L-モーメント法を用いた高波の極値統計解析におけるプロッティングポジション公式の選択と各種統計量の信頼区間

Selection of Plotting Position Formulas in the *L*-moment method applied for Extreme Wave Analysis and Confidence Interval of Statistical Estimates

合田良実

Yoshimi GODA

土木学会名誉会員　工博 横浜国立大学名誉教授　(株) エコー（〒110-0014 東京都台東区北上野 2-6-4）

A Monte Carlo simulation study has been carried out on the reliability of the parameter and quantile estimation of the GEV, GPA, and Weibull distributions by means of the *L*-moment method for application to extreme wave statistics. The three population distributions are assigned with five to six values of the shape parameter each, and the scale and location parameters are so specified as to yield the 10-year and 50-year wave heights of 8.33 and 10.0 m, respectively. Samples of extreme wave heights are generated from 16 populations at the size varying from 10 to 1000, and 50,000 samples for respective conditions are analyzed for statistical characteristics. The unbiased plotting position recommended by Hosking and Wallis (1997) is confirmed to yield minimal biases on the parameter estimates, but it yields small biases on quantile estimates. A new plotting position estimator is proposed, which yields minimal quantile biases. A set of empirical formulas are derived for estimation of the coefficient of variations of the *L*-moment ratios and the quantile of the three distributions.

Key Words: *Extreme wave statistics, GEV distribution, GPA distribution, Weibull distribution, L-moments, plotting position, quantile variability*

1．まえがき

先に著者ら (2009a, 2009b) は極値波浪統計解析に対する*L*-モーメント法の適用法について検討し，Hosking · Wallis (1997) が取り扱っていないワイブル分布への適用について解説した．そして，日本沿岸の波浪観測地点のうち8個所を選び，*L*-モーメント法を使って確率波高の推定を行った．さらに著者ら (2010) はNOWPHASのデータベースから日本海沿岸11地点を選び，地域頻度解析の手法を適用してこの沿岸域の確率波高の推定を行った.

Hosking · Wallis (1997) が紹介している *L*-モーメント法による極値解析では，*L*-モーメントの計算に当たって使用するプロッティングポジションとして式(1) を示し，定数値として $\alpha = 0, \beta = 0$ の組み合わせと，α=0.35, β=0 の二つを提示している.

$$P_i = \frac{i-\alpha}{n+\beta} \tag{1}$$

ここに，i は昇順に並べ替えられた極値データの順位，n は極値データの総数，すなわち標本の大きさである.

一般化極値分布 (Generalized Extreme Value distribution: GEV) に対して確率重み付き積率法 (Probability-weighted moments: PWM) を適用する場合については，Hosking ほか (1985) が α = 0.35, β = 0 の組み合わせが良い結果を与えると報告している．また，一般化パレート分布 (Generalized Pareto distribution: GPA) に対しても，α = 0.35, β = 0 の組み合わせが適切であることを Hosking · Wallis (1987) が報告している．しかし，その

後のHosking・Wallis (1995) の検討では，$\alpha = 0, \beta = 0$ の組み合わせのほうが優れているとして，この組合せを不偏公式 (unbiased plotting position) と呼んでいる．この検討結果に基づいて，Hosking・Wallis (1997) はL-モーメント法の適用に際してプロッティングポジションとして不偏公式を用いることを推奨している．

極値統計解析に最小2乗法を用いる場合には，推計値の偏りが生じないようにプロッティングポジションについて吟味し，極値データに当てはめる分布関数ごとに最適のプロッティングポジション公式を選択する．これについては合田 (2008, 340-343頁) が詳しく述べている．

L-モーメント法の母体となったPWM法では，分布関数の母数や再現確率量の推計値の精度について多くの研究が行われており，推計値の真値からの偏りなどが吟味されてきた．それらによると，標本の大きさが100程度以下ではかなり大きな偏りを示すことが珍しくない（たとえば，Landwehrほか (1979a, 1979b)，Boesほか(1989) など）．

しかしながら，PWM法ならびにL-モーメント法に関する研究では，プロッティングポジション公式の選定についてあまり吟味されていないようである．そこで本研究では，L-モーメント法で使用するのに適切なプロッティングポジション公式を見いだすことを目的として，モンテカルロ法による大規模な数値実験を行うものである．

全体の構成は次の通りである．まず，2. においてここで使用する3種類の極値分布の関数形ならびにL-モーメント比を紹介する．L-モーメント法そのものについては先の著者らの論文 (2009a) で記述したので，ここでは割愛する．次に3. では，L-モーメント法を用いて極値統計解析を行う手順を説明し，4. で数値実験の対象とする極値分布や計算条件などの概要を説明する．数値計算の結果は，まず5. においてL-モーメントの推計値の信頼性について記述し，極値分布の母数推計については6.，確率波高の推計値の偏り率については7.，確率波高の信頼区間については8. で述べる．また，既往の諸研究との比較を9. で論じ，さらに10. において地域頻度解析における地域分割の際の不均質性の判断基準を新しく提案し，先に検討した日本海沿岸の高波への適用例を示す．最後に全体とまとめと今後の課題を11. に述べる．

2．極値分布関数とL-モーメント比の関係

Hosking (1990) は，各種の極値分布関数に対するL-モーメントλ_1とλ_2およびL-モーメント比τ_3とτ_4の理論値ならびに母数の推定式を導いた。なお，1次L-モーメントλ_1は平均値であり，1次L-モーメントと2次L-モーメントとの比 $\tau = \lambda_1 / \lambda_2$ はL-Coeffcient of Variation (L-CV) と名付けられている．また，L-モーメント比τ_3はL-skewness，L-モーメント比τ_4はL-kurtosisと呼ばれるので，本論文でもこれらの名称を使用する．L-モーメントおよびL-モーメント比ならびに分布関数の母数推定式は，Hosking・Wallis (1997) の解説書にも付録としてまとめられている。そのうちから，高波の極値統計解析に使用する分布関数に関わるものを以下に再録する。ただし，ワイブル分布については前論文（合田ら2009a, 2009b）においてGreenwoodほか(1978) を参照して独自に導いたものである。

(A) 指数型分布

a) 分布関数：

$$F(x) = 1 - \exp\{-(x-B)/A\}, \quad B \le x < \infty \tag{2}$$

b) 確率密度関数：

$$f(x) = 1/A \times \exp\{-(x-B)/A\} \tag{3}$$

c) 平均値と分散：

$$E[x] = B + A, \quad Var[x] = A^2 \tag{4}$$

d) 確率統計量(quantile)：

$$x = F^{-1}(P) = B - A\ln(1-P) \tag{5}$$

e) L-モーメント：

$$\lambda_1 = B + A,\ \lambda_2 = (1/2)A,\ \tau_3 = 1/3,\ \tau_4 = 1/6 \tag{6}$$

f) 母数推定式：

$$\hat{A} = 2l_2, \quad \hat{B} = l_1 - \hat{A} \tag{7}$$

ここで，$F(x)$ は非超過確率を表し，Aは尺度母数，Bは位置母数と呼ばれるものである．指数型分布は後述の一般化パレート分布で形状母数が$k = 0$，ワイブル分布で$k = 1.0$の場合に相当する．なお，$\hat{A}, \hat{B}$ の ˆ の記号は推計値であることを示す．以下においても，推計値には ˆ の記号を付す．また，l_1, l_2は標本について計算した1次と2次のL-モーメントを表す．

(B) ガンベル分布（極値I型分布）

a) 分布関数：

$$F(x) = \exp[-\exp\{-(x-B)/A\}], \ -\infty \le x < \infty \tag{8}$$

b) 確率密度関数：

$$f(x) = \frac{1}{A}\exp\left[-\frac{x-B}{A} - \exp\left(\frac{x-B}{A}\right)\right] \tag{9}$$

c) 平均値と分散：

$$E[x] = B + A\gamma, \quad Var[x] = \pi^2 A^2/6 \tag{10}$$

d) 確率統計量(quantile)：

$$x = F^{-1}(P) = B - A\ln(1-\ln P) \tag{11}$$

e) L-モーメント：

$$\lambda_1 = B + A\gamma,\ \lambda_2 = A\ln 2,\ \tau_3 \cong 0.1699,\ \tau_4 \cong 0.1504 \tag{12}$$

f) 母数推定式：

$$\hat{A} = l_2/\ln 2,\quad \hat{B} = l_1 - \hat{A}\gamma \tag{13}$$

ここにγ はオイラーの定数, 0.5772....である。ガンベル分布は次項の一般化極値分布で形状母数が $k = 0$ の場合に相当する.

