

# ベイズ手法に基づく大量アンサンブル気候予測データを用いた 極値降雨量の信頼区間・予測区間の将来変化推定

清水啓太\*, 山田正\*\*, 山田朋人\*\*\*

## Projection for Future Change of Confidence Interval and Prediction Interval Based on Bayesian Method

Keita SHIMIZU\*, Tadashi YAMADA\*\* and Tomohito YAMADA\*\*\*

### Abstract

One of the difficulties of hydrological frequency analysis is how to manage nonstationarity of rainfall which is caused by climate change accompanied global warming. To manage nonstationarity of rainfall, using Bayesian statistics is effective. In this research, Bayesian update of extreme value distribution fitted with past observed data was conducted by using a large ensemble data base, d4PDF. d4PDF is constructed of an enormous amount of calculated values of climate models which simulated several thousand years for past climate and future climate. By using the information of d4PDF, we can estimate extremes which could occur in future time when global warming could proceed to some extent. Also, using the theory of probability limit method test to Bayesian updated extreme value distributions, both confidence intervals and prediction intervals in future climate can be constructed. Update of confidence interval and prediction interval based on Bayesian method can evaluate future change of design level rainfall and occurrence risk of catastrophic rainfalls in future climate and contribute construction of flood control management corresponding to climate change.

### 1 はじめに

従来慣用されてきた水文頻度解析手法では、過去の観測極値のみに基づき、計画確率規模に関連付けた唯一の確率雨量を算定するという、決定論的確率を扱ってきた。一方、現在利用可能な観測極値は有限であるため、これらの観測極値に基づく計画降雨量は、不確実性を大いに有する。さらに、従来の水文頻度解析は、水文量の定常性を仮定している。ここで、定常性とは、水文量を支配する確率法則（確率分布）が時間的に変化しないことである。すなわち、従来手法では、水文量が従う確率分布が時間的に変化しないものとして扱ってきた。しかしながら、近年では、地球温暖化に伴う気候変動の影響による水文量の非定常性が指摘されてきている。ここで、非定常性とは、水文量の時系列がトレンドや周期をもつことあるいはそれらが従う確率分布が時間的に変化することをいう。我が国においては、地球温暖化に伴う気候変動により、将来期間において、降雨量ならびに洪水の発生頻度が増加をするという予測が示されている [1]。上述したような気候変動に伴う降雨の激甚化を考慮すると、水文量の非定常性を考慮した統計解析手法の構築・発展は、洪水対策を立案する上で、極めて重要である。近年では、水文量の非定常性に着目した研究や報告が多数示されている。例えば、Milly et al. [2] は、気候変動の影響を踏まえ、水文分野における定常性の仮定に警鐘を鳴らし、計画と管理の在り方に対して、水文量の確率分布の非定常性を考慮すべきであると述べている。加えて、近年、我が国では、過去経験のないような規模の豪雨が多発しており、各地で人命ならびに社会基盤等に対して大規模な被害をもたらしている。このような現状を鑑みると、上述したよ

\* 中央大学大学院理工学研究科都市人間環境学専攻 〒112-8551 東京都文京区春日1-13-27

\*\* 中央大学理工学部都市環境学科 〒112-8551 東京都文京区春日1-13-27

\*\*\* 北海道大学大学院工学研究院 〒060-8628 北海道札幌市北区北十三条西8丁目

うな近年における降雨の激甚化ならびに地球温暖化に伴う豪雨の頻発傾向を考慮した洪水対策の策定は喫緊の課題である。また、同洪水対策の構築に向けて、近年では、アンサンブル気候予測データベースを用いた洪水リスク評価手法が提示されている [3] [4]。アンサンブル気候予測データベースとは、気象モデルによる過去気候および将来気候の計算結果を格納したデータベースである。このデータは、我々が過去経験する可能性のあった降雨ならびに将来経験する可能性のある降雨の観測値として捉えることができる。従来の洪水対策では、実績として存在するただ一つの観測標本に基づき計画降雨量を設定してきたが、同データを用いることで、計画降雨量の“振れ幅”を合理的に推定することができる。清水ら [5] は、従来手法において想定外として扱われるような激甚豪雨に対応するために、確率雨量が有する不確実性を信頼区間という形で定量化し、その幅の中の雨量が生起するリスクを議論するという、確率論的確率を導入した水文頻度解析手法を提案している。当該手法は、アンサンブル気候予測データを用いた物理的モンテカルロ法あるいは数理統計理論により構築できる。清水ら [5] は、数理統計理論である確率限界法検定 [6] に基づく水文頻度解析手法を構築した。その手法論を構築するにあたり、重要となる理論の一つは、過去存在しえた観測値の振れ幅の推定を可能とする Kolmogorov-Smirnov 検定である。一方、Kolmogorov-Smirnov 検定は、洪水対策上重要となる、確率分布の裾部に相当する激甚豪雨に対する推定精度が弱い。清水ら [5] は、Kolmogorov-Smirnov 検定が有するこの弱点を改善した、確率限界法検定を採用し、当該検定に基づく信頼区間・予測区間を構築した。なお、予測区間とは、将来観測されうる雨量がとり得る値の幅を表すものである。確率限界法検定に基づく信頼区間・予測区間の導入により、信頼区間・予測区間をどこまで議論するかによって洪水対策におけるリスクが表現可能となる。しかし、当該手法は、時間軸上のある一点に立脚しており、将来の観測値が有する情報を活かしていない。一方、ベイズ統計理論を確率限界法検定に基づく水文頻度解析手法に導入することで、新たに取得した観測値の情報を取り込み、より合理的な信頼区間・予測区間が算定可能となる。このため、同理論を用いることで、アンサンブル気候予測データベースの将来気候予測実験結果を取り込み、採用確率分布とその信頼区間・予測区間を更新することができる。本研究では、マルコフ連鎖モンテカルロ法 (Markov Chain Monte Carlo method, MCMC 法) を用いて、過去数十年間分の観測実績から算定した採用確率分布とその信頼区間・予測区間に、将来気候予測実験結果を取り込み、それらを更新する手法を構築した。同手法の適用は、地球温暖化時における確率雨量の評価・検討の実現可能性を提示するものである。なお、本研究では、安倍川流域牛妻観測所における 1960 年から 2003 年に渡る 44 年間分の年最大日降水量の観測値を解析用データとしている。以降での統計解析では、主に Gumbel 分布を用いているため、式(1)に Gumbel 分布の累積分布関数  $F_X(x)$  を示しておく。

$$F_X(x) = \exp\left[-\exp\left\{-\left(\frac{x-\mu}{\sigma}\right)\right\}\right] \quad (1)$$

ここに、 $\mu$ : 位置母数,  $\sigma$ : 尺度母数である。

## 2 確率限界法検定を用いた信頼区間・予測区間の構成手法

以降では、実績の観測値を表す確率変数を  $X$ ,  $X$  の累積分布関数を  $F_X(x)$ ,  $X$  の累積分布関数の逆関数である確率表現関数 [7] を  $\chi_X(u)$  と表記する。なお、本論では、確率限界法検定を「当該検定」、確率限界法検定における棄却限界値を「確率限界値」とも称する。

任意の連続確率分布の累積確率  $U$  は、標準一様分布 (区間  $[0,1]$  の一様分布) に従い、また、標準一様分布からの順序統計量  $U_{(i)}$  は、母数  $(i, n-i+1)$  のベータ分布に従う。このことを式(2)により表す。

$$F_{U_{(i)}}(u) = I_u(i, n-i+1) \quad (2)$$

ここに、 $F_{U_{(i)}}(u)$ : 第  $i$  順序統計量  $U_{(i)}$  の累積分布関数,  $I_u(i, n-i+1)$ : 母数  $(i, n-i+1)$  のベータ分布の累積分布関数,  $n$ : 標本サイズ (使用する観測値の総数),  $i$ : 観測値を小さい順に並び替えた場合における小さい方からの順位である。

標準一様分布における確率限界値を算出するためには、式(3)に示す試行により、“標準一様分布から大きく外れるような一様乱数値の生起確率  $\alpha_{\min}$ ” を算出する。

$$\alpha_{\min} = \min_{1 \leq i \leq n} \left\{ \text{Min} \left[ I_u(i, n-i+1), I_{1-u}(n-i+1, i) \right]_{u=U_{(i)}} \right\} \quad (3)$$

