇本和昌, 大崎紘一: 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.