(C) <u>一般化極値分布 (Generalized Extreme-Value Distribution: GEV)</u>

a) 分布関数：

$$F(x) = \begin{cases} \exp[-\{1-k(x-B)/A\}^{1/k}], & k>0:\ -\infty \le x < B + A/k \\ & k<0:\ B - A/k < x < \infty \\ \exp[-\exp\{-(x-B)/A\}], & k=0:\ -\infty < x < \infty \end{cases} \tag{14}$$

b) 確率密度関数：

$$\begin{aligned} f(x) &= [1-k(x-B)/A]^{1/k-1} \\ &\quad \times \exp\{-[1-k(x-B)/A]^{1/k}/A \quad : \quad k \ne 0 \end{aligned} \tag{15}$$

c) 平均値と分散：

$$\left.\begin{aligned} E[x] &= B + A[1-\Gamma(1+k)]/k \\ Var[x] &= A^2[\Gamma(1+2k) - \Gamma^2(1+2k)]/k^2 \end{aligned}\right\} \tag{16}$$

d) 確率統計量(quantile)：

$$x = F^{-1}(P) = \begin{cases} B + A\{1-(-\ln P)^k\}/k, & k \ne 0 \\ B - A\ln(-\ln P), & k = 0 \end{cases} \tag{17}$$

e) L-モーメント：

$$\begin{aligned} &\lambda_1 = B + A\{1-\Gamma(1+k)\}/k, \quad \lambda_2 = A(1-2^{-k})\Gamma(1+k)/k, \\ &\tau_3 = 2(1-3^{-k})/(1-2^{-k}) - 3, \\ &\tau_4 = \{5(1-4^{-k}) - 10(1-3^{-k}) + 6(1-2^{-k})\}/(1-2^{-k}) \end{aligned} \tag{18}$$

f) 母数推定式：

$$\begin{aligned} &\hat{k} \approx 7.8590c + 2.9554c^2, \quad c = \frac{2}{3+t_3} - \frac{\log 2}{\log 3} \\ &\hat{A} = \frac{l_2\hat{k}}{(1-2^{-\hat{k}})\Gamma(1+\hat{k})}, \quad \hat{B} = l_1 - \hat{A}\{1-\Gamma(1+\hat{k})\}/\hat{k} \end{aligned} \tag{19}$$

ここに，t_3は標本について計算したL-skewnessである.

一般化極値分布は，式(14)の表現では$k<0$のときに極値II型分布，$k>0$のときに極値III型分布となり，前者では確率統計量が下限値を持ち，後者では上限値を持つ。また，$k=0$のときは極値I型分布である。なお，高波の極値統計で使われている極値II型は式(14)の形状母数を$k'=-1/k$と変換したものに等しい。

また，式(19)における形状母数kは近似推定式であり，Hosking (1990) が導いたものである。

(D) <u>一般化パレート分布 (Generalized Pareto Distribution : GPA)</u>

a) 分布関数：

$$F(x) = \begin{cases} 1-[1-k(x-B)/A]^{1/k} & 1 \ge k > 0:\ B \le x < B + A/k \\ & k<0:\quad B \le x < \infty \\ 1-\exp\{-(x-B)/A\}, & k=0:\quad B \le x < \infty \end{cases} \tag{20}$$

b) 確率密度関数：

$$f(x) = [1-k(x-B)/A]^{1/k-1}/A \quad : B \le x,\ k \ne 0 \tag{21}$$

c) 平均値と分散：

$$\left.\begin{aligned} E[x] &= B + A/(1+k) \\ Var[x] &= A^2/[(1+k)^2(1+2k)] \end{aligned}\right\} \tag{22}$$

d) 確率統計量(quantile)：

$$x = F^{-1}(P) = \begin{cases} B + A\{1-(1-P)^k\}/k, & k \ne 0 \\ B - A\ln(1-P), & k = 0 \end{cases} \tag{23}$$

e) L-モーメント：

$$\begin{aligned} &\lambda_1 = B + A/(1+k), \quad \lambda_2 = A/\{(1+k)(2+k)\}, \\ &\tau_3 = (1-k)/(3+k), \quad \tau_4 = (1-k)(2-k)/\{(3+k)(4+k)\} \end{aligned} \tag{24}$$

f) 母数推定式：

$$\begin{aligned} &\hat{k} = (1-3t_3)/(1+t_3), \\ &\hat{A} = (1+\hat{k})(2+\hat{k})l_2, \quad \hat{B} = l_1 - (2+\hat{k})l_2 \end{aligned} \tag{25}$$

一般化パレート分布は一般化極値分布のような極大・極小値に対する漸近関数ではないけれども，ポアソン分布に従う確率変量のうち，ある閾値を超えるものに着目する場合などを表示するのに適切な分布関数といわれる。北野ほか (2002) は Kodiak 沖の高波データについてこの分布の適用性を検討している。なお，この分布関数の特性等については Coles (2001) などを参照されたい。

(E) ワイブル分布

a) 分布関数：

$$F(x) = 1 - \exp[-\{(x-B)/A\}^k], \quad B < x < \infty \tag{26}$$

b) 確率密度関数：

$$\begin{aligned} f(x) &= k/A \times [(x-B)/A]^{k-1} \\ &\quad \times \exp[-\{(x-B)/A\}^k] \quad : \quad B < x < \infty \end{aligned} \tag{27}$$

c) 平均値と分散：

$$\left.\begin{aligned} E[x] &= B + A\Gamma(1+1/k) \\ Var[x] &= A^2\left[\Gamma(1+2/k) - \Gamma^2(1+1/k)\right] \end{aligned}\right\} \tag{28}$$

d) 確率統計量 (quantile)：

$$x = F^{-1}(P) = B + A[\{-\ln(1-P)\}^{1/k}] \tag{29}$$

e) L-モーメント：

$$\begin{aligned} &\lambda_1 = B + A\Gamma(1+1/k), \quad \lambda_2 = A(1-2^{-1/k})\,\Gamma(1+1/k), \\ &\tau_3 = 3 - 2\,(1-3^{-1/k})/(1-2^{-1/k}), \\ &\tau_4 = \{5\,(1-4^{-1/k}) - 10\,(1-3^{-1/k}) + 6\,(1-2^{-1/k})\}/(1-2^{-1/k}) \end{aligned} \tag{30}$$

f) 母数推定式：

$$\hat{k} = 285.3t_3^6 - 658.6t_3^5 + 622.8t_3^4 - 317.2t_3^3 + 98.52t_3^2 - 21.256t_3 + 3.5160 \tag{31}$$

$$\hat{A} = \frac{l_2}{(1-2^{-1/\hat{k}})\Gamma(1+1/\hat{k})}, \quad \hat{B} = l_1 - \hat{A}\,\Gamma(1+1/\hat{k}) \tag{32}$$

なお，著者ら (2009b) の論文では，その中の**表－2**に記載した尺度母数Aの推定式が式 (32) の分子にkを乗じた形となっているけれども，それは校正漏れによる間違いである．ただし，論文中の計算結果はすべて式 (32) に基づいているので，訂正の必要はない．

３．*L*-モーメント法による極値統計解析の手順

(1) データの並べ替えと*L*-モーメントの計算

高波の資料（標本）が得られたならば，標本中のデータを昇順に並べ替え，1～4次のL-モーメントl_1～l_4，とその比$t=l_2/l_1, t_3=l_3/l_2$, および$t_4=l_4/l_2$を計算する．この計算には，Hoskingが開発したFORTRANプログラムlmoments[1]をベースにしてプログラムを作成し，それを使用する．

(2) 当てはめ候補の分布関数の母数値の推定

標本のL-skewness [t_3] が分かれば，式 (19), (25), (31) によって形状母数kの値が推定できる．指数分布とガンベル分布は形状が固定されているので，この手順は省略される．

尺度母数 $\hat{A}$ は推定された $\hat{k}$ 値と2次のL-モーメントl_2から求められ，位置母数 $\hat{B}$ はこれらの推計値に1次のL-モーメントl_1の情報を追加することで求められる．

(3) 最適合の分布関数の選定

標本に当てはめる分布関数として複数の候補を取り上げるときには，標本にもっとも適合する関数を選ぶ必要がある．著者ら (2009a, 2009b) は先に TUD 指標 (Twenty-Up Deviation) を提案した．これは，標本中で大きなほうから20個のデータを取り上げ，当てはめた分布関数を用いたこれら20位のデータに対する確率統計量の推計値と観測値の相対誤差の2乗平均の平方根値のもっとも小さいものを最適合と判断するものであった．