ここに、 $L_u(i, n-i+1)|_{u=u(i)}$ ：第*i*順序統計量 $u(i)$ の非超過確率、 $I_{1-u}(n-i+1, i)|_{u=u(i)}$ ：第*i*順序統計量 $u(i)$ の超過確率である。

式(3)の試行を*N*回繰り返すことで、*N*個の $\alpha_{\min}$ が得られる。次に、これら*N*個の $\alpha_{\min}$ を $-\log_{10}(2\alpha_{\min})$ に対数変換する。この $-\log_{10}(2\alpha_{\min})$ の分布は、標準一様分布から抽出した標本中において、母数 $(i, n-i+1)$ のベータ分布の裾部に相当する値の分布を表している。なお、 $-\log_{10}(2\alpha_{\min})$ は、当該標本中の最大値となるため、極値分布への当てはめが妥当である。Fig-1に、試行回数*N*を50000とした場合における $-\log_{10}(2\alpha_{\min})$ の標本に当てはめたGumbel分布の確率表現関数 $\chi_\alpha(u)$ の一例を示す。例えば、有意水準*p*% (両側確率*p*%) で標本サイズ*n*の観測標本 $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ とこれらの観測値が従うと仮定した累積分布関数 $F_X(x)$ に対する当該検定を行う場合、この有意水準 (両側確率) に対応する $\alpha$ は、方程式 $\chi_\alpha(1-p) = -\log_{10}(2\alpha_{\min})$ を満たす $\alpha_{\min}$ である。式(2)に示す確率が $\alpha$ となる $u$ が $z_L(i)$ 、当該確率が $1-\alpha$ となる $u$ が $z_U(i)$ である。すなわち、標準一様分布の上方確率限界値の標本 $z_U$ は $\{z_U(1), z_U(2), \dots, z_U(n)\}$ 、下方確率限界値の標本 $z_L$ は $\{z_L(1), z_L(2), \dots, z_L(n)\}$ となる。ここで、Fig-2は両側確率5%、観測値の総数*n*が44の場合での確率限界法検定に基づく信頼区間の構成過程を示したものである。同図には、標準一様分布における確率限界値 (緑点)、解析用データ (黒点)、解析用データに当てはめたGumbel分布 (青実線) および確率限界法検定に基づく当該Gumbel分布の95%信頼区間 (青破線) が示されている。

標準一様分布の確率限界値から構成される幅 $[z_L(i), z_U(i)]$ は、*i*番目の順序統計量 $X_{(i)}$ を生じさせる累積確率 $u$ がとり得る幅を表している。このため、Fig-2が示すように、観測値が従うと仮定した確率分布における上方確率限界値 $\chi_X(z_U(i))$ と下方確率限界値 $\chi_X(z_L(i))$ は、*i*番目の累積確率である $U_{(i)}$ が有意水準*p*の下でとり得る幅 $[z_L(i), z_U(i)]$ を観測実績から算定した確率表現関数 $\chi_X(u)$ に代入することで得られる。このため、区間 $[\chi_X(z_L), \chi_X(z_U)]$ は、観測値自体がとり得る幅を表し、この幅は、過去存在しえた極値がとり得る幅として捉えることができる。このため、任意の有意水準の下で決定する、観測値自体がとりうるの限界値 (確率限界値) の標本に、採用確率分布と同じ関数形をあてはめることで、採用確率分布自体の限界曲線が得られる。したがって、本研究では、信頼限界線は、確率限界値の標本に採用確率分布と同じ関数形で当てはめることで得られると仮定している。また、Fig-2から、採用確率分布であるGumbel分布を確率限界値の標本 (薄青点) に当てはめることで、信頼限界線 (青破線) が構成されていることが分かる。

次に、確率限界法検定に基づく予測区間の構成方法を述べる。確率限界法検定の採択域は、我々が過去経験しえた大雨の振れ幅を高精度で推定するものであるため、確率限界値を大雨の振れ幅の限界値として捉えることができる。したがって、確率限界値の標本に対して、高い適合度を有する確率分布モデルを選定することにより、当該確率分布モデルは、データの振れ幅の限界値を与えるものとして捉えることができる。加えて、当該確率分布モデルにより、使用したデータの外部に対しても、外挿値という形で値を得ることができる。このことは、将来期間に対しても、データの振れ幅の限界値を予測していることを表している。したがって、本研究では、当該確率分布モデルを予測限界線とし、両限界線から成る幅が予測区間である。

## 2. 1 確率限界法検定に基づく信頼区間と統計的モンテカルロ法により算定した信頼区間の適合度評価

確率限界法検定による信頼区間の精度を検証するために、当該信頼区間と統計的モンテカルロ法により構成した信頼区間との比較を行った。なお、統計的モンテカルロ法による信頼区間は、仮定した確率分布に従う乱数を観測値の総数と同じ数だけサンプリングし、これらの乱数から構成される“統計的アンサンブル標本”に当てはめた確率分布が与えるクオンタイル値 (確率水文量) の分布から求めることができる。具体例を挙げると、同クオンタイル値の分布の2.5%ile値と97.5%ile値から成る区間が統計的モンテカルロ法による95%信頼区間である。精度検証にあたり、実施した統計的モンテカルロ法の手順を次に示す。

(手順1)：利根川水系八斗島観測所における54年間分の年最大日降水量の観測値に当てはめたGumbel分布を求める。なお、本節で扱うデータ (利根川水系八斗島観測所における54年間分の年最大日降水量の観測値) は、上述した解析用データとは異なる点に注意されたい。

(手順2)：(手順1)において、観測値に当てはめたGumbel分布に従う乱数を54個 (実際の観測値と同数) だけ発生させ、これら54個の乱数に当てはめたGumbel分布を求める。

(手順3) : (手順2)を5000回だけ繰り返す。この操作により、統計的アンサンブル標本に当てはめたGumbel分布を5000組得る。ここで、Fig-3は以上の手順で示した諸量を記入した確率紙を示したものである。

(手順4) : 統計的アンサンブル標本に当てはめたGumbel分布が与えるクオンタイル値を用いて、対象とする確率年における確率年最大日降水量の頻度分布を得る。

(手順5) : 確率年最大日降水量の分布が信頼区間に含まれる割合を算出する。なお、以降ではこの割合を被覆率と表記する。

本研究では、確率年を100年、200年、500年、1000年、5000年および10000年とした場合における被覆率を算出した。Fig-4は信頼係数と被覆率の関係を確率年別示したものである。この図から、任意の信頼係数に着目し、被覆率を比較すると、いずれの場合も、数%程度の変化が存在することが分かる。この結果は、本研究が提案する信頼区間の精度が確率年にほとんど依存しないことを示している。Fig-5は100年確率年最大日降水量の頻度分布と信頼区間の関係を示したものである。同図から、信頼限界値は信頼係数 $p$  ( $0 < p < 1$ )を大きくするほど、確率年最大日降水量の頻度分布の100  $((1-p)/2)$ %ile値および100  $(1-(1-p)/2)$ %ile値に漸近するという特徴をもつことが分かる。ここで、Fig-4より信頼区間の精度は確率年にほとんど依存しないことが示されており、この特徴は任意の確率年において成立することが分かる。以上の結果は、信頼係数 $p$ が0.9以上の場合、確率限界法検定による信頼区間と統計的モンテカルロ法による信頼区間はほとんど一致することを示すものである。このため、統計的モンテカルロ法による確率雨量の頻度分布の形状を定量化せずとも、100%上側信頼限界線上の確率水文量をもつ超過確率を $(1-p)/2$ として近似可能であることが示された。これは、確率限界法検定の理論および提案手法の特徴であると言える。

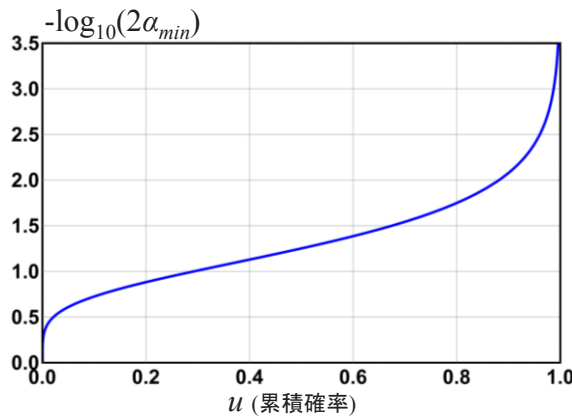


Fig-1 Probability representing function of Gumbel distribution fitted with a sample of  $-\log_{10}(2\alpha_{min})$  (Where, N is 50000.) ※In this research, to preserve the stability of probability distribution fitted with a sample of  $-\log_{10}(2\alpha_{min})$ , the number of trials is set as 50000.

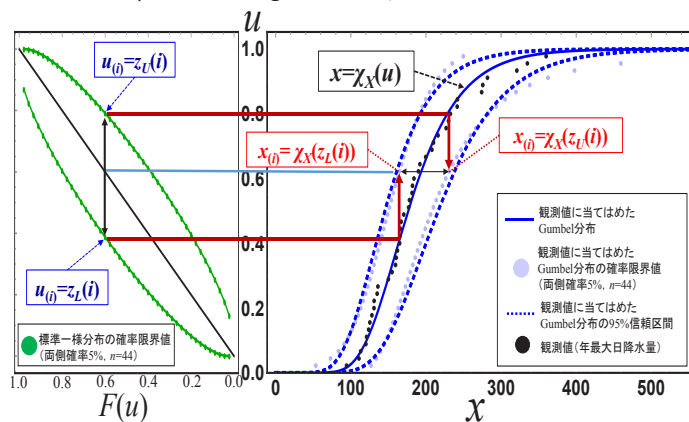


Fig-2 The construction process of 95% confidence interval of Gumbel distribution fitted with observed data using probability limit method test  
 ※ $x$  is the realized values of probability variable  $X$  which expressed observed data.



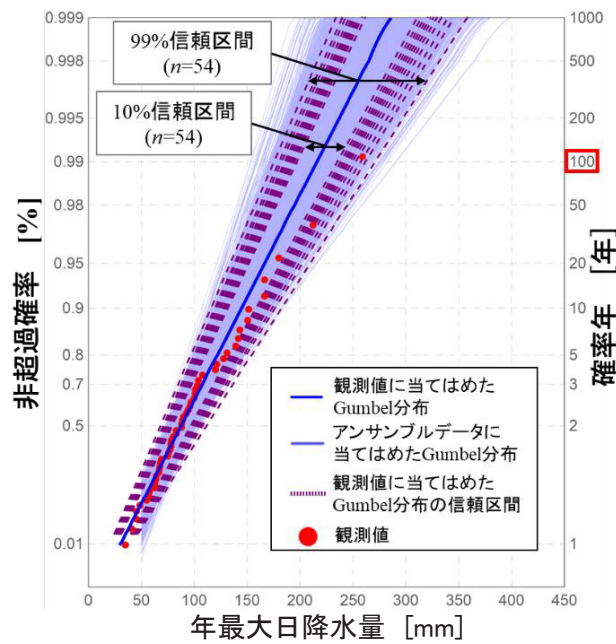


Fig-3 Observed annual maximum daily rainfall data for 54 years at Yattajima observatory, Gumbel distribution fitted with these observed data and Gumbel distributions fitted with statistical ensemble data. Where, n in Fig-3 means total number of observed data.

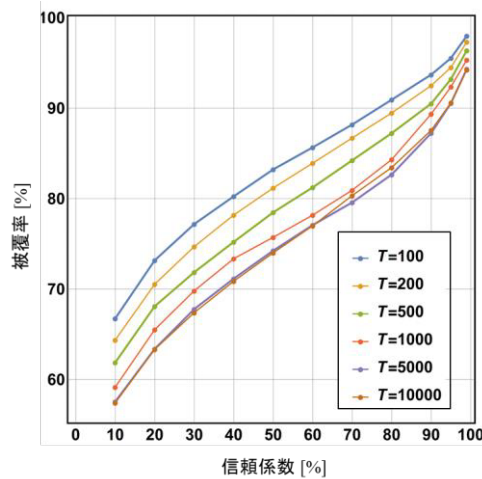


Fig-4 Relation between confidence coefficient and coverage probability. T means return period.

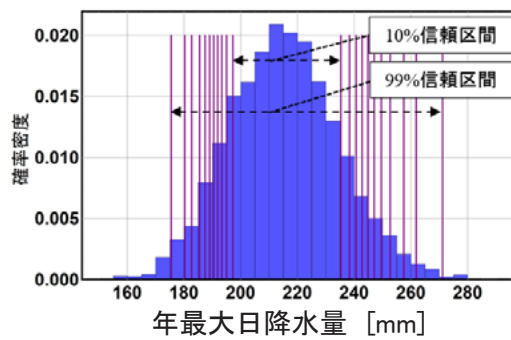


Fig-5 Relation between frequency distribution of 100-year annual maximum daily rainfall and confidence interval based on probability limit method test.

In Fig-5, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% and 99% confidence intervals are shown.

### 3 マルコフ連鎖モンテカルロ法を用いた信頼区間・予測区間の更新

本章では、マルコフ連鎖モンテカルロ法を用いることで、観測実績に基づき算定した極値分布に将来実験結果を取り込ませ、将来気候での極値分布およびその信頼区間を構成する手法を示す。

母数の事後分布は、式(4)に示すベイズの定理より求めることができる。

$$f(\theta|x) = \frac{f(x|\theta)f(\theta)}{\int_{\Theta} f(x|\theta)f(\theta)d\theta} \propto f(x|\theta)f(\theta) \quad (4)$$

ここに、 $\theta$ : 母数、 $f(\theta)$ :  $\theta$ の事前分布、 $f(x|\theta)$ : 尤度関数、 $f(\theta|x)$ :  $\theta$ の事後分布である。

式(4)は、観測値を新たに取得するごとに母数に関する情報量が増加し、母数に関する事前情報を表す事前分布 $f(\theta)$ が事後分布 $f(\theta|x)$ に更新されることを示している。一般的に、極値分布の母数の事後分布は、式(4)に示すベイズの定理を用いて解析的に求めることが困難である。そのため、本研究では、極値分布の母数の事後分布を数値的に推定可能なマルコフ連鎖モンテカルロ法を採用した。また、MCMC法の一手法であるメトロポリス法を用いて、最尤法により解析用データに当てはめた Gumbel 分布 (Fig-2 の青実線) の母数の事後分布を求めた。

#### 3.1 メトロポリス法の概要

以下にメトロポリス法 [8] の手順を示す。

(手順1):  $\theta$ の初期値を設定する。

(手順2): 母数の候補を与える提案分布 $q(\theta')$ に従って、 $\theta'$ を独立に発生させる。

(手順3): 確率 $\beta$ を式(5)より計算する。

$$\beta = \text{Min} \left\{ \frac{f(\theta'|x)}{f(\theta|x)}, 1 \right\} \quad (5)$$

(手順4): 確率 $\beta$ で $\theta$ を $\theta'$ に推移させて、 $\theta=\theta'$ と更新する。具体的には、確率 $1-\beta$ で元の $\theta$ を採用し、推移を行わない。

(手順5): (手順2): ~ (手順4)を繰り返す。

(手順6):  $\theta$ のサンプル列が収束した場合、反復計算を打ち切る。

#### 3.2 使用データの概要

本研究では、「地球温暖化に資するアンサンブル気候予測データベース、database for Policy Decision making for Future climate change (d4PDF)」[9]の領域実験を使用した。d4PDF領域実験は領域気候モデルにより作成された水平解像度20kmの気候予測データであり、長期観測データが得られる1951年から2010年の60年間を対象に異なる初期値、海氷と海面水温に摂動を加えた50メンバの数値計算である過去実験(60年×50メンバ=3000年)、産業革命以前に比べて全球平均気温が4°C上昇した世界を想定し、6種類の海面水温パターンとそれに摂動を与えた15メンバの数値計算を実施した4°C上昇実験(60年×海面水温6パターン×15メンバ=5400年)で構成される。本研究で使用したd4PDFデータは、安倍川流域牛妻観測所に相当するd4PDFのモデル格子における3000年間分の過去実験結果の年最大日降水量および5400年間分の+4°C将来実験結果での年最大日降水量である。なお、上記の6種類の海面水温パターンとは、第5次結合モデル相互比較計画(CMIP5)における6種類の主要モデルから計算されたものである。また、本論は、観測実績を過去気候における降水量の母集団から実現した標本として扱う。

#### 3.3 メトロポリス法の実施結果

メトロポリス法における母数の初期値は、解析用データに当てはめた Gumbel 分布の最尤推定値として、上記手順に基づきメトロポリス法を実施すると、Fig-6に示す当該 Gumbel 分布の事後分布が有する母数のサンプル列が得られた。なお、本研究は、提案分布として指数分布を採用している。Fig-6から、最尤推定値を初期値として選定することで、初期値近傍では顕著な振動が見られるものの、母数のサンプル列は安

定しており、計算上捨てるべき助走区間（バーンイン区間）がほとんどないことが分かる。このため、MCMC法による事後分布の母数のサンプル値群を用いた検討の妥当性が確認された。

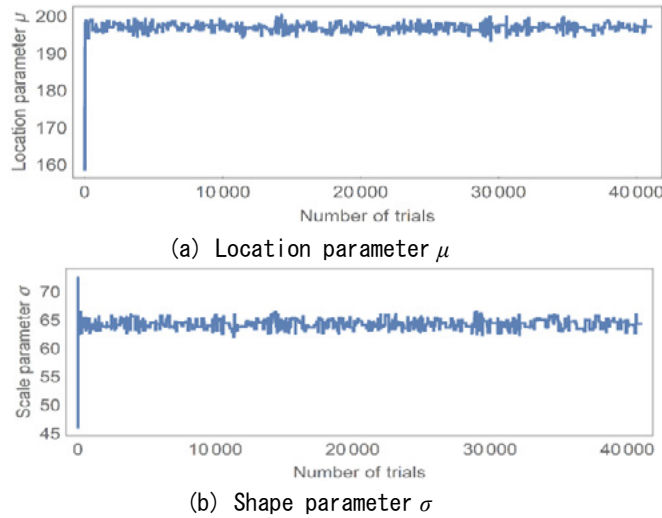


Fig-6 Sample series of parameters of posterior distribution of Gumbel distribution fitted with targeted data generated by metropolis method.

### 3. 4 将来気候における極値の予測分布の構成

極値の予測分布は、式(6)により得られる。このとき、将来実験結果を尤度関数に与えたMCMC法により算定される極値の予測分布は、将来気候における極値分布として扱える。

$$P(Y \leq y | \mathbf{x}) = \int_{\theta} F(Y \leq y | \theta) f(\theta | \mathbf{x}) d\theta \quad (6)$$

ここに、 $Y$ ：将来の極値を表す確率変数、 $P(Y \leq y | \mathbf{x})$ ：極値の予測分布の累積分布関数、 $F(Y \leq y | \theta)$ ：母数  $\theta$  に対応する  $Y$  の累積分布関数、 $f(\theta | \mathbf{x})$ ：観測値群から算定した母数の事後密度、 $\chi_Y(u)$ ：極値の予測分布の確率表現関数である。また、母数  $\theta$  をMCMC法の活用により生成可能である母数のサンプル値群とすると、極値の予測分布を式(7)により表すことができる [10]。

$$P(Y \leq y | \mathbf{x}) \approx \frac{1}{s} \sum_{i=1}^s F_Y(y | \theta_i) \quad (7)$$

$F_Y(y)$ は将来観測値の母分布として仮定した確率分布である。式(7)は、MCMC法により生成された母数のサンプル値群  $\theta_i$  を  $F_Y(y | \theta_i)$  に代入し算出される将来観測値  $y$  の累積確率の合計を、採用したサンプル値群  $\theta$  の総数  $s$  で平均することで、極値の予測分布の累積分布関数  $P(Y \leq y | \mathbf{x})$  が算定可能であることを示している。また、 $s$  とは、MCMC法により生成した母数のサンプル値群の総数から、計算上捨てるべき助走区間（バーンイン区間）分に相当するサンプル値群の総数を除いた値である。本研究では、母分布  $F_Y(y)$  にGumbel分布を採用している。また、MCMC法の反復計算回数は、41000であり、このうち1000をバーンイン区間の総数としているため、 $s$  は41000から1000を引いた値である40000とした。

MCMC法によるベイズ更新後の確率分布（実績観測値に当てはめた確率分布の予測分布）に対して、上述した、確率限界法検定による信頼区間・予測区間の構成手法を適用することで、当該分布の信頼区間・予測区間が得られる。Fig-7は、両側確率5%、観測値の総数  $n$  が44の場合での当該検定を用いて、将来実験結果をMCMC法により取り込み更新されたGumbel分布の95%信頼区間を構成する過程を示している。同図には、標準一様分布における確率限界値（緑点）、将来実験結果にMCMC法を適用しベイズ更新されたGumbel分布（赤実線）、および確率限界法検定に基づく当該Gumbel分布の95%信頼区間（赤破線）

が示されている。なお、標準一様分布の確率限界値は実績の観測値の総数に応じて決定するため、観測実績（過去気候）と将来気候の標準一様分布における確率限界値は、同一である。

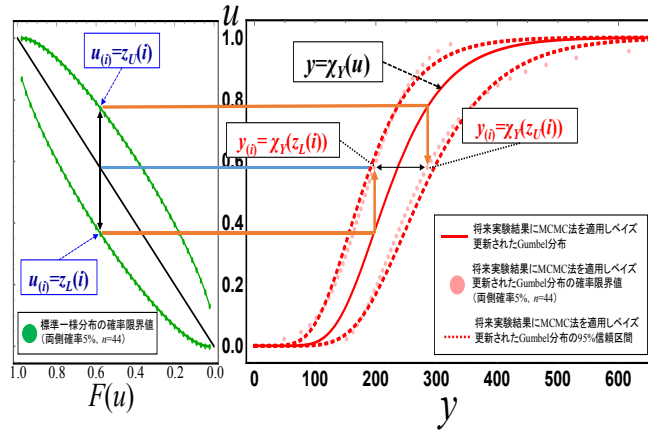


Fig-7 Construction process of 95% confidence interval based on probability limit method in future climate

Where, “y” is the probability variable which expresses future extremes.

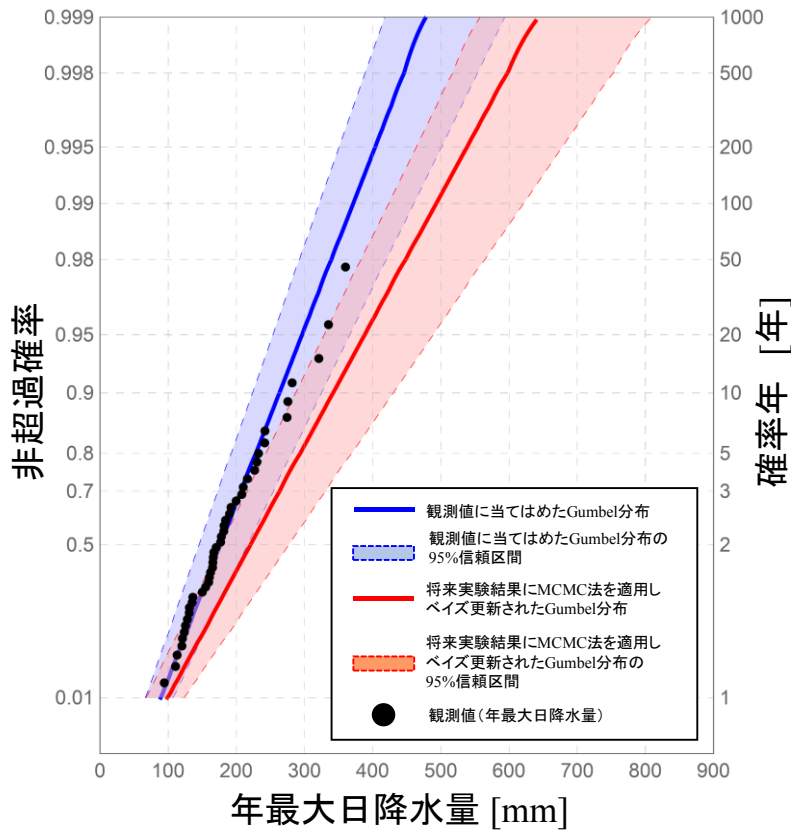


Fig-8 Update of confidence interval using Bayesian method which assimilates ensemble climate projection database

#### 4 ベイズ手法の導入による信頼区間の更新

Fig-8 は、解析用データ（黒点）、解析用データに当てはめた Gumbel 分布（青実線）および確率限界法検定に基づく当該 Gumbel 分布の 95%信頼区間（青色で表示された範囲）と MCMC 法に基づき解析用デー



タに当てはめた Gumbel 分布に将来実験結果を取り入れベイズ更新された Gumbel 分布（解析用データに当てはめた Gumbel 分布の予測分布，赤実線），確率限界法検定に基づくベイズ更新された当該 Gumbel 分布の 95%信頼区間（赤色で表示された範囲）である。過去気候における信頼区間と将来気候における信頼区間が重複する範囲より，将来気候における年最大日降水量は過去気候においても発生する可能性があったことが分かる。また，同図より，将来気候では，過去気候に対して信頼区間が広がり，確率雨量が増加することが確認された。さらに，計画規模に相当する 100 年確率に着目すると，対象地点において，将来気候の 100 年確率年最大日降水量（493.3mm）は，観測実績に基づく過去気候の 100 年確率年最大日降水量（370.9mm）の約 1.33 倍の値となる。また，将来気候の 100 年確率年最大日降水量の 95%上側信頼限界値（621.4mm）は，観測実績に基づく過去気候の 100 年確率年最大日降水量の 95%上側信頼限界値（460.3mm）の約 1.35 倍の値となる。

統計理論であるベイズ手法により，従来採用してきた確率分布に気候変動予測結果を取り込み，信頼区間を更新することは，気候変動下での確率雨量の検討に大いに資するものであると期待される。

#### 5 確率限界法検定に基づく信頼区間の物理的モンテカルロ法に基づく信頼区間に対する被覆率

過去実験結果・将来実験結果における確率雨量の振れ幅や度数分布をより鮮明化し，当該検定に基づく信頼区間との比較を実施するために，両実験結果にリサンプリングを行った。具体的には，過去実験 3000 年間・将来実験 5400 年間分の年最大日降水量を実績観測値の総数である 44 と同じ数だけリサンプリングし，5000 個のリサンプリング標本を作成した。このため，各リサンプリング標本の標本サイズは，観測実績の総数と同じ 44 である。次に，各リサンプリング標本に Gumbel 分布を当てはめ，確率評価を実施した。

Fig-9 は，解析用データに当てはめた Gumbel 分布（青実線），確率限界法検定に基づく当該 Gumbel 分布の 95%信頼区間および過去実験結果のリサンプリング標本に当てはめた Gumbel 分布（緑細実線，5000 個）を示している。また，Fig-10 は，将来気候での Gumbel 分布（解析用データに当てはめた Gumbel 分布の予測分布，赤実線），確率限界法検定に基づく当該 Gumbel 分布の 95%信頼区間（赤破線）および将来実験結果のリサンプリング標本に当てはめた Gumbel 分布（橙細実線，5000 個）を示している。Fig-9 と Fig-10 から，リサンプリングを行うことで，過去実験と将来実験における確率雨量の振れ幅がより鮮明となり，その振れ幅は確率限界法検定に基づく信頼区間と極めて良く適合していることが分かる。特に，Fig-9 に示す信頼区間（青破線）は，d4PDF の過去実験結果を MCMC 法により取り込まず，数理統計理論である確率限界法検定のみを用いて導出したものであり，当該検定に基づく信頼区間が物理的モンテカルロ法による確率雨量の振れ幅（信頼区間）に極めて良く適合することは，当該検定の有用性を大いに示す結果である。次に，当該検定に基づく信頼区間と物理的モンテカルロ法による信頼区間の適合度を被覆率の観点から評価する。被覆率とは，対象推定量が信頼区間に含まれる割合である。本研究で扱う被覆率は，確率雨量が信頼区間に含まれる割合である。被覆率を算出することで，信頼区間の精度を評価することができる。すなわち，信頼区間は，設定した信頼水準（信頼係数）の下で，その被覆率が合理的なものであり，できるだけ幅の狭いものが望ましい。例えば，95%信頼区間を構成した際，その被覆率が 95%に近い値の場合，当該信頼区間の精度は良いとされる。

Fig-11 は，リサンプリングを実施した過去実験における 100 年確率年最大日降水量の頻度分布と過去気候における当該検定に基づく 95%信頼区間の関係，Fig-12 は，リサンプリングを実施した将来実験における 100 年確率年最大日降水量の頻度分布と将来気候における当該検定に基づく 95%信頼区間の関係を示している。Fig-11 で示している当該検定に基づく 95%信頼区間（青破線）の被覆率は 92.7%であり Fig-12 で示している当該検定に基づく 95%信頼区間（赤破線）の被覆率は 94.6%である。将来気候での検討結果については，ベイズ手法によりその将来気候変動予測実験結果を反映させた信頼区間を用いているため，当該信頼区間の適合度ならびに被覆率は，過去気候におけるそれらよりも高くなっている。また，100 年確率以外を対象とした検討においても，95%信頼区間の被覆率は上記のものと同程度のものとなり，被覆率の観点からも同検定に基づく信頼区間が物理的モンテカルロ法による確率雨量の頻度分布を高精度で表現可能であることが示された。

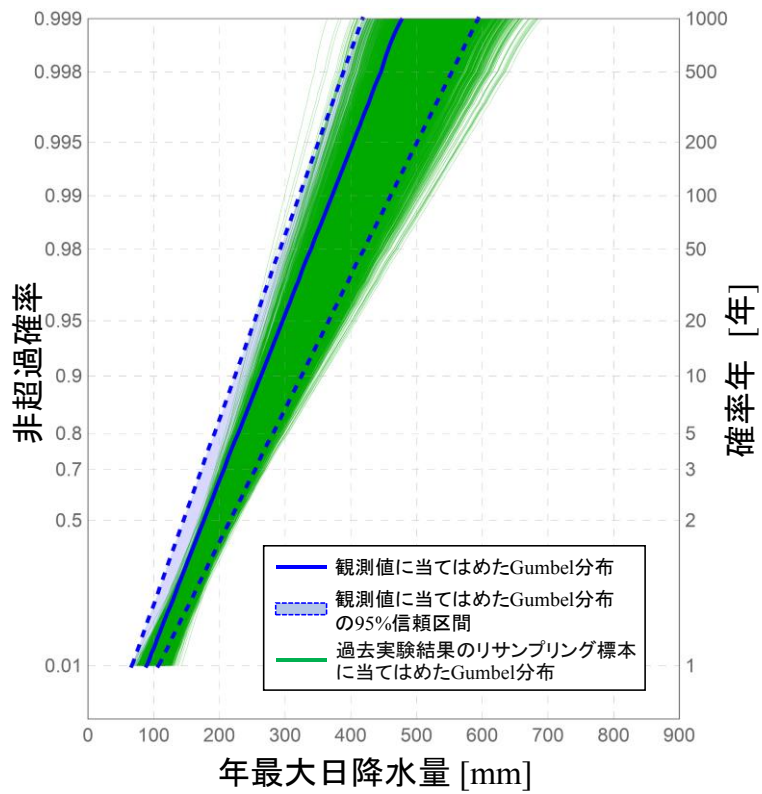


Fig-9 95% confidence interval based on probability limit method test in past climate and interval of T-year annual maximum daily precipitation generated by resampling method

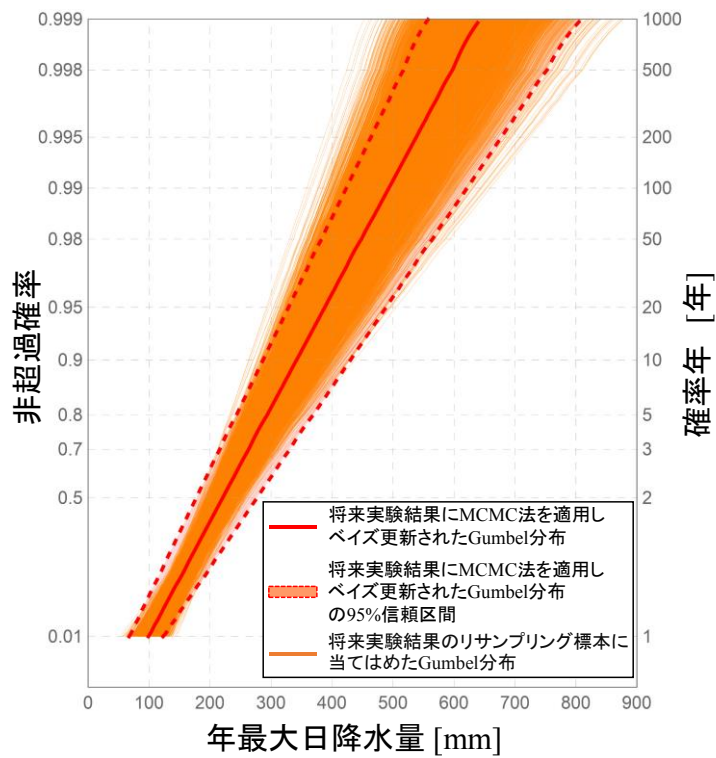


Fig-10 95% confidence interval based on probability limit method test in future climate and interval of T-year annual maximum daily precipitation generated by resampling method

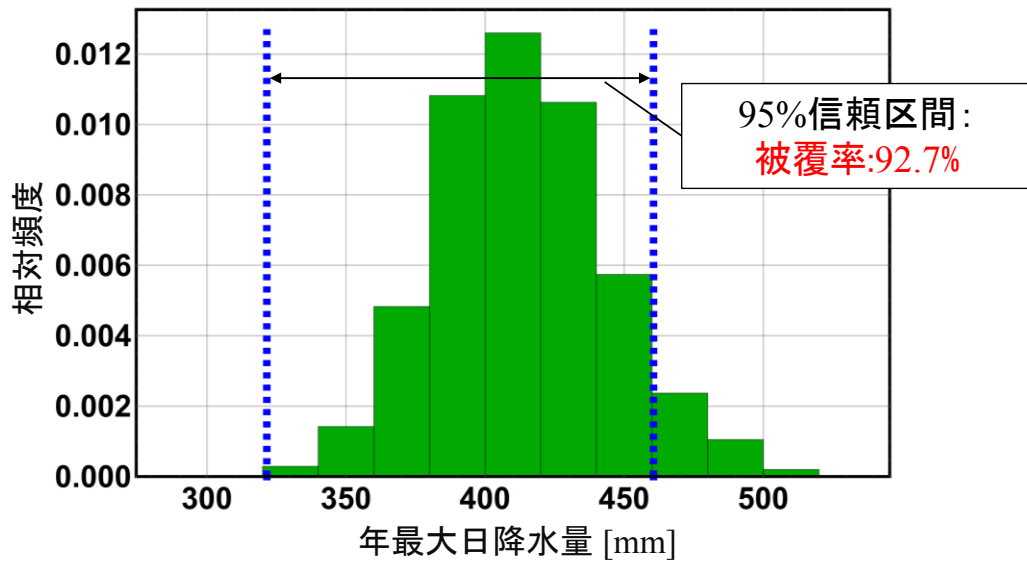


Fig-11 The Relation between 95% confidence interval based on probability limit method test in past climate and interval of 100-year annual maximum daily precipitation generated by resampling method

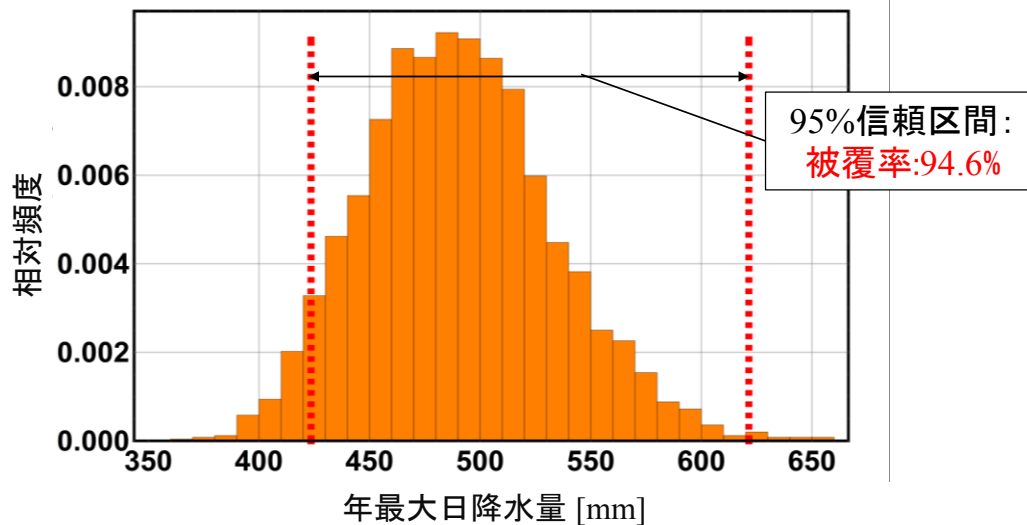


Fig-12 The Relation between 95% confidence interval based on probability limit method test in future climate and interval of 100-year annual maximum daily precipitation generated by resampling method

## 6 ベイズ手法の導入による予測区間の更新

予測区間を構成するには、確率限界値の標本に最良の適合度を有する確率分布を採用する必要があることは既述した通りである。そのため、確率限界値と過去実験結果ならびに将来実験結果を分析することで、両者の適合度を見る。すなわち、確率限界値の両側から成る確率雨量の幅（採択域）と過去・将来実験結果

から得られる確率雨量の幅との間が良く合うのであれば、本研究が提案する予測区間の妥当性（精度の良さ）が物理的モンテカルロ法による数値実験結果の観点からも担保されることとなる。

まず、観測実績から算定した確率限界値と過去実験結果の比較を確率紙へのプロットに対する目視により行う。Fig-13は、解析用データに当てはめた Gumbel 分布（青実線）、確率限界法検定による当該 Gumbel 分布における両側確率 5%の確率限界値（青点）および過去実験結果のリサンプリング標本の経験分布（緑点、500 例を図示）を示している。また、Fig-14は、将来気候での Gumbel 分布（解析用データに当てはめた Gumbel 分布の予測分布、赤実線）、確率限界法検定による当該 Gumbel 分布における両側確率 5%の確率限界値（青点）および将来実験結果のリサンプリング標本の経験分布（赤褐点、500 例を図示）を示している。Fig-13より、過去実験より算定される確率雨量の振れ幅は、両側確率 5%の理論的に算定される確率限界法検定の採択域にほとんど含まれることが分かる。さらに、Fig-14より、将来実験結果の最上位の値は、採択域の上限を大きく外れる場合があることが多いことを示唆している。なお、Fig-14では、図示されている 500 例の最上位値の 3 つが採択域の上限を超えた値となっている。将来実験では、過去実験に対して、最上位の値の振れ幅が大きくなることを示唆される一方で、将来実験より算定される確率雨量の振れ幅は、両側確率 5%の理論的に算定される確率限界法検定の採択域に概ね一致することが分かる。このように、目視により、物理的モンテカルロ法による順序統計量の頻度分布と確率限界法検定の採択域が同程度の幅を取ることを確認した。次は、より具体的に両者の適合度を明らかとする。すなわち、両側確率 5%で構成した採択域の物理的モンテカルロ法による順序統計量の頻度分布に対する被覆率を算出する。両側確率 5%で構成した採択域の物理的モンテカルロ法による順序統計量の頻度分布に対する被覆率が 95%に近いほど、両者の適合度は高いということを意味している。Fig-15は、確率限界法検定による解析用データに当てはめた Gumbel 分布における両側確率 5%の採択域と過去実験結果から抽出した 5000 個のリサンプリング標本から算定した第 44 順序統計量の頻度分布である。このとき、採択域の被覆率は 97.2%となるため、被覆率の観点からも確率限界法検定の有用性が確認された。また、Fig-16は、確率限界法検定による MCMC 法に基づき解析用データに当てはめた Gumbel 分布に将来実験結果を取り入れベイズ更新された Gumbel 分布における両側確率 5%の採択域と将来実験結果から抽出した 5000 個のリサンプリング標本から算定した第 44 順序統計量の頻度分布である。このとき、採択域の被覆率は 97.4%となるため、ベイズ手法の適用により、将来気候においても確率限界法検定による手法の有用性が確認された。よって、信頼係数 95%（両側確率 5%）で算定した採択域は、物理的モンテカルロ法による確率雨量の分布を高精度で推定できるため、採択域の上限値群と下限値群をそれぞれ最良の適合度を有する確率分布で外挿することで将来期間の大雨の振れ幅（95%予測区間）を算定可能であることが分かる。なお、本研究では、Fig-13およびFig-14に示されている両側確率 5%の確率限界値群に対して、それぞれ一般化極値分布を当てはめることで、過去気候および将来気候における 95%予測区間を構成した。ここで、一般化極値分布の累積分布関数の関数形  $F_X(x)$  を式(8)に示す。

$$F_X(x) = \exp \left[ - \left\{ 1 + \frac{\xi(x-\mu)}{\sigma} \right\}^{-\frac{1}{\xi}} \right] \quad \left( \xi \neq 0, 1 + \frac{\xi(x-\mu)}{\sigma} > 0 \right) \quad (8)$$

ここに、 $\mu$ : 位置母数,  $\sigma$ : 尺度母数,  $\xi$ : 形状母数である。

次に予測区間の将来変化について述べる。Fig-17は、解析用データ（黒点）、解析用データに当てはめた Gumbel 分布（青実線）および確率限界法検定に基づく当該 Gumbel 分布における 95%予測区間（青色で表示された範囲）と MCMC 法に基づき解析用データに当てはめた Gumbel 分布に将来実験結果を取り入れベイズ更新された Gumbel 分布（解析用データに当てはめた Gumbel 分布の予測分布、赤実線）、確率限界法検定に基づくベイズ更新された当該 Gumbel 分布における 95%予測区間（赤色で表示された範囲）である。過去気候における予測区間と将来気候における予測区間が重複する範囲より、将来気候における年最大日降水量は過去気候においても発生する可能性があったことが分かる。また、同図より、将来気候では、過去気候に対して予測区間が広がり、地球温暖化進行時の年最大日降水量が増加することが確認された。さらに、計画規模に相当する 100 年確率に着目すると、対象地点において、将来気候の 100 年確率年最大日降水量の 95%上側予測限界値（883.8mm）は、観測実績に基づく過去気候の 100 年確率年最大日降水量の

95%上側予測限界値 (638.0mm) の約 1.39 倍の値となる。このように、予測区間の更新により、地球温暖化進行時における極値降雨量の規模ならびにその生起リスクが評価可能となる。

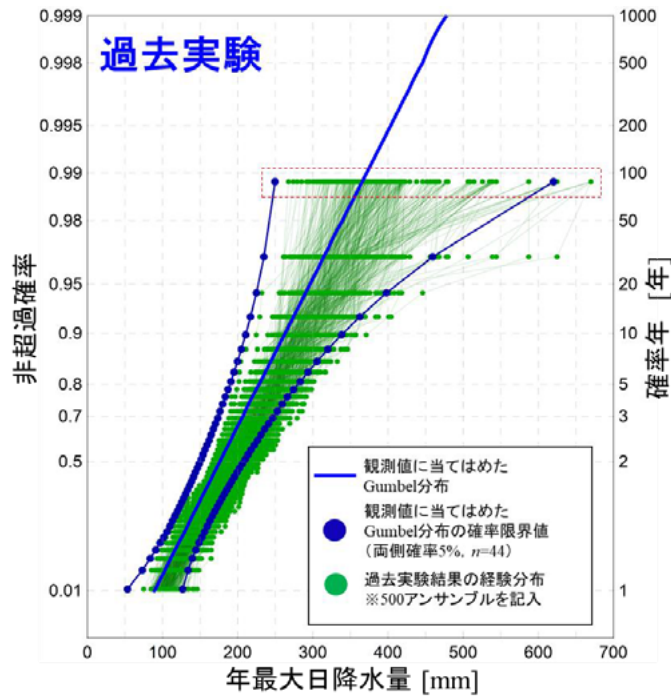


Fig-13 Gumbel distribution fitted with analytical data, probability limit values of this Gumbel distribution with two-side probability 5% and 500 sets of empirical distribution constructed of resampling samples for past experiments.

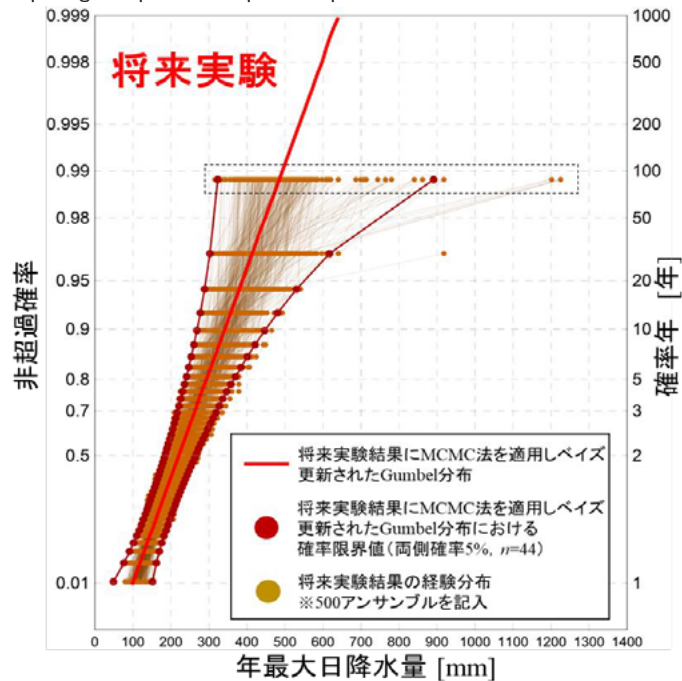


Fig-14 Gumbel distribution in future climate (Predicted distribution of Gumbel distribution fitted with analytical data), probability limit values of this Gumbel distribution with two-side probability 5% and 500 sets of empirical distribution constructed of resampling samples for future experiments



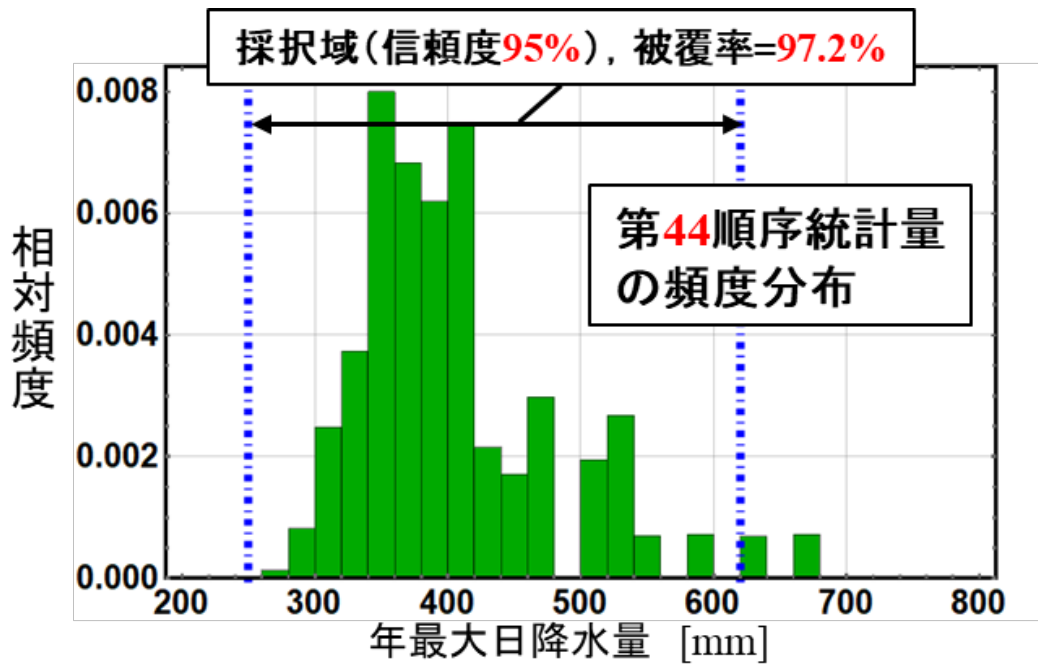


Fig-15 Relation between an adopted region of probability limit method test with two-side probability 5% for Gumbel distribution fitted with analytical data and a frequency distribution of 44<sup>th</sup> order statistics constructed of 5000 sets of resampling samples from past experiments

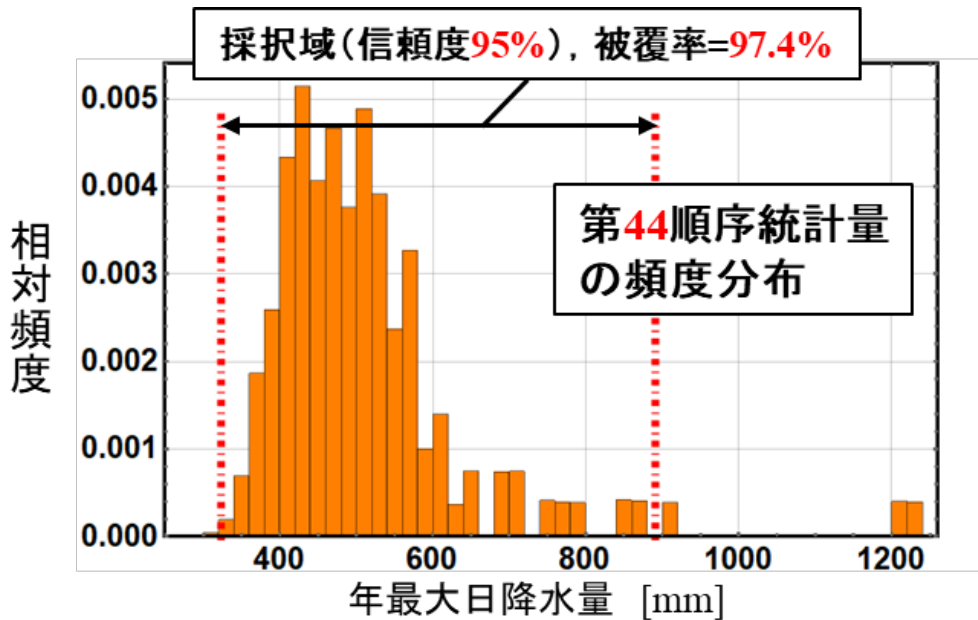


Fig-16 Relation between an adopted region of probability limit method test with two-side probability 5% for predicted distribution of Gumbel distribution fitted with analytical data and a frequency distribution of 44<sup>th</sup> order statistics constructed of 5000 sets of resampling samples from future experiments

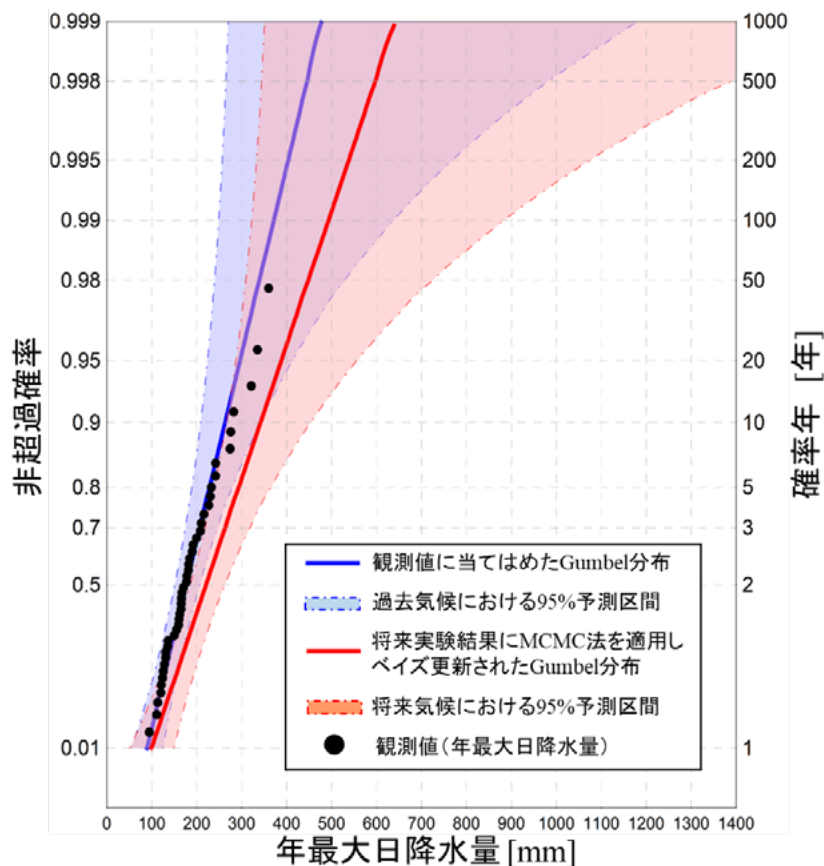


Fig-17 Update of prediction interval using Bayesian method which assimilates ensemble climate projection database

## 7 まとめ

本研究では、マルコフ連鎖モンテカルロ法により、気候変動予測結果を取り込み、信頼区間・予測区間を更新する手法を構築した。

地球温暖化に対応した洪水対策を検討するためには、今後降雨観測を継続していく中で、新たに得られた水文資料の統計的特徴を十分吟味し、降雨現象がどのように変化していくのかを把握していく必要があると言える。本論で挙げた成果は、検討流域において、将来  $n$  年間分の降雨資料を得た時点で、地球温暖化の影響の進行度合いを評価するとともに信頼区間の更新を行い、更新された信頼区間に基づき事業オプションを選定していくというような、時間軸の概念を取り入れた洪水リスク評価の実現可能性を示唆している。また、本論では、アンサンブル気候予測データを活用した物理的モンテカルロ法、数理統計理論である確率限界法検定、統計的モンテカルロ法により構成される信頼区間には、極めて調和的な関係が成立することを明らかとした。この成果は、今後の水文頻度解析に対して幅を有する考え方の科学的妥当性を提示している。

## 参考文献

- [1] 国土交通省 国土技術政策総合研究所 気候変動適応研究本部：気候変動適応策に関する研究（中間報告），国総研資料 749 号，II-112-154，2013.
- [2] Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z.W., Lettenmaier, D.P. and Stouffer,

- R.J. : Stationarity Is Dead: Whither Water Management ?, *Science*, 319, 573-574, 2008.
- [3] 北海道地方における気候変動予測（水分野）技術検討委員会  
([https://www.hkd.mlit.go.jp/ky/ky/kawa\\_kei/splaat000000vdyw.html](https://www.hkd.mlit.go.jp/ky/ky/kawa_kei/splaat000000vdyw.html))
- [4] 山田朋人, 星野剛ら : 北海道における気候変動に伴う洪水外力の変化, 河川技術論文集, 第 24 巻, pp.391-396, 2018.
- [5] 清水啓太, 山田朋人, 山田 正 : 確率限界法検定に基づく確率分布モデルの信頼区間を導入した新しい水文頻度解析手法, 土木学会論文集 B1(水工学) Vol.74, No.4, I\_331-I\_336, 2018.
- [6] 森口繁一 : 確率表現関数の検定について—Kolmogorov-Smirnov 検定を考え直す—, pp.233-244, 日本統計学会, 第 25 巻, 1995.
- [7] 森口繁一 : 確率表現関数, 東京大学出版, 1995.
- [8] 伊庭幸人, 種村正美 : 計算統計 II マルコフ連鎖モンテカルロ法とその周辺 (統計科学のフロンティア 12), 岩波書店, pp.10,37,49, 2005.
- [9] Ryo Mizuta, Akihiko Murata, and Masayoshi Ishii : Over 5,000 Years of Ensemble Future Climate Simulations by 60-km Global and 20-km Regional Atmospheric Models, *Bull.Am. Meteorol. Soc.*, pp.1383-1393, 2016.
- [10] 北野利一, 高橋倫也, 田中茂信 : ベイズ手法による降水量の極値の予測分布～最尤法による相違と相互補完, 土木学会論文集 B1 (水工学), 72(4), pp.I\_463-I\_468, 2016.