

北海道内降水量に対する新しい確率降水量計算法の開発

Development of New Calculation Method for Probable Precipitation in Hokkaido

佐渡公明
Kimiteru SADO

北見工業大学 名誉教授

要 旨

最近の地球温暖化や気候変動が顕在化している気象状況下においては、降水量極値（年最大日雨量、年最大無降雨連続日数）の母集団は変化している可能性がある。従って、治水、利水計画のための確率日雨量、確率無降雨連続日数の推定には、母集団変化に対応できる非定常頻度解析を用いる必要がある。

本論文では、まず北海道内 22 気象官署の開所年～2006 年までの年最大日雨量、年最大無降雨連続日数データに対し、各官署の全統計年数を前半、後半部分時系列の 2 つに分け、この 2 つの部分標本に対し平均値と分散のジャンプ検定を行う。直近の 30 年間以上にわたるハイジャンプが有意な官署については、ハイジャンプが有意となった後半部分標本に対してのみ T 年確率日雨量、T 年確率無降雨連続日数を推定する。これにより、従来全統計年数に対して定常頻度解析により推定された確率降水量を更新する非定常頻度解析を提案する。

北海道内 22 気象官署の年最大日雨量データについては、ハイジャンプが有意な 12 官署の 50, 100, 150, 200 年確率日雨量は、従来の推定値の 0.1～49%増の更新が必要であると得られた。また、年最大無降雨連続日数データについては、ハイジャンプが有意な 5 官署に対する 5, 10, 15, 20 年確率無降雨連続日数は、従来の推定値の 2～14%増の更新が必要である。

《キーワード：年最大日雨量；年最大無降雨連続日数；平均値ジャンプ；分散ジャンプ；T 年確率日雨量；T 年確率無降雨連続日数；治水安全率；利水安全率》

1. はじめに

河川の基本高水のピーク流量や維持流量の算定に用いられるT年確率日雨量やT年確率無降雨連続日数の推定は、年最大日雨量や年最大無降雨連続日数が定常確率過程を示すという仮定のもとに行われている。しかし、最近の地球温暖化や気候変動が顕在化している気象状況下においては、降水量極値（年最大日雨量、年最大無降雨連続日数）の母集団は変化している可能性がある。従って、母集団変化に対応できる非定常頻度解析を用いた確率日雨量、確率無降雨連続日数の推定が必要である。このような指摘は、Z. W. Kundzewicz¹⁾、Z. W. Kundzewicz and A. J. Robson²⁾、徐・竹内・石平³⁾、西岡・寶⁴⁾、寒川・中村⁵⁾ など多くの論文に取り上げられている。筆者らも北海道内22気象官署の年最大日雨量データを用いて100年確率日雨量の経年変化を明らかにし、6官署において治水安全率が低下していることを示している⁶⁾。

本論文では、まず北海道内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対し、各官署の全統計年数 N を大きさ n_1 の前半部分標本及び大きさ n_2 の後半部分標本に分け($N=n_1+n_2$)、2つの部分標本に対し平均値と分散のジャンプ検定を行なう⁷⁾。直近の30年間以上($n_2 \geq 30$)にわたるハイジャンプ(n_2 期間の平均値、分散が n_1 期間の平均値、分散より大きい)が有意な官署については、ハイジャンプが有意となった n_2 期間のみの部分標本に対しT年確率日雨量、T年確率無降雨連続日数を推定することにより、 N 期間の確率降水量を更新する非定常頻度解析を提案する。他方、直近の30年間以上のハイジャンプが有意でないときまたはロージャンプ(n_2 期間の平均値、分散が n_1 期間の平均値、分散より小さい)が有意なときは、従来と同じく全統計年数データを用いた定常頻度解析により確率降水量を推定する。これにより、その地域の治水・利水安全率を高めることができ、その理論的根拠を示すことができる。

普通、気候変動解析に用いられるトレンド成分解析では、気候変動による確率水文量への影響を推定出来ないので、本論文ではジャンプ検定を用いて確率水文量への影響を調べる。最近時々発生している超過洪水は、基本高水のピーク流量のリターンピリオドが計画当時よりは下がっていることが原因で発生していることも予想され、これに関連して本研究は、治水安全率の更新の必要性を暗示している。

2. 最近30年間以上にわたる年最大日雨量の平均値、分散のジャンプ検出

北海道内22気象官署における年最大日雨量の観測時系列 $X(=x_1, \dots, x_N)$ を、大きさ n_1 の前半部分標本 $X_1(=x_1, \dots, x_{n_1})$ と大きさ n_2 の後半部分標本 $X_2(=x_{n_1+1}, \dots, x_{n_1+n_2})$ ($n_1+n_2=N$)に分け、 X_1 、 X_2 に対し平均値と分散のジャンプ検定を行なう。平均値のジャンプについてはt検定(以下t-tと略記)、Mann-Whitney検定(MW)、ブートストラップt検定(BS-t)、ブートストラップMann-Whitney検定(平均値BS-MW)を用い、さらに分散ジャンプについてはF検定(F-t)、ブートストラップF検定(BS-F)、著者らが提案している分散ジャンプ用ブートストラップMann-Whitney検定(分散BS-MW)⁷⁾を用いる。ブートストラップジャンプ検定法については付録に詳述している。

表-1 札幌の年最大日雨量の有意な平均値ジャンプと n_2 期間のT年確率日雨量

No.	$\bar{x}_1$ (mm/ 日)	$\bar{x}_2$ (mm/ 日)	Jump (mm/ 日)	n_1 年	n_2 年	4 検定手法				X50 (mm/ 日)	X100 (mm/ 日)	X150 (mm/ 日)	X200 (mm/ 日)	Model
						t-t	MW	BS -t	BS- MW					
1	65.1	79.2	14.1	79	45	1	1	1	1	187.7	216.4	234.1	246.9	SqrtEt
2	65.0	79.0	14.0	78	46	1	1	1	1	186.3	214.6	232.0	244.7	SqrtEt
3	65.0	78.7	13.7	77	47	1	1	1	1	184.4	212.4	229.5	242.0	SqrtEt
4	65.2	78.2	13.1	76	48	1	1	1	1	182.2	209.6	226.4	238.7	SqrtEt
5	64.7	78.4	13.6	74	50	1	1	1	1	181.9	209.3	226.0	238.2	SqrtEt
6	64.5	78.4	13.8	73	51	1	1	1	1	181.5	208.7	225.4	237.5	SqrtEt
7	64.9	77.5	12.6	72	52	1	1	1	1	181.4	208.8	225.7	237.9	SqrtEt
8	64.9	77.2	12.3	70	54	1	1	1	1	180.5	207.8	224.5	236.8	SqrtEt
9	65.3	77.8	12.5	75	49	1	1	1	1	179.8	206.8	223.3	235.3	SqrtEt
10	65.3	76.9	11.6	71	53	1	1	1	1	179.5	206.7	223.4	235.5	SqrtEt
11	64.4	76.8	12.4	66	58	1	1	1	1	179.4	206.7	223.4	235.6	SqrtEt
12	65.1	76.6	11.5	69	55	1	1	1	1	178.5	205.4	221.9	234.0	SqrtEt
13	64.1	77.4	13.3	67	57	1	1	1	1	177.7	203.0	218.4	229.6	Iwai
14	64.3	76.5	12.2	64	60	1	1	1	1	176.6	201.7	216.9	228.0	Iwai

データ数と年最大日雨量の確率水文量との関係について、寒川ら⁸⁾は長野県下の3観測所について年最大日雨量の分布にGumbel分布を用い、データ数が30～40個以上になると50年、100年確率日雨量が安定してくることを示した。ここでは、 X_1 と X_2 に対するハイジャンプが有意なときは、 n_2 期間の部分標本に対してT年確率日雨量を計算する。従って、 n_2 を30個以上とした。

一般に標本の大きさが大きければブートストラップ反復回数 M も大きくとる必要がある⁹⁾。 $M=3000$ とすると、 $6!=720 < M=3000 < 7!=5040$ の関係から、標本の大きさが6のとき、最少でも $3000-720=2280$ 組のブートストラップ標本は一度抽出された標本と全く同じ標本となり不自然である¹⁰⁾。これを防ぐために、 $n_1 \geq 7$ とした。

22官署のジャンプ検出を行なう前に、角屋の方法¹¹⁾を用いて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、苫小牧における1950年の447.9mm/day、広尾における1998年の356.0mm/dayが異常値と判定された⁶⁾。そこで、データに異常値が含まれると確率水文量に大きな誤差を与える可能性があるため、これらの異常値を前後年の平均値で置き換えた。

(1) 平均値のジャンプ検出

表-1は一例として札幌の年最大日雨量データに対し、平均値のジャンプ検出を行なった結果である。 $\bar{x}_1$, $\bar{x}_2$ は X_1 , X_2 の平均値, $Jump = \bar{x}_2 - \bar{x}_1$, 4検定手法の欄の1はハイジャンプが有意なことを示し, $X50 \sim X200$ は最適確率分布を用いて計算した n_2 期間に対する50～200年確率日雨量の値である。最後の欄のModelは、9種類の確率分布モデル（一般化極値分布(GEV), PearsonIII型分布(P3), 対数PearsonIII型分布(LP3), 2母数対数正規分布(LN2), 3母数対数正規分布(LN3), 岩井法(Iwai), Gumbel分布(Gumbel), Gumbel-Chow法(G-C), 平方根指数型最大値分布(SqrtEt))の中から、標準最小二乗規準 (SLSC)¹²⁾を最小にさせるという判断基準を用いて採用した最適確率分布である。 $n_1 \geq 7$, $n_2 \geq 30$, $n_1 + n_2 = 124$ を満足する部分標本 X_1 , X_2 の88通りの組み合わせについて平均値ジャンプの検定を行ない、4検定手法全て

でハイジャンプが有意となった14通りの組合わせを50年確率日雨量 X_{50} の大小順に示している。リターンピリオド T が異なると、それに対応する確率水文量の大小順が一般に異なる。

No.1の組合わせが、平均値ジャンプと T 年確率日雨量が共に最大である。このときの時系列を図-1(a)に示す。青の点線が N 年間の平均値、赤の実線が1961年に平均値ジャンプ14mm/dayが発生したときの、ジャンプ発生前の時系列 X_1 と発生後の時系列 X_2 の平均値を示している。

以下同様に、他の官署についても平均値ジャンプ検定を行ない、4検定手法全てにおいてハイジャンプが有意となった官署は、苫小牧、浦河、寿都、倶知安、小樽、室蘭であり、道内14支庁の内、道南の胆振・日高・後志地方に集中している。逆に、4検定手法全てにおいてロージャンプが有意となった官署はなかった。ハイジャンプが最大となった苫小牧に対し、 X_{100} が最大となる X_1 、 X_2 の組合せについて、その時系列を示したのが図-1(b)である。 N 年間の平均値を挟んでハイジャンプが生じている様子が分かる。

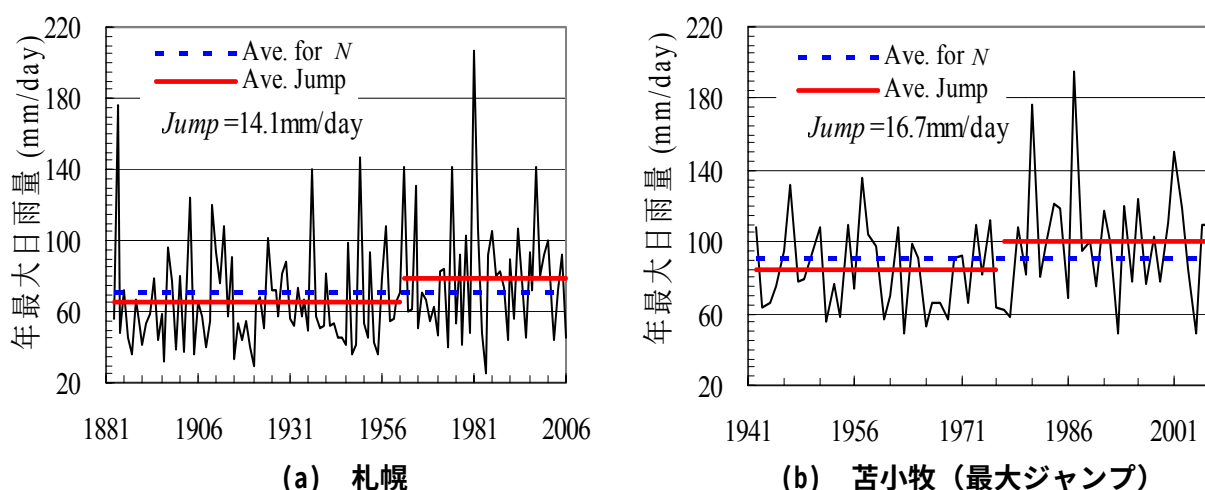


図-1 n_2 期間の100年確率日雨量が最大となる平均値ジャンプの例（(a) 札幌 (b) 苫小牧）

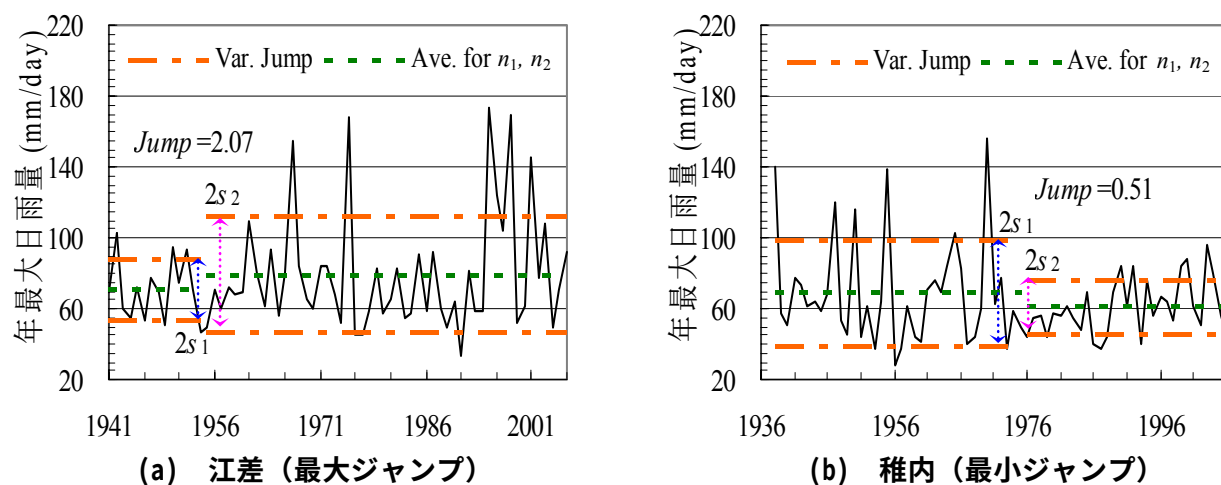


図-2 n_2 期間の100年確率日雨量が最大となる分散ジャンプの例（(a) 江差 (b) 稚内）

(2) 分散のジャンプ検出

前節と同様に各官署に対し、 $n_1 \geq 7$, $n_2 \geq 30$, $n_1 + n_2 = N$ を満足する部分標本 X_1 , X_2 の全組合わせについて分散ジャンプの検定を行ない、3検定手法全てでハイジャンプが有意となった組合せについて X_{50} , X_{100} , X_{150} , X_{200} を求めた。

3検定手法全てにおいてハイジャンプが有意となった官署は、網走、寿都、帯広、根室、旭川、江差、小樽、逆にロージャンプが有意となった官署は函館、稚内であり、全道的に分布している。分散ジャンプが有意となった官署に対し、 X_{100} が最大となる X_1, X_2 の組合わせについて、ジャンプの最大(江差)、最小(稚内)の場合の時系列を図-2(a), (b)に示す。図中の $Jump$ は、 $Jump = s_2/s_1$ (s_1, s_2 は X_1, X_2 の標準偏差)である。

3. 平均値、分散のジャンプ検出によるT年確率日雨量の更新

第2章の最近30年間以上にわたる年最大日雨量の平均値、分散のジャンプ検定結果を用いて、統計年数 N 年間に対し定常頻度解析により推定した T 年確率日雨量の更新を考える。本章では、 N 年間に対し最適確率分布を用いて推定した T 年確率日雨量を XT 、また、平均値ジャンプまたは分散ジャンプが有意で n_2 年間に対し、最適確率分布を用いて推定した T 年確率日雨量をそれぞれ YT , ZT とする。

ロージャンプが有意となった函館、稚内における $Jump$ は0.68, 0.51であった。このとき函館、稚内の X_{100} は170.6, 158.3mm/day (最適確率分布は第2章(1)の9種類の中から共にGEV)であるのに対し、 Z_{100} はそれぞれ159.2, 107.7mm/day (最適確率分布はそれぞれLP3, LN2)と低下している。このように、ロージャンプが有意なときは、 XT に対し YT , ZT は当然低下するが、行政的には治水施設整備の低下を避けるための XT の更新は不要とする。

ハイジャンプが有意なときに、 $XT < YT$, $XT < ZT$ となり、 XT の更新が必要となる。この場合、次の3通りが考えられる。

- I 型…平均値ジャンプのみが有意で XT を YT で更新
- II 型…分散ジャンプのみが有意で XT を ZT で更新
- III 型…平均値と分散ジャンプの両方が有意で、 XT を YT と ZT の大きい方で更新

以上の結果を増加率 YT/XT または ZT/XT の大きい順にまとめたのが表-2である。I 型には札幌、苫小牧、浦河、倶知安、室蘭が入り、II 型には網走、帯広、根室、旭川、江差が入り、III 型には寿都、小樽が入っている。No.1の網走は、開所年～2006年の統計年数117年に対し、GEVを用いて X_{50} , 100, 150, 200=111.5, 128.2, 138.6, 146.4mm/dayと推定された。しかし、 $T=50$, 100年に対し $n_1=87$ と $n_2=30$ ($T=150$, 200年では $n_1=86$ と $n_2=31$)の部分標本で ZT を最大とさせる分散ジャンプが発生し、 Z_{50} , 100, 150, 200=150.1, 181.3, 202.1, 218.3mm/dayと増加し、最大の増加率35, 41, 46, 49% ($T=50$, 100, 150, 200に対し)を示している。更新すべき T 年確率日雨量の増加率 YT/XT または ZT/XT は、 $T=50$ 年に対し1.4～35%, $T=100$ 年で0.4～41%, $T=150$ 年で0.1～46%, $T=200$ 年で4.4～49%である。同一官署では T の増加と共に増加率が増える傾向にある。表-2において増加率が1未満の YT, ZT は更新の対象から外す。

表-2 全統計年数に対するT年確率日雨量 XT から
 平均値ジャンプによる YT ，分散ジャンプによる ZT への更新

No.	官署	N 年	最適 PDF	X50 mm/日	更新 型	Jump		n_2 年	最適 PDF	Y50 mm/日	Z50 mm/日	Y50/X5 0	Z50/X5 0
						$\bar{x}_2 - \bar{x}_1$	s_2/s_1						
1	網走	117	GEV	111.5	II	—	1.64	30	LP3	—	150.1	—	1.346
2	寿都	123	GEV	141.5	III	—	1.79	48	LP3	—	183.9	—	1.300
						10.5	—	84	LP3	153.8	—	1.087	—
3	帯広	115	Iwai	147.9	II	—	1.36	30	LN2	—	182.7	—	1.235
4	根室	123	P3	142.4	II	—	1.32	44	LN2	—	172.3	—	1.210
5	札幌	124	LP3	159.3	I	14.1	—	45	SqrtEt	187.7	—	1.178	—
6	旭川	115	P3	150.3	II	—	1.51	53	Iwai	—	175.8	—	1.170
7	苫小牧	65	Gumbel	168.7	I	16.7	—	30	LP3	193.7	—	1.148	—
8	江差	66	LP3	167.4	II	—	2.07	49	Iwai	—	183.2	—	1.094
9	浦河	80	LN3	151.9	I	11.6	—	45	GEV	163.8	—	1.078	—
10	小樽	64	LP3	136.2	III	—	1.75	48	LN2	—	144.7	—	1.062
						16.7	—	52	Gumbel	139.0	—	1.021	—
11	倶知安	63	GEV	145.1	I	14.2	—	46	LP3	154.0	—	1.061	—
12	室蘭	84	GEV	147.2	I	12.6	—	64	P3	149.2	—	1.014	—

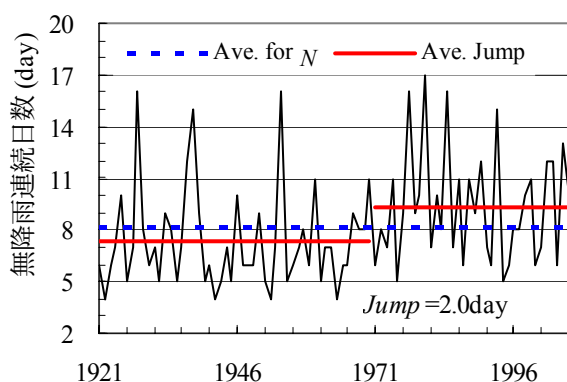
No.	官署	X100 mm/日	Y100 mm/日	Z100 mm/日	Y100/X10 0	Z100/X10 0	X200 mm/日	Y200 mm/日	Z200 mm/日	Y200/X20 0	Z200/X20 0
1	網走	128.2	—	181.3	—	1.414	146.4	—	218.3	—	1.491
2	寿都	164.0	—	224.8	—	1.371	189.1	—	273.3	—	1.445
			180.3	—	1.099	—		210.2	—	1.112	—
3	帯広	163.9	—	204.4	—	1.247	167.7	—	214.8	—	1.281
4	根室	155.2	—	193.3	—	1.245	180.1	—	226.5	—	1.258
5	旭川	171.2	—	206.1	—	1.204	192.3	—	238.8	—	1.242
6	札幌	183.5	216.4	—	1.179	—	209.7	246.9	—	1.177	—
7	苫小牧	184.9	214.2	—	1.158	—	201.0	235.1	—	1.170	—
8	浦河	175.1	193.1	—	1.103	—	200.1	227.1	—	1.135	—
9	江差	194.6	—	211.8	—	1.088	225.0	—	243.4	—	1.082
10	小樽	152.8	—	162.3	—	1.062	170.0	—	180.2	—	1.060
			154.1	—	1.009	—		169.2	—	0.995	—
11	倶知安	171.2	180.4	—	1.054	—	201.3	210.2	—	1.044	—
12	室蘭	162.9	163.5	—	1.004	—	179.1	177.8	—	0.993	—

4. 平均値，分散のジャンプ検出によるT年確率無降雨連続日数の更新

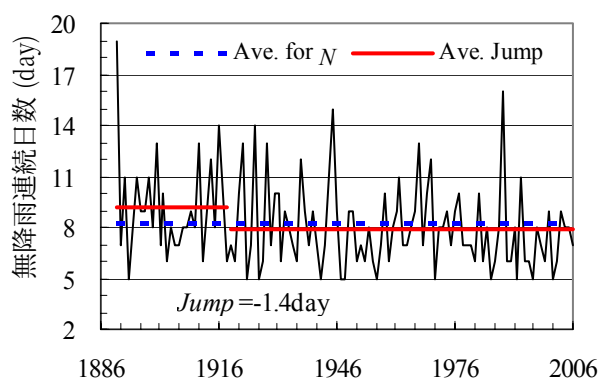
本章では， N 年間に對し最適確率分布を用いて推定したT年確率無降雨連続日数を XT ，平均値または分散ジャンプが有意で n_2 年間に對し最適確率分布を用いて推定したT年確率無降雨連続日数をそれぞれ YT ， ZT とする。

(1) 平均値，分散のジャンプ検出

北海道内22気象官署の無降雨連続日数データに対する平均値，分散のジャンプ検定を第2章(1)，(2)と同様の手法で行なった。

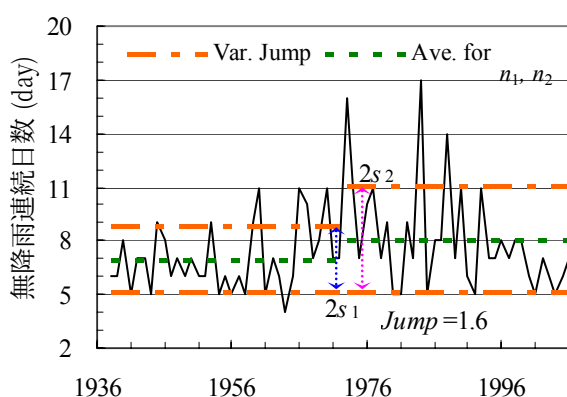


(a) 羽幌 (最大ジャンプ)

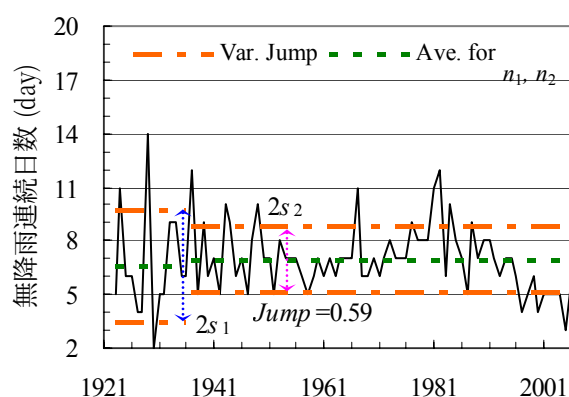


(b) 網走 (最小ジャンプ)

図-3 n_2 期間の10年確率無降雨連続日数が最大となる平均値ジャンプの例 ((a) 羽幌 (b) 網走)



(a) 稚内 (最大ジャンプ)



(b) 室蘭 (最小ジャンプ)

図-4 n_2 期間の10年確率無降雨連続日数が最大となる分散ジャンプの例 ((a) 稚内 (b) 室蘭)

リターンピリオドが10年の場合、平均値ジャンプが有意な官署名とY10を最大にするジャンプをジャンプの大小順に列記すると、ハイジャンプが羽幌($Jump=2.0$ 日)、稚内(1.3)、ロージャンプが根室(-1.0)、函館(-1.0)、浦河(-1.1)、旭川(-1.2)、網走(-1.4)であった。図-3に羽幌、網走の時系列を示す。同様に、分散ジャンプが有意な官署名とZ10を最大にするジャンプを列記すると、ハイジャンプが稚内($Jump=1.64$)、釧路(1.61)、浦河(1.54)、ロージャンプが寿都(0.71)、北見枝幸(0.64)、小樽(0.61)、室蘭(0.59)であった。図-4に稚内、室蘭の時系列を示す。

(2) T年確率無降雨連続日数の更新

第3章と同様にハイジャンプが有意なときに、 XT を YT または ZT で更新することになり、更新型にはI、II、III型がある。T=5、10、20年に対する結果を表-3に示す。羽幌がI型、釧路と小樽がII型、稚内と浦河がIII型である。

更新すべきT年確率無降雨連続日数の増加率は、T=5年に対し2～12%、T=10年に対し3～12%、T=15、20年に対し5～14%であり、T年確率日雨量の増加率と比べ小さいことが分かる。T=5年のとき最大増加率12%は羽幌で起こり、 $X5=10.5$ 日を平均値ジャンプによる $Y5=11.8$ 日で更新することが必要である。T=10、15、20年の場合の最大増加率12、14、14%は、全て稚内で生じており、 $X10, 15, 20=10.6, 11.6,$

表-3 T年確率無降雨連続日数 X_T から Y_T , Z_T への更新

No.	官署	N 年	最適 PDF	X5 日	更新型	Jump		n_2 年	最適 PDF	Y5 日	Z5 日	Y5/X5	Z5/X5
						$\bar{x}_2 - \bar{x}_1$	s_2/s_1						
1	羽幌	86	P3	10.5	III	2.0	—	36	LN2	11.8	—	1.124	—
2	稚内	69	GEV	9.0	I	1.7	—	41	GEV	9.9	—	1.100	—
3	釧路	97	GEV	8.7	II	—	1.64	34	LP3	—	9.9	—	1.100
4	浦河	80	Iwai	8.4	III	—	1.54	54	LP3	—	8.6	—	1.024
5	小樽	64	GEV	10.2	II	-1.10	—	39	LN3	7.6	—	0.905	—
						—	0.61	55	Iwai	—	10.4	—	1.020

No.	官署	X10 日	Y10 日	Z10 日	Y10/X10	Z10/X10	X20 日	Y20 日	Z20 日	Y20/X20	Z20/X20
1	羽幌	12.4	13.7	—	1.105	—	14.2	15.5	—	1.092	—
2	稚内	10.6	11.8	—	1.113	—	12.4	13.8	—	1.113	—
3	釧路	10.2	—	11.9	—	1.123	—	—	14.1	—	1.137
4	浦河	9.8	—	11.2	—	1.098	11.9	—	13.5	—	1.134
5	小樽	12.2	—	10.1	—	1.031	11.1	—	11.7	—	1.054
			8.9	—	0.91	—	10.2	—	0.92	—	—
			—	12.2	—	1.000	14.3	—	14.0	—	0.979

12.4日を分散ジャンプによるZ10, 15, 20=11.9, 13.2, 14.1日で更新することが必要である。

以上のように、年最大日雨量についてはハイジャンプの官署数が多く、年最大無降雨連続日数についてはロージャンプの官署数が多い。また、更新すべきT年確率日雨量の増加率はT年確率無降雨連続日数の増加率より大きい。これは、最近の北海道内の気候変動の影響が、利水計画に関連する年最大無降雨連続日数よりは、治水計画に関連する年最大日雨量に対して強く影響していることを示唆していることになる。

5. 結 論

本研究により得られた結果を以下にまとめる。

- (1) 最近の気候変動を考慮した確率水文量の非定常頻度解析の一方法として、平均値、分散のジャンプ検出に基づいて、従来の全統計年数に対する定常頻度解析によるT年確率水文量を更新する方法を提案した。
- (2) 北海道内22気象官署の年最大日雨量データにジャンプ検定を行なった結果、平均値ジャンプが有意な官署数はハイジャンプが7官署、ロージャンプが0、分散ジャンプが有意な官署数はハイジャンプが7、ロージャンプが2であった。
- (3) (2)のハイジャンプが有意な12官署に対するT年確率日雨量は、従来の推定値の0.1～49%増の更新が必要であり、同一官署においてはリターンピリオドの増加と共に更新割合が増える傾向にある。
- (4) 年最大無降雨連続日数データについて(2)と同様のジャンプ検定を行なった結果、有意な官署数は平均値のハイ、ロージャンプについてそれぞれ2, 5官署、分散のハイ、ロージャンプについてそれぞれ3, 4官署であった。

- (5) (4)のハイジャンプが有意な5官署に対するT年確率無降雨連続日数は、従来の推定値の2～14%増の更新が必要である。
- (6) 最近の北海道内の気候変動の影響は、利水計画に関連する年最大無降雨連続日数よりは治水計画に関連する年最大日雨量に対して強く影響していることを示唆している。

謝辞：本研究は、(財)北海道河川防災研究センターの平成19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。

付録

1. 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 $n1$ 個標本 $X_1(=x_1, \dots, x_{n1})$ とジャンプ発生後の $n2$ 個の標本 $X_2(=x_{n1+1}, x_{n1+2}, \dots, x_{n1+n2})$ ($n1+n2=N$) に分ける。両標本の母集団平均値を μ_1, μ_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。

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{u_1^2}{n1} + \frac{u_2^2}{n2}}} \quad (1)$$

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 δ 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⁹⁾。

$$\begin{aligned} y_i &= x_i - \bar{x}_1 + \bar{X}, \quad i=1, \dots, n1 \\ y_{n1+j} &= x_{n1+j} - \bar{x}_2 + \bar{X}, \quad j=1, \dots, n2 \end{aligned} \quad (2)$$

式(2)の y_i, y_{n1+j} からのブートストラップ標本をそれぞれ、 $y_i^*, i=1, \dots, n1$ および $y_{n1+j}^*, j=1, \dots, n2$ とし、ブートストラップ検定統計量は次式で与えられる。

$$\hat{t}^* = \frac{\bar{y}_1^* - \bar{y}_2^*}{\sqrt{\frac{u_1^{*2}}{n1} + \frac{u_2^{*2}}{n2}}} \quad (3)$$

ただし、 $\bar{y}_1^*, \bar{y}_2^*$ および u_1^{*2}, u_2^{*2} は y_i^*, y_{n1+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ィングポジション 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 $m_i \sim m_i + 1$ の範囲に入ったとき、 t_0 の非超過確率は $p_i = (m_i - 0.4)/(M + 0.2)$ 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。

$$\left. \begin{array}{ll} p_t < 0.025 & \cdots H_0 \text{を棄却, ハイジャンプ} \\ 0.025 \leq p_t \leq 0.975 & \cdots H_0 \text{を採択, ジャンプなし} \\ 0.975 < p_t & \cdots H_0 \text{を棄却, ロージャンプ} \end{array} \right\} \quad (5)$$

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. **1.(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 $x_i, i=1, \dots, n1$ と $x_{n1+j}, j=1, \dots, n2$ 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_{x0} を次式の順位和で求める.

$$T_{x0} = \sum_{i=1}^{n1} R(x_i) \quad (6)$$

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_{n1+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_{x0} が $\hat{T}_x^*$ の $m_T \sim m_T+1$ の範囲に入ったとき, T_{x0} 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

1.(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 $x_i, i=1, \dots, n1$ と $x_{n1+j}, j=1, \dots, n2$ に分ける ($n1+n2=N$). 両標本の正規母集団分散を σ_1^2, σ_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_{n1+j} の分散 $V(y_i), V(y_{n1+j})$ が X の分散 $V(X)$ に等しくなるように次の変換を行う.

$$\left. \begin{array}{l} y_i = \left(\frac{V(X)}{V(x_i)} \right)^{\frac{1}{2}} x_i, i=1, \dots, n1 \\ y_{n1+j} = \left(\frac{V(X)}{V(x_{n1+j})} \right)^{\frac{1}{2}} x_{n1+j}, j=1, \dots, n2 \end{array} \right\} \quad (8)$$

式(8)からのブートストラップ標本を y_i^*, y_{n1+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_{n1+j}^* の標本不偏分散) とし, F_0 の非超過確率 p_F は, **1.(1)**と同様にカナン公式から得られ, $\alpha=0.05$ の両側検定の場合, 式(5)の p_i を p_F 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

1.(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 $|x_i - \bar{x}|$ 差で置き換え, 発生後時系列 x_{n1+j} を絶対偏差 $|x_{n1+j} - \bar{x}_2|$ で置き換える. その後は, **1.(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.

参考文献

- 1) Kundzewicz, Z.W. : Searching for change in hydrological data, *Hydrol. Sci. J.* , Vol.49, No.1, pp.3-6, 2004.
- 2) 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.
- 3) 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及びトレンドに関する研究, 水工学論文集, 第46巻, pp.121-126, 2002.
- 4) 西岡昌秋, 寶 馨: Mann-Kendall検定による水文時系列の傾向変動, 京都大学防災研究所年報, 第46号B, pp.181-192, 2003.
- 5) 寒川典昭, 中村 哲: 日高川流域の月・季節・年降水量の非定常頻度解析, 水工学論文集, 第49巻, pp.7-12, 2005.
- 6) 杉山一郎, 佐渡公明: 北海道内22気象官署における降水量の非定常頻度解析, 水工学論文集, 第50巻, pp.187-192, 2006.
- 7) 佐渡公明, 中尾隆志, 杉山一郎: 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, 水工学論文集, 第52巻, 36, 2008.
- 8) 寒川典昭, 荒木正夫, 渡辺輝彦: 確率分布の推定母数の不確定性評価法, 土木学会論文集, 第375号/II-6, pp.133-141, 1986.
- 9) 汪 金芳, 田栗正章, 手塚 集, 樺島祥介, 上田修功: 計算統計 I - 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) 奥村晴彦: パソコンによるデータ解析入門—数理とプログラミング実習, 技術評論社, pp.215-224, 1990.
- 11) 角屋 睦: 異常(確率)水文量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.
- 12) 寶 馨, 高棹琢馬: 水文頻度解析における確率分布モデルの評価規準, 土木学会論文集, 第393号/II-9, pp.151-160, 1988.