ここで，分布関数によるデータの推計値というのは，一般に次のように表される．

$$\hat{x}_i = F^{-1}(P_i) = \hat{B} + \hat{A}\,y_i \tag{33}$$

式 (33) 右辺の y_iは順位iの非超過確率P_iに対応する基準化変量と称するもので，分布関数毎に次のように与えられる．

1) 一般化極値分布 (GEV)：

$$y_i = \begin{cases} \{1-(-\ln P_i)^k\}/k, & k \neq 0 \\ -\ln(-\ln P_i), & k = 0 \end{cases} \tag{34}$$

2) 一般化パレート分布 (GPA)：

$$y_i = \begin{cases} [1-(1-P_i)^k]/k, & k \neq 0 \\ \ln(1-P_i), & k = 0 \end{cases} \tag{35}$$

3) ワイブル分布：

$$y_i = \{-\ln(1-P_i)\}^{1/k} \tag{36}$$

前述のように，GEV分布で$k=0$というのはガンベル分布であり，GPA分布で$k=0$というのは指数分布であるので，式 (34) と (35) はそれぞれガンベル分布と指数分布を包含している．

最小2乗法では，観測されたデータx_iと基準化変量y_iの間に式 (33) を仮定し，直線回帰式を適用して母数AとBの値を推計している．

TUD指標では上位 20 個のデータに限定したけれども，本研究では$i = 1, 2, \ldots, n$の全データについて観測値x_iとその推計値$\hat{x}_i$ の間の相関係数 $r(x_i, \hat{x}_i)$ を計算し，その上限値である1からの残差 $\Delta r = 1 - r(x_i, \hat{x}_i)$ が最小となるものを最適合と判断する．いわば，**最小残差法** (Minimum Residue of Correlation coefficient: MRC) を採用する．この方式は，“Probability Plot Correlation Method (PPC)” とも呼ばれている（Choudhury et al. 1991）

合田 (1988) が以前に提案したMIR法では，当てはめる分布関数と標本の大きさに応じた残差の平均値をあ

[1] http://lib.stat.cmu.edu からダウンロード可能である．

らかじめ算定しておき，残差平均値に対する比の大小で最適分布を判定した．しかし，L-モーメント法では相関係数の残差の平均値などを算定することができないため，残差の絶対値を判定の指標とするものである．

具体的には，観測された極値あるいはモンテカルロ法によって発生させた極値の標本に対して複数の極値分布関数を当てはめ，それぞれについて相関係数 $r(x_i, \hat{x}_i)$ の1からの残差 $\Delta r = 1 - r(x_i, \hat{x}_i)$ を算定する．そして，当てはめ候補の分布関数の中で残差が最小のものを最適合と見なす方式である．

(4) 確率波高の推計

標本に対して当てはめた分布関数の形状・尺度・位置の3母数が推定されると，所定の再現期間 R に対する確率波高 H_R は式 (33) を用いて推定できる．このときの基準化変量 y_i は，分布関数に応じて式 (34)～(36) によって計算する．その際の非超過確率 $P_i = P_R$ は次式で与えればよい．

$$P_R = 1 - \frac{1}{\lambda R} \tag{37}$$

ここに，λ は極値解析を行った高波データの年間の平均発生率である．

4．数値計算の条件と計算方法

(1) 発生させる高波の特性

数値シミュレーションによって高波の標本を多数作成する際には，高波として現地の波浪特性を反映させたものであることが必要である．高波の極値分布の特性として重要なのは，合田 (2002) が提唱した裾長度パラメータである．これは10年確率波高 H_{10} に対する50年確率波高 H_{50} の比として次のように定義される．

$$\gamma_{50} = H_{50} / H_{10} \tag{38}$$

わが国沿岸の高波の極値統計解析の結果では，裾長度パラメータ γ_{50} が 1.13～1.27 の値をとっている（合田 2008，375頁）．そこで，本研究では裾長度パラメータとして γ_{50} = 1.10, 1.20, および1.30の3通りの値を用いることとした．

また，50年確率波高としては H_{50} = 10.0 m に設定する．これは数値の換算が容易なように区切りの良い数値としたものである．

裾長度パラメータをある値に設定すると，分布関数の尺度母数と位置母数は任意の値を与えることができず，次式によって一義的に決定される値を使わなければならない．

$$\left.\begin{aligned} A &= H_{50} \frac{1 - 1/\gamma_{50}}{y_{50} - y_{10}} \\ B &= H_{50} \left[1 - \frac{1 - 1/\gamma_{50}}{1 - y_{10}/y_{50}} \right] \end{aligned}\right\} \tag{39}$$

ここに，y_{10} と y_{50} は再現期間10年と50年に対する基準化変量であり，$R = 10$ および50として式 (37) で非超過確率を求め，それを式 (34)～(36) に代入して算定される．たとえば，裾長度を $\gamma_{50} = 1.20$，平均発生率を $\lambda = 2$ に設定すると，指数分布であれば $H_{50} = 10.0$ m に対して自動的に $A = 1.036$ m，$B = 5.23$ m と定まる．

(2) 発生させる高波の極値分布関数

各種の極値分布は，L-モーメント比 τ_3 (L-skewness) と τ_4 (L-kurtoness) の相関関係によってその特性を表すことができる．**図-1** は，GEV，GPA，およびワイブル分布の特性曲線を示したもので，図中の◆，■，●の記号はそれぞれの分布関数で形状母数 k が特定の値をとる場合を示している．

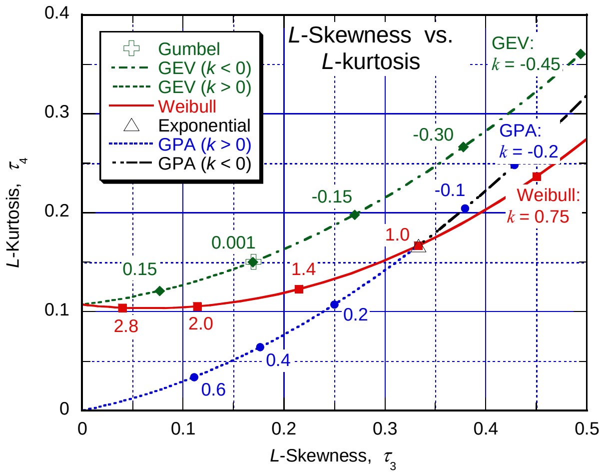


図-1　高波の母分布関数の L-kurtosis と L-skewness の関係

本研究では数値シミュレーションで発生させる高波の母分布関数として GEV分布，GPA分布，およびワイブル分布の3種類を考え，形状母数をそれぞれ5～6通りに変えた，全体で16種類の分布関数を使用する．分布関数と形状母数の組合せを**表-1** に示す．これらの分布関数の τ_3 と τ_4 の関係は**図-1** に表すとおりである．

なお，GEV分布の k = 0.001 はガンベル分布を近似的に表現したものである．また，ワイブル分布の k = 1.0 は指数分布であり，GPA分布の k = 0 と共通である．ただし，確認のためにGPA分布の k = 0.001 のケースも計算に加えた．

表-1　高波の発生に使用する分布関数

分布関数	形状母数 k					
一般化極値 (GEV)	-0.45	-0.3	-0.15	0.001	0.15	–
一般化ﾊﾟﾚｰﾄ (GPA)	-0.2	-0.1	0.001	0.2	0.4	0.6
ワイブル	0.75	1.0	1.4	2.0	2.8	–

(3) 高波の平均発生率と発生年数ならびに確率波高算定の再現期間

極値の統計的変動性はもっぱら標本の大きさ，すなわち標本中のデータ個数 n に支配される．今回の数値実験では，標本の大きさを n = 10, 20, 50, 100, 200, 500, および1000と7通りに変化させた．平均発生率は $\lambda = 2$ 個/年に設定したので，高波発生の期間が K = 5, 10, 25, 50, 100, 250, および500年に相当する．確率波高を算定する再現期間としては，R = 1, 2, 5, 10, 20, 50, 100, 200, 500, および1000年の10通りについて検討した．非超過確率では P_R = 0.5, 0.75, 0.90, 0.95, 0.975, 0.99, 0.995, 0.9975, 0.999, および0.9995にそれぞれ相当する．

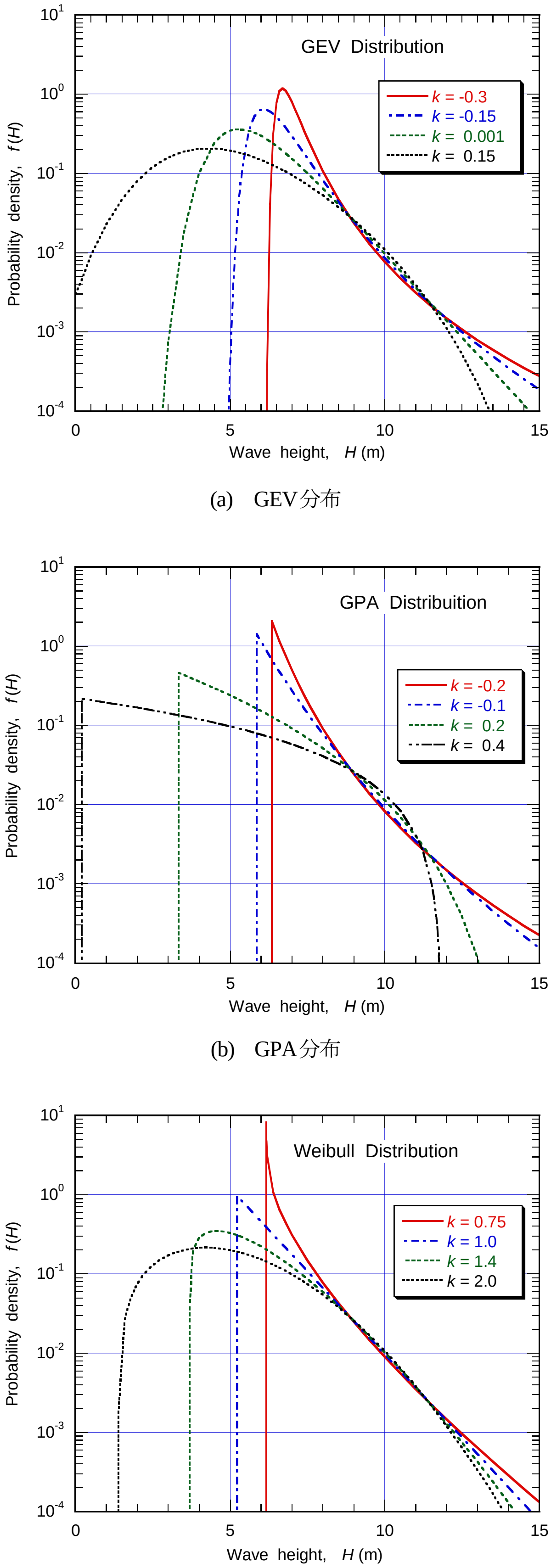
統計解析にあたっては，標本の大きさ毎に N = 50,000個の標本を作製し，解析結果の信頼度を高めるようにした．

(4) 計算に使用する尺度・位置母数の値

既に(1)で述べたように，極値分布関数が与えられ，裾長度パラメータが設定されたときには，尺度母数 A と位置母数 B が式(39)によって一義的に決定される．今回の数値計算では H_{50} = 10.0 m とし，裾長度を3通りに変えているので，この条件の下で**表-1**の母分布関数の尺度・位置母数の値を式(39)で計算したところ，**表-2～4**の母数値が得られた．

表で明らかなように，どの分布においても形状母数が大きくなるにつれて尺度母数と L-CV が増加し，位置母数と平均値が減少する．また，裾長度が大きくなると尺度母数と L-CV が増加し，位置母数と平均値が減少する．L-skewness と L-kurtosis は裾長度の影響を受けず，形状母数のみの関数である．

このようにして設定した極大値波高の確率密度関数をグラフ表示すると，裾長度が γ_{50} = 1.20の場合について**図-2**のようになる．裾長度をそのように設定したことによって，波高 H = 8～11 m くらいの範囲では，分布関数による差異が小さい．分布関数の差異が大きくなるのは，もっぱら波高の小さい範囲である．また，ワイブル分布とGPA分布ではそれぞれの位置母数の値以下の波高には確率密度関数が存在しない．



(a)　GEV分布

(b)　GPA分布

(c)　ワイブル分布

図-2　数値実験に用いる極値分布関数の確率密度関数（裾長度 γ_{50} = 1.20 の場合）

表-2　GEV分布の母数値，平均値，およびL−モーメント比（H_{50} = 10.0 m，平均発生率λ = 2）

k	L-skw	L-krt	γ_{50} = 1.1				γ_{50} = 1.2				γ_{50} = 1.3			
	(τ_3)	(τ_4)	A (m)	B (m)	平均値	L-CV	A (m)	B (m)	平均値	L-CV	A (m)	B (m)	平均値	L-CV
-0.45	0.4940	0.3602	0.0999	8.458	8.595	0.0153	0.182	7.198	7.424	0.0325	0.254	6.086	6.433	0. 0518
-0.3	0.3779	0.2663	0.179	8.223	8.400	0.0213	0.325	6.775	7.067	0.0464	0.454	5.488	5.939	0.0765
-0.15	0.2700	0.1977	0.319	7.886	8.125	0.0319	0.584	6.124	6.562	0.0723	0.809	4.633	5.240	0.1255
0	0.1693	0.1501	0.567	7.395	7.722	0.0509	1.026	5.289	5.823	0.1236	1.439	3.387	4.216	0.2364
0.15	0.0771	0.1209	0.997	6.686	7.131	0.0858	1.800	4.020	4.741	0.2367	2.530	1.589	2.718	0.5717

表-3　GPA分布の母数値，平均値，およびL−モーメント比（H_{50} = 10.0 m，平均発生率λ = 2）

k	L-skw	L-krt	γ_{50} = 1.1				γ_{50} = 1.2				γ_{50} = 1.3			
	(τ_3)	(τ_4)	A (m)	B (m)	平均値	L-CV	A (m)	B (m)	平均値	L-CV	A (m)	B (m)	平均値	L-CV
-0.2	0.4286	0.2481	0.263	8.012	8.341	0.0219	0.482	6.355	6.958	0.0481	0.668	4.953	5.788	0.0801
-0.1	0.3793	0.2042	0.386	7.743	8.172	0.0276	0.707	5.863	6.649	0.0622	0.980	4.271	5.360	0.1069
0.001	0.3329	0.1663	0.567	7.395	7.961	0.0356	1.040	5.224	6.262	0.0829	1.439	3.387	4.825	0.1489
0.2	0.2500	0.1071	1.203	6.381	7.383	0.0617	2.205	3.364	5.202	0.1606	3.053	0.812	3.356	0.3446
0.4	0.1765	0.0642	2.539	4.659	6.472	0.1168	4.655	0.207	3.532	0.3922	4.885	0.342	3.831	0.3794
0.6	0.1111	0.0338	3.507	4.715	6.906	0.1221	6.429	0.310	4.328	0.3570	–	–	–	–

注：γ_{50} = 1.2 の k = 0.6 については平均発生率が λ = 1 である．また，γ_{50} = 1.3 では k = 0.6 のケースを計算できない．

表-4　ワイブル分布の母数値，平均値，およびL−モーメント比（H_{50} = 10.0 m，平均発生率λ = 2）

k	L-skw	L-krt	γ_{50} = 1.1				γ_{50} = 1.2				γ_{50} = 1.3			
	(τ_3)	(τ_4)	A (m)	B (m)	平均値	L-CV	A (m)	B (m)	平均値	L-CV	A (m)	B (m)	平均値	L-CV
0.75	0.4505	0.2365	0.272	7.917	8.240	0.0237	0.499	6.180	6.774	0.0529	0.690	4.711	5.533	0.0896
1.0	0.3333	0.1667	0.565	7.399	7.964	0.0355	1.036	5.231	6.267	0.0826	1.434	3.397	4.831	0.1484
1.4	0.2150	0.1227	1.155	6.562	7.615	0.0538	2.117	3.698	5.627	0.1339	2.932	1.274	3.945	0.2644
2.0	0.1140	0.1054	2.190	5.301	7.241	0.0785	4.015	1.385	4.943	0.2108	5.012	0.087	4.529	0.2872
2.8	0.0401	0.1040	3.701	3.614	6.910	0.1046	5.937	0.336	5.623	0.2062	–	–	–	–

注：γ_{50} = 1.2 の k = 2.8 および γ_{50} = 1.3 の k = 2.0 については平均発生率が λ = 1 である．また，γ_{50} = 1.3 では k = 2.8 のケースを計算できない．

(5) 位置母数の制約条件

極値分布関数のうちで$k>0$の場合のGPA分布とワイブル分布は下限値がBであり，波高は負の値をとることがないので，位置母数Bは正の値でなければならない．GEV 分布で$k<0$の場合には下限値が$B-A/k$であるので，位置母数BはA/kよりも大きくなければならない．

位置母数Bが正となる条件は，式 (39) の第二式を書き換えることにより，次のように求められる．

$$\frac{B}{H_{50}}=1-\frac{1-1/\gamma_{50}}{1-y_{10}/y_{50}}>0$$

$$1>\frac{1-1/\gamma_{50}}{1-y_{10}/y_{50}}, \qquad 1-y_{10}/y_{50}>1-1/\gamma_{50}$$

$$\therefore \quad \frac{y_{50}}{y_{10}}>\gamma_{50} \tag{40}$$

基準化変量y_{50}とy_{10}は，再現期間をそれぞれ$R=50$および10年として式 (37) で非超過確率P_Rを算定し，それを式 (34)～(36) に代入して計算する．すなわち，基準化変量y_{50}とy_{10}は平均発生率λの関数であり，平均発生率が大きな場合には式 (40) の条件を満たさないことが起きる．GPA分布とワイブル分布について，裾長度がγ_{50} = 1.2と1.3のときに許容される平均発生率の最大値を試算すると，**表-5**の結果が得られる．

**表-5　位置母数を正値にとどめるための
平均発生率の許容最大値 λ_{max}**

ワイブル分布			GPA分布			GEV分布		
k	γ_{50} = 1.2	γ_{50} = 1.3	k	γ_{50} = 1.2	γ_{50} = 1.3	k	γ_{50} = 1.2	γ_{50} = 1.3
1.2	72.1	7.74	0.1	25.9	5.59	0.1	26.1	5.72
1.4	25.3	3.75	0.2	7.57	2.59	0.15	13.0	3.75
1.6	11.5	2.17	0.3	3.53	1.55	0.2	7.68	2.69
2.0	3.87	1.03	0.4	2.09	1.07	0.3	3.63	1.63
2.8	1.03	< 1.0	0.6	1.04	<1.0	0.4	2.17	1.14

この表に示した平均発生率の許容最大値は，モン

テカルロ法で数値シミュレーションを行う場合に考慮しなければならない数値である．実際の高波の極大値データを極値統計解析する場合には，極大値を規定する波高の閾値によって高波の平均発生率が定まる．すなわち，閾値を低くするほど平均発生率が大きくなる．高波のデータに対する極値分布がある程度予測されるときには，**表-2**～**4** の値を参考として波高の閾値をある程度高めに設定する必要がある．

(6) 検討対象のプロッティングポジション式

式 (1) のプロッティングポジションにおいては，定数 α と β にどのような値を与えるかによって極値データの L-モーメントの値が異なり，その結果として確率再現波高も異なるものとなる．

極値統計において標本に分布関数を当てはめる方法の選択に際しては，**不偏性** (unbiasedness) と**有効性** (efficiency) の二つの条件について検討しなければならない．不偏性というのは，同一の母集団に属する多数の標本について推計した分布関数の母数や確率再現統計量の平均値が母集団の値からずれていないという条件である．また，それぞれの標本から推定した統計量は平均値の周りに分散するけれども，そのばらつきの度合いが小さいことを有効性と称する．

プロッティングポジションの与え方によって，不偏性と有効性の条件を満たす様態が異なると考えられるため，本研究では**表-5** に示す 5 種類のプロッティングポジション式を取り上げた．

A は，Hosking · Wallis (1995) が Wakeby 分布 (5 母数の分布) と GEV 分布について数値実験を行って B よりも優れていると結論付け，地域頻度解析の解説書で推奨しているものである．

表-5　検討したプロッティングポジション式

記号	α	β	備考
A	0	0	Unbiased position
B	0.35	0	Landwehr et al.
C	0.45	0	New proposal
D	0.50	0	Temporal trial
E	0.44	0.12	Gringorten
F	0.375	0.25	Blom

B は Landwehr ほか (1979) が提案したもので，Hosking · Wallis (1995, 1997) はこれを Plotting Position 式と呼んでいる．しかし，プロッティングポジションとしては他にもいろいろあるので，これだけをプロッティングポジション式と呼ぶのは適切ではない．

C は，B のプロッティングポジションを用いたときの L-モーメントその他の偏りが無視できない大きさであったため，その改良として今回の数値実験の過程で提案するものである．D は C よりも定数 α の値を少し大きくしてみたもので，ワイブル分布の k = 1.0 のケースに適用したところ結果が不具合であったため，他のケースでは試みていない．

E は，ガンベル分布に対して Gringorten (1963) が提案したもので，最小 2 乗法で当てはめを行うときには偏りが生じないことを合田 (1988) が数値実験で確認している．F は，正規分布および対数正規分布に対して不偏性を保持するプロッティングポジションとして Blom (1965) が提案したもので，不偏性については同じく合田 (1988) が確認している．

E と F のプロッティングポジション式の定数は次の条件を満たしている．

$$2\alpha + \beta = 1 \tag{41}$$

このように設定すると，式 (1) の昇順の序数 i を降順の序数 m に替えても定数 α と β として同じ値を使うことができる．

プロッティングポジション式としては，極値 II 型分布を対象として提案された合田・小野澤 (1990) の式，ワイブル分布に対して提案された合田 (1988) の式その他がある．ただし，この両者とも形状母数をあらかじめ仮定することが必要なため，母分布関数が不確定のときには適用がむずかしい．そのため，本研究では**表-5** の 5 種類のプロッティングポジション式を比較検討の対象としたものである．

(7) モンテカルロ法による標本の作成法

モンテカルロ法では，[0, 1] の範囲に一様に分布する疑似乱数をコンピュータで発生させ，その値を非超過発生確率 P_i として式 (17), (23), (29) に代入して，それぞれ GEV, GPA, ワイブル分布の確率統計量 x_i を計算する．この操作を繰り返すことによって，所定の大きさの極値の標本が作成できる．なお，一様疑似乱数の発生には合田 (2008, p. 294) のアルゴリズムを使用した．

発生させた標本が所定の分布に従っていることは，それぞれの標本について平均値と分散を求め，標本の大きさごとに 50,000 回繰り返した結果の総平均と総分散が，統計的変動性の範囲内で理論値に一致することで確認した．

今回の対象とした 3 種類の極値分布の平均値と分散は式 (16), (22), (28) で与えられる．このうち，平均

値については**表-2**～**4**に記載済みである．なお，各標本の平均値は，これらの表に記載した平均値の周りに$\sigma/\sqrt{n}$の標準偏差を持って変動する．たとえば，ワイブル分布で裾長度が$\gamma_{50}=1.2$であって$k=1.0$，$n=10$の標本であれば，式(28)による標準偏差が$\sigma=1.036$ mであるので，標本の平均値6.27 mに対してその90%信頼区間が5.73～6.80 mと推定される．

(8) 推計値の偏り率（Bias）の定義

統計解析で偏りというのは，一つの標本から推定された値が真値から離れている度合いをいう．ただし，個々の標本は統計的に変動するため，多数の標本からそれぞれ得られた推計値の平均を用いて真値からの偏りを判断する．

本研究ではL-モーメント諸量，分布関数の母数値，確率波高などを推計するので，こうした標本ごとの推計値をξ_i，母集団の真値をξ_0で表す．推計値の平均$\mu[\xi]$ならびに標準偏差$\sigma[\xi]$は次のように定義される．

$$\mu[\xi]=\frac{1}{N}\sum_{i=1}^{N}\xi_i \tag{42}$$

$$\sigma[\xi]=\left[\frac{1}{N}\sum_{i}^{N}\xi_i^2-\mu^2[\xi]\right]^{1/2} \tag{43}$$

ここに，Nは標本の総個数であり，本研究ではすべて$N=50{,}000$である．

偏りの度合いは絶対値ならびに相対比の二つの方法で定義することができる．すなわち，

$$\text{Absolute Bias}=\mu[\xi]-\xi_0 \tag{44}$$

$$\text{Relative Bias}=(\mu[\xi]-\xi_0)/\xi_0 \tag{45}$$

式(44)の偏りはここでは偏り量と呼び，これは，分布関数の形状母数などその値が0となる可能性のあるものに対して用いる．式(45)の偏りの相対比はここでは偏り率と呼ぶことにする．なお，表示に際しては実数値ではなく，百分率で表すこともある．

推計値のばらつきの度合いは，式(43)で算定される標準偏差で表されるが，推計値の偏りも含めた2乗平均誤差RMSEで表されることも多い．これは次のように計算される．

$$\text{RMSE}=\left[(\mu[\xi]-\xi_0)^2+\sigma^2[\xi]\right]^{1/2} \tag{46}$$

５．L−モーメント比の推計値の偏り率

(1) 母集団のL−モーメント比

本研究で母集団として用いている3種類の分布関数について**表-1**の形状母数を使って計算したL-モーメント比の理論値は，**表-2**～**4**に記載したとおりである．

L-モーメント比のうちで，L-skewness (τ_3) と L-kurtosis (τ_4) は式(18), (24), (30)で明らかなように形状母数kのみの関数であり，裾長度の影響を受けない．しかし，L-CV (τ) は2次と1次のL-モーメント比であるため，尺度母数と位置母数の値によって変化する．この辺りの状況は，**表-2**～**4**の数値を比較すれば明らかである．

(2) 標本の大きさによる偏り率の変化

標本から得られる各種の推計値は，標本が小さいほど偏りが大きく，変動も激しくなる．**図-3**は，ワイブル分布で形状母数が$k=1.0$の場合，すなわち指数分布についてL-CV, L-skewness, およびL-kurotosisの偏り率（百分率表示）を示したものである．裾長度は$\gamma_{50}=1.20$の場合である．いずれも**表-5**に記した6通りのプロッティングポジション式を用いた結果を表示している．

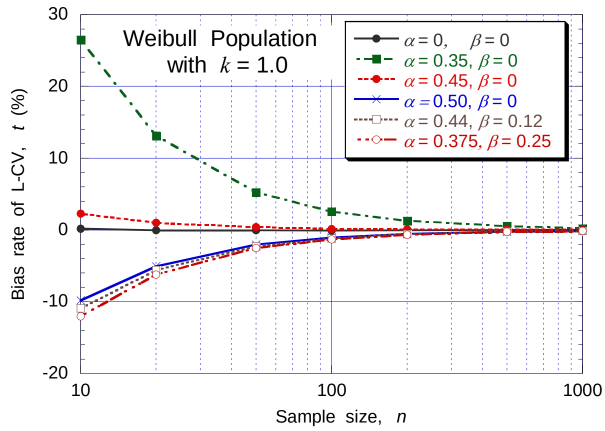
3種類のL-モーメント比のいずれについても，Unbiased Position ($\alpha=0, \beta=0$) は偏り率が極めて小さく，その名称にふさわしい結果である．

Landwehrほか(1979)によるプロッティングポジション($\alpha=0.35, \beta=0$)は，L-CVとL-kurotosisに関する偏り率が大きく，これを用いると不適切な推定結果をもたらす危険性がある．これに対して新しく提案した定数の式($\alpha=0.45, \beta=0$)は，偏り率を大きく縮減しており，Unbiased Positionによる結果にかなり近づいている．試みにこの定数の値を$\alpha=0.45$から$\alpha=0.50$に変えてみたところ，L-CVとL-kurotosisに関する偏り率が負の値に大きく振れてしまい，プロッティングポジションとして不具合な結果であった．

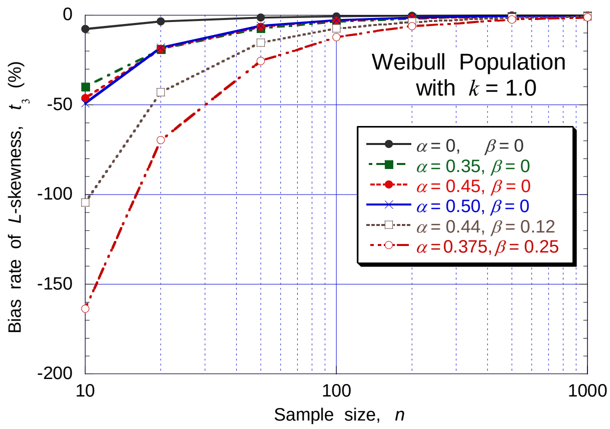
$\alpha=0.50$のプロッティングポジションはこの分布関数以外には試してみなかったけれども，$\alpha=0.35, 0.45$, および0.50の比較結果から見て$\alpha=0.45$のプロッティングポジションが実用的であると判断した．

L-skewnessの偏り率に関しては，定数の値を$\alpha=0.35$から$\alpha=0.50$に変えてもほとんど差異が見られず，$\alpha=0$のUnbiased Positionよりはやや劣る結果であった．

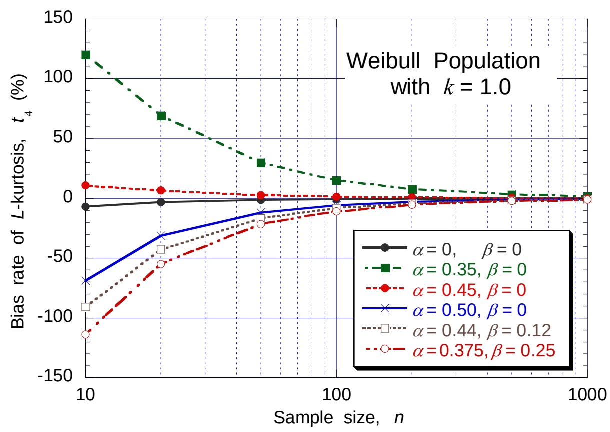
一方，最小2乗法で使われるGringorten式とBlom式は，L-CV, L-skewness, およびL-kurotosisのいずれ



(a) *L*-CV



(b) *L*-skewness



(c) *L*-kurtosis

図-3　ワイブル分布 (k = 1.0) に対する *L*-モーメント比の偏り率

についても偏り率が負の大きな値となり，*L*-モーメントの計算には不向きであることが例示される．

数値シミュレーションでは，裾長度が γ_{50} = 1.20 の場合について A～F のプロッティングポジションのうち，D を除くすべてを使って標本を作成して統計量の解析を行った．しかし，**図-3** の例に見られるように，α = 0 の Unbiased Position と，新提案の α = 0.45 以外のプロッティングポジション式は *L*-モーメント比をはじめとする各種の推計量の偏りが大きいことが確認された．このため，*L*-モーメント法による極値解析ではこれらのプロッティングポジション式を推奨することができない．したがって以下においては，プロッティングポジション式を α = 0 と α = 0.45 の二つに限定して記述する．

(3) 正規化された *L*–モーメント比の偏り率

図-3 は *L*-モーメント比の偏り率を標本の大きさに対して表示したものであり，これをみると偏り率がほぼ標本の大きさ n に反比例している．この関係は理論的にも予測されるもののようで，Hosking ほか (1985) は一般化極値分布 (GEV) の偏り量に標本の大きさ n を乗じた量についての漸近理論値を示している．

そこで，これ以降においては偏り量あるいは偏り率に n を乗じた値を標本の大きさ毎に求め，それを複数の標本の大きさについて平均したものを分布関数の代表値として用いる．ここでは，これを正規化された偏り量・偏り率と称する．平均値を求める際には，一般に標本の大きさが n = 50～500 のものを対象としたが，偏り率と標本の大きさの反比例関係が確認できる場合には n = 20～1000 の範囲にまで拡げたケースもある．

図-4 は，一般化極値分布 (GEV)，一般化パレート分布 (GPA)，およびワイブル分布のそれぞれについて，*L*-CV, *L*-skewness, および *L*-kurtosis の正規化された偏り率と母集団の *L*-skewness [τ_3] との関係を示したものである．プロッティングポジションとしては，α = 0 と 0.45 の 2 種類を用いた結果を区別して示している．前者の計算結果は黒塗りの記号，後者は白抜きの記号で表示している．

裾長度を 3 通りに変えて数値計算した結果のうち，*L*-skewness は裾長度の影響がほとんど見られないのに対し，*L*-CV と *L*-kurtosis は裾長度によって異なる値を示すこともあった．ここでは，全体の傾向を示すために，3 通りの裾長度の計算結果の算術平均の値で表示している．

図-4(a) は *L*-CV の偏り率を示している．α = 0 のプロッティングポジションでは偏り率がゼロといってよい．α = 0.45 では偏り率がやや大きく，特に一般化極値分布では *L*-skewness が 0.3 以上のときに偏り率が 1.0/n を超えている．

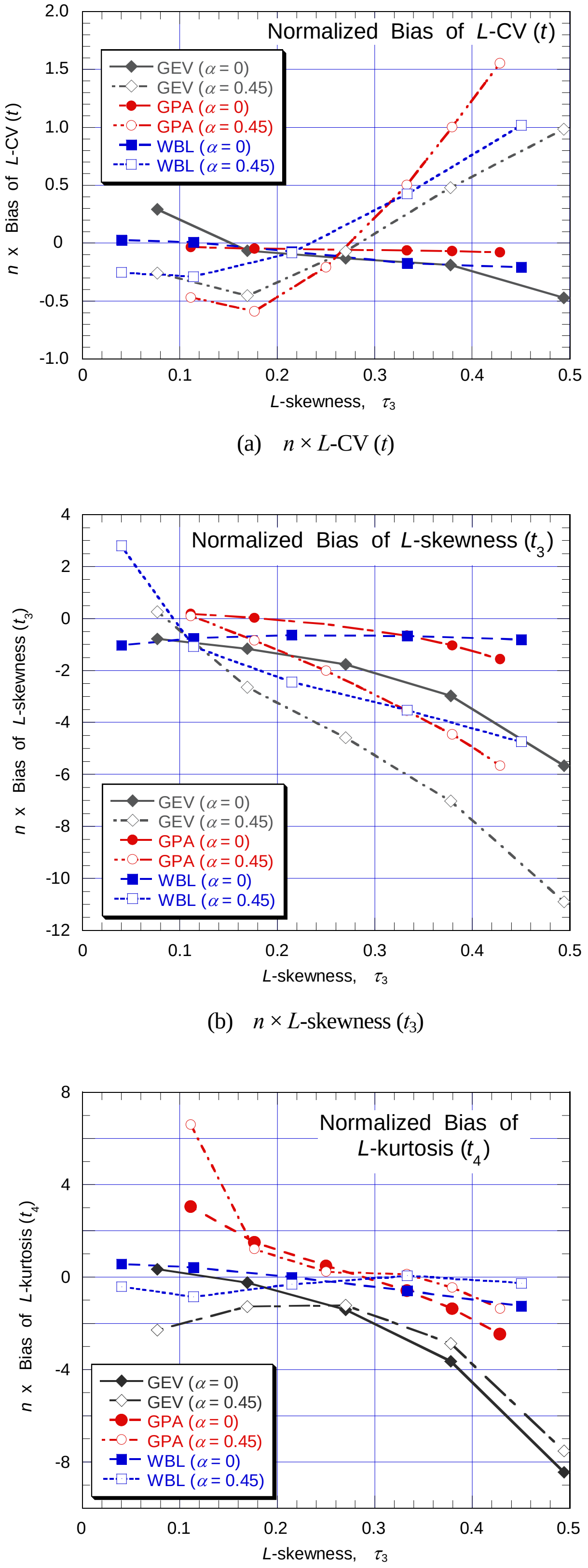


図-4　正規化されたL-モーメント比の偏り量

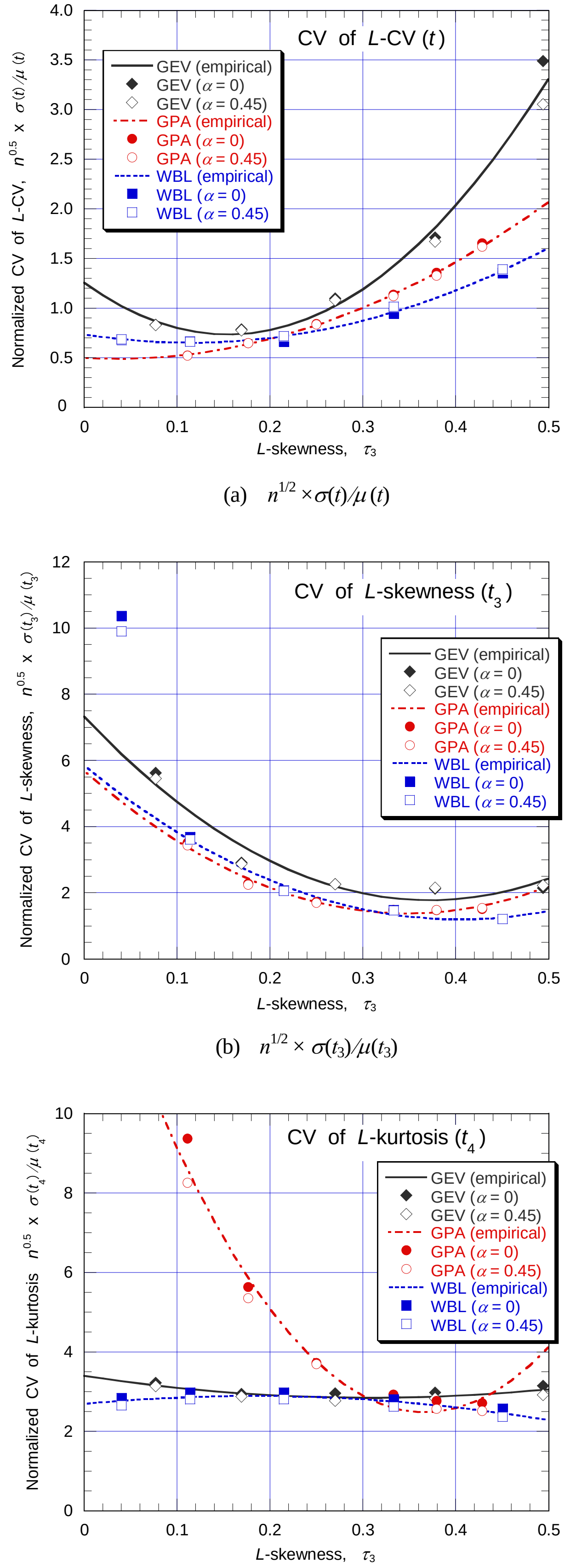


図-5　正規化されたL-モーメント比の変動係数

L-skewness に関しては**図-4**(b) に示しており，偏り率はL-CV よりも大きく，負の値を取るのが大半である．二つのプロッティングポジションを比較すると，$\alpha = 0$ のケースが全体に小さく，$\alpha = 0.45$ では偏り率が大きくなる．ワイブル分布の $\alpha = 0$ のケース以外は，L-skewness が大きくなる（確率密度分布の右裾が長くなる）ほど偏り率が大きくなっている．また，$\alpha = 0.45$ のプロッティングポジションは $\alpha = 0$ による偏り率の2倍以上の場合がある．

図-4(c) はL-kurtosis の偏り率を示したものである．これに関しては，二つのプロッティングポジションによる結果には有意な差がなく，ほぼ同じ偏り率を示す．分布関数の中では，一般化極値分布でL-skewness が $t_3 = 0.494$ ($k = -0.45$) のときには約 $8/n$ の大きな偏り率を示している．

(4) 正規化されたL−モーメント比の変動係数

L-モーメント比の推計値に関しては，平均値の偏りのみならず，推計値の変動の幅も検討する必要がある．先に Hosking ほか (1985) は，一般化極値分布 (GEV) の母数推計値の分散についても，それが標本の大きさに反比例して減少することを前提にした漸近理論値の図表を提示している．標準偏差は標本の大きさの平方根，すなわち $n^{1/2}$ に反比例すると考えられる．そこで，L-モーメント比の推計値の標準偏差（絶対値）については，それと $n^{1/2}$ の積を計算し，それを標本の大きさのある範囲（$n = 50$～1000）について平均して代表値とすることとした．

L-モーメント比のうちで L-skewness と L-kurtosis は形状母数 k のみの関数であるので，裾長度には依存しない．しかし，L-CV は尺度母数と位置母数の値にも依存するので，その値は形状母数のみでなく裾長度によって変化する．しかし，標準偏差と平均値との比，すなわち変動係数として整理すると，裾長度に依存しないことが数値計算の結果として確認された．そこで，標準偏差に $n^{1/2}$ を乗じたものを平均値で除した値を求め，これを正規化された変動係数と呼ぶ．

正規化された変動係数に関しては，$\alpha = 0$ と $\alpha = 0.45$ の二つのプロッティングポジションによる結果にほとんど差がなかったので，分布関数毎に両者の平均として整理した．一般化極値分布 (GEV)，一般化パレート分布，およびワイブル分布のそれぞれについて整理した結果を**図-5** に示す．

3通りの裾長度による計算結果による変動係数の値は，L-CV と L-skewness に関してはほとんど差異がなく，最大値と最小値の差は多くは 1% 以下，最大でも 5% 未満であった．ただし，L-kurtosis については裾長度によって数 % の差異が見られることが多く，最大では 20% 近い差が生じた．しかしデータ整理としては，3通りの裾長度による計算結果の平均値を用いることとした．

このようにして求めた $\alpha = 0$ のプロッティングポジションによる結果は黒塗りの記号，$\alpha = 0.45$ のを用いた結果は白抜きの記号で示してある．**図-5**(a), (b), (c) は，それぞれ L-CV, L-skewness, および L-kurtosis の変動係数を母集団の L-skewness (τ_3) に対してプロットしたものである．二つのプロッティングポジションによる計算結果にほとんど差が見られないところから，両者の平均値に対して2次曲線を当てはめた．それによって得られた算定式は**表-6** に示すとおりである．**図-5** に描かれている曲線は，この近似推定式によるものである．

表-6　L-モーメント比の変動係数の推定式

L-モーメント比	分布関数	変動係数CV の推定式
L-CV (t)	一般化極値分布 (GEV)	$\mathrm{CV}(t) = (21.6745t_3^2 - 6.7226t_3 + 1.256)/n^{0.5}$
	一般化パレート分布 (GPA)	$\mathrm{CV}(t) = (7.2791t_3^2 - 0.4957t_3 + 0.4973)/n^{0.5}$
	ワイブル分布 (WBL)	$\mathrm{CV}(t) = (6.4538t_3^2 - 1.4805t_3 + 0.7362)/n^{0.5}$
L-skewness (t_3)	一般化極値分布 (GEV)	$\mathrm{CV}(t_3) = (40.1193t_3^2 - 29.8590t_3 + 7.3288)/n^{0.5}$
	一般化パレート分布 (GPA)	$\mathrm{CV}(t_3) = (35.5486t_3^2 - 24.7527t_3 + 5.6864)/n^{0.5}$
	ワイブル分布 (WBL)	$\mathrm{CV}(t_3) = (28.5170t_3^2 - 23.0420t_3 + 5.8494)/n^{0.5}$
L-kurtosis (t_4)	一般化極値分布 (GEV)	$\mathrm{CV}(t_4) = (5.8445t_3^2 - 3.5959t_3 + 3.4001)/n^{0.5}$
	一般化パレート分布 (GPA)	$\mathrm{CV}(t_4) = (93.1882t_3^2 - 68.4313t_3 + 15.0430)/n^{0.5}$
	ワイブル分布 (WBL)	$\mathrm{CV}(t_4) = (-6.0637t_3^2 + 2.2134t_3 + 2.6893)/n^{0.5}$

表-7　形状母数の推計値の標準偏差 [$\sigma(k)$] の変動係数の推定式

分布関数	標準偏差の推定式
一般化極値分布 (GEV)	$\sigma(k) = (2.4457t_3^2 - 0.0543t_3 + 0.6998)/n^{0.5}$
一般化パレート分布 (GPA)	$\sigma(k) = (7.1472t_3^2 - 3.9073t_3 + 1.6025)/n^{0.5}$
ワイブル分布 (WBL)	$\sigma(k) = (26.8575t_3^2 - 22.8583t_3 + 5.8874)/n^{0.5}$

ここで，*L*-skewness の変動係数のうちのワイブル分布に関しては，$k = 2.8$ のケースが $k = 2.0$ 以下のケースからかけ離れて大きいな変動係数を示したため，これを除外して2次曲線の当てはめを行っている．

なお，**図-4, 5** においては横軸の *L*-skewness が $\tau_3 = 0.3333$ の値を取る縦軸上でにおいて，一般化パレート分布 (GPA) とワイブル分布 (WBL) の線ならびに計算点が交差している．ここではGPA が $k = 0.001$ の値を取り，WBL が $k = 1.0$ の値であって，ともに指数分布を代表している．**図-5**(c) で両者が完全に一致しないのは，標本の統計的変動性が強く現れたためと思われる．また，推定式がこの点だけでなく，他の形状母数の点も含めた当てはめによって設定されていることも関係している．

表-6 に記載した推定式を使うと，**図-5** に示した *L*-skewness (τ_3) の値の範囲で，標本から推計された *L*-モーメント比の変動係数，ひいては *L*-モーメント比の信頼区間を推計できる．たとえば，大きさ $n = 500$ の標本から *L*-モーメント比として $t = 0.1096$, $t_3 = 0.2589$, $t_4 = 0.1299$ の値が得られたとする．この標本にワイブル分布を当てはめると，その形状母数は *L*-skewness である t_3 の値から $k = 1.23$ と推計される（式 (31) による）．一方，**表-6** 中の近似式を用いて *L*-モーメント比の変動係数を計算し，それにそれぞれのて *L*-モーメント比の値を乗じると，標準偏差が次のように推定される．

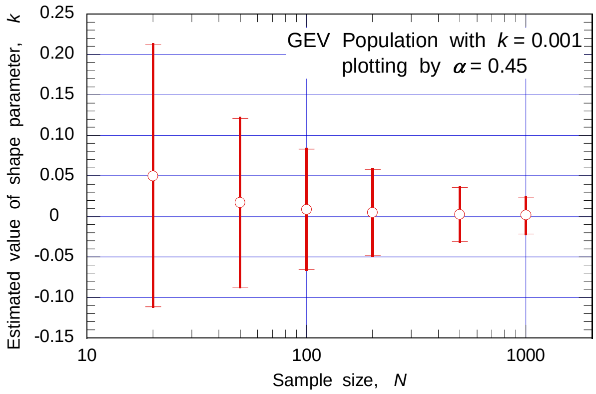
$$\sigma(t) = 0.0039,\ \sigma(t_3) = 0.0208,\ \sigma(t_4) = 0.0166.$$

推計された *L*-モーメント比の信頼区間を標準偏差の ±1.64 倍に取ると，*L*-CV, *L*-skewness, および *L*-kurtosis についてそれぞれ次のように推定される．

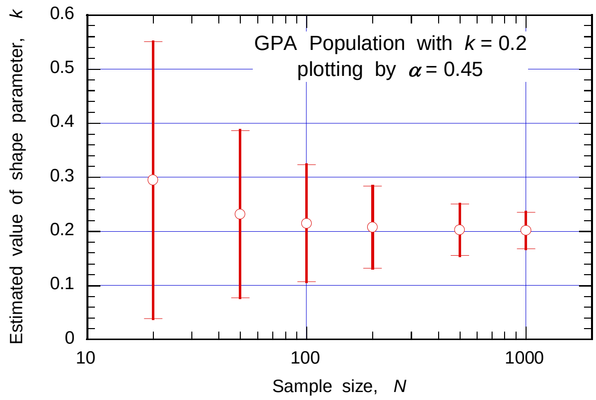
$$(t)_{90\%} = 0.1033〜0.1159,$$

$$(t_3)_{90\%} = 0.2248〜0.2930,$$

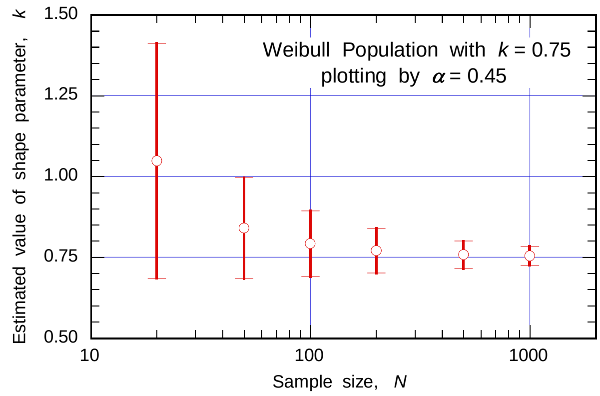
$$(t_4)_{90\%} = 0.1027〜0.1571$$



(a)　一般化極値分布 (GEV)：$k = 0.001$



(b)　一般化パレート分布 (GPA)：$k = 0.2$



(c)　ワイブル分布：$k = 0.75$

図-6　標本の大きさによる形状母数の推計値の変化

６．分布関数の母数推計値の偏り量と偏り率

(1) 標本の大きさによる形状母数の推計値の変化

一つの標本に対して極値分布関数を当てはめるときには3(2) に述べたように，まず形状母数を推計する．標本が小さいときには形状母数の推計値 $\hat{k}$ が真値 k_0 から偏るだけでなく，推計値が標本毎に統計的に大きく変動する．

図-6 は一般化極値分布 (GEV)，一般化パレート分布 (GPA)，およびワイブル分布のそれぞれにつ

いて形状母数の平均値とその 68% 信頼区間 (標準偏差の ±1 倍の範囲) を示したものである．標本の大きさが n = 10 のデータについては信頼区間の幅が大きすぎるため，表示から除外した．

(2) 正規化された形状母数の推計値の偏り量と標準偏差

図-6 から分かるように，標本の大きさが 50 未満では信頼区間の幅が大きすぎ，形状母数の値を適切に推定することが困難である．また，形状母数の推計値と真値との差 $\Delta k = \hat{k} - k_0$ は，標本の大きさに反比例して減少しているように見える．そこで，これまでと同じように標本の大きさ n を乗じて正規化された偏り量をとりあげ，標本の大きさが n = