2 変量同時極値の数値シミュレーションを用いた降水量の依存性の統計特性

名古屋工業大学大学院工学研究科　学生会員　　国門卓也
名古屋工業大学大学院工学研究科　学生会員　○今井遼太
名古屋工業大学大学院工学研究科　正 会 員　　北野利一

1.　研究背景と目的

平成 30 年 7 月 6 月 28 から 7 月 8 日において梅雨前線の記録的な豪雨はや台風 7 号の影響による記録的な豪雨は死者・行方不明者 230 人を超える平成最悪の豪雨災害であった [1]．このとき，岡山県高梁川水系小田川の支流では，流域内の多地点で同時に極大な降雨となり，異なった支川の合流部にて，堤防が決壊し，外水氾濫が起こった．このような多地点(2 地点)の極大降水量の同時生起に対して，同時生起の頻度や空間上の依存性の程度を 2 変量極値解析で予測することが，上記にふれたような災害に対する準備・対策につながる．2 変量(多変量)の極値理論は 2 地点の依存性について考慮しなくてはならないため，1 変量極値理論より複雑となる．さらに，依存性を表す理論分布は限りなく定義できるため，実データの解析には課題点が残る．よって，理論的な分布を与え，それに準拠した乱数を生成し，ノンパラメトリックな手法でデータ解析を行う．与えた分布の特性を乱数データの解析結果によって再現できるかどうかを判断し，実データ解析への知見を得る．

2.　2 変量極値理論に基づく数値シミュレーション

2. 1.　　同時・包括生起率

各地点 A，B における降水量などの外力レベルから求めた再現期間に相当するパラメータをデータとして扱う．これらは生起率の逆数の関係である．閾値 x_A , x_B を超える数を片側生起数として $k_A(x_A)$， $k_B(x_B)$ で表す．2 変量の場合には，(x_A,x_B) のどちらともを超える個数である同時生起数 $k_{AB}(x_A,x_B)$，どちらか片方でも超える個数である包括生起数 $k_*(x_A,x_B)$ が 1 変量に加えて定義され，これらは集合における共通部分と和集合の関係にある．**図-1** では閾値を 1 年として，$k_A(1)=2$， $k_B(1)=1$， $k_{AB}(1,1)=1$， $k_*(1,1)=2$ である．また，生起率は，生起数の期待値をとったものである．同時生起率 $\Lambda(x_A,x_B)$，包括生起率 $V(x_A,x_B)$ は，同時・包括生起数の期待値を取ったものである．

2. 2.　　閾値を超える 2 変量極値の乱数生成

一雨の降水量の再現期間を乱数生成している．(x_A,x_B) を疑似極座標変換により，(r,t) に変換することで，同時生起率 Λ の生起率密度が次の形で表される．

$$\frac{\partial \Lambda(x_A,x_B)}{\partial x_A \partial x_B} dx_A dx_B = \frac{2 dr dH(t)}{r^2} \qquad (1)$$

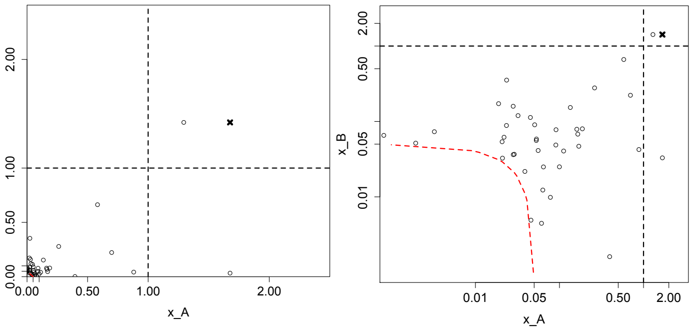


図-1　閾値モデルを模した再現期間の乱数(1 年分)
（左:ノーマルスケール　右:対数スケール）

式(1)より Λ は r と t の関数に独立な形に分解できることで容易に乱数生成が可能になる [2]．よって，(r,t) の乱数をそれぞれに生成している．1 年間に平均 20 個(2~3 週間に 1 回の頻度に相当し，2 地点で平均 40 個)の降水量が極値をとると考え，乱数を発生させる．閾値(本研究では，**図-1** の赤点線である $r = 0.05$ 年)を超えるものを極値とするものが閾値モデル(Peaks-Over-Threshold，以下，POT)である．一方，2 変量の期間最大値は期間(1 年)最大値のペア(**図-1**　×印)である Component-wise Maxima であり，データが同時極値とは限らず，閾値モデルとは極値抽出の手法が異なる．

3.　　ノンパラメトリックな手法による 2 変量極値解析

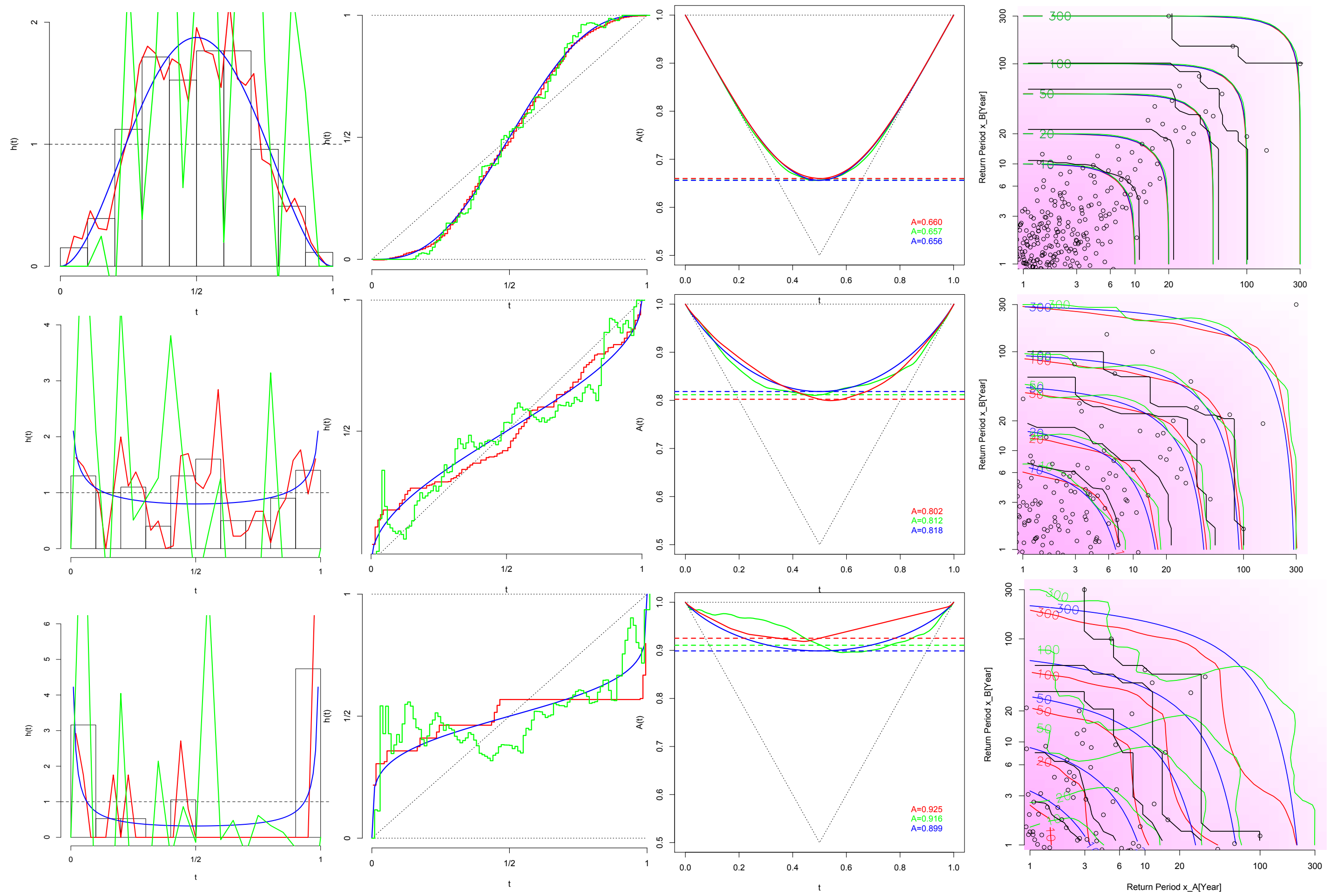
3. 1.　　ノンパラメトリックな手法を用いる理由

式(1)で生起率密度が (r,t) の関数に独立に分解できることに対して，生起率密度を積分し，包括生起率 $V(x_A,x_B)$ を得る関係は以下のようになる．

$$V(x_A,x_B) = 2\int_0^1 \left(\frac{t}{x_A} \vee \frac{1-t}{x_B} \right) dH(t) \qquad (2)$$

式(2)で定義される横断分布が定まることで，依存関数，包括生起率，同時生起率も求まる．しかし，横断分布には平均が 1/2，全確率が 1 という条件のみで，理論分布が限りなく定義できる [3]．このような状況で特定の理論分布を仮定した解析結果が最適であるか判断できない．包括生起率 V が式(2)のように表現できることからノンパラメトリックな方法が適用でき，解析手法として用いる．乱数生成のために，数式が複雑でなく理論分布の特性を柔軟に表せる 2 変量ベータ分布を与えた．

キーワード　数値シミュレーション，再現期間，ノンパラメトリック，依存関数
連絡先　　〒466-8555　愛知県名古屋市昭和区御器所町，ＴＥＬ:080-9987-1710，E-mail:28116019@stn.nitech.ac.jp



(a)横断密度 $h(t)$　(b)横断分布 $H(t)$　(c)依存関数 $h(t)$　(d)再現期間 $R_{AB}(x_A,x_B)$

図-2　数値シミュレーションによる真値(青線)と
POT による推定線(赤線)，Component-wise Maxima による推定線(緑線)の比較
上段：$beta(2,2)$(依存性：高)　中段：$beta(0.5,0.5)$(依存性：中)　下段：$beta(0.2,0.2)$(依存性：低)

3.2.　同時生起の再現期間の算出

データは前述した乱数を 300 年分生成し，1 サンプルとした．依存関数 A は包括生起率 V を個々の生起率の和で除すことで求まるという関係があり，依存関数を求めれば，同時生起の再現期間(同時生起率の逆数)が，

$$R_{AB}(x_A,x_B)=1\Bigg/\left(\frac{1}{x_A}+\frac{1}{x_B}\right)\left\{1-A\left(\frac{x_A}{x_A+x_B}\right)\right\} \tag{3}$$

で求められる [4]．ノンパラメトリックな解析により，依存関数 A を求め，依存関数を微分していくことにより，横断分布 H・横断密度 h が求められる．依存性を変えて乱数を生成し，得られた理論値になるもの(青線)と推定線(POT:赤線，Component-wise Maxima:緑線)を**図-2** に示す．

4.　まとめと今後の課題

2 地点間の依存性が高い時は，POT と Component-wise Maxima の推定結果と与えた分布は，ほぼ一致している(**図-2** 上中段)．一方，依存性が低い時(**図-2** 下段)では，Component-wise Maxima の推定結果は，与えた理論分布とのずれが大きかった．

POT が同時極値(直接)のデータから同時極値の依存関数を求めるのに対して，Component-wise Maxima は間接的なデータから同時極値の依存関数を求めているため，適合性が悪くなると考えられる．ノンパラメトリックな手法は外力の順位の情報を解析に用いているが，降水量の値が小さいデータを正確に扱えないという欠点がある．特定の理論分布を仮定することで，データ数が少ない場合の推定も可能になるパラメトリックな方法が今後の課題である．また，ほとんど独立な場合は，依存性の高いときのような解析では信頼性が低いことが分かった．このような場合の解析手法の検討も今後の課題である．

参考文献

(1) 二瓶泰雄:小田川における洪水氾濫状況(平成30年7月豪雨)，http://www.isad.or.jp/wp/wp-content/uploads/2019/06/no136_12p.pdf，2020年2月16日閲覧

(2) 北野利一：恋する極値統計，統計数理研究所　共同研究集会「極値理論の工学への応用」，課題番号2019-ISMCPR-5013，pp12-18，2019.

(3) Kotz，S.　and S.　Nadarajah: Extreme value distributions，theory and applications，Imperial College Press，pp. 112-113，2000.

(4) 北野利一・志村隆彰・田中茂信：d4PDF の多数アンサンブルを活かした極大降水量の2 地点の依存性についてのノンパラメトリック解析，土木学会論文集 B1(水工学)，第 75 巻，pp. I_289-I_294，2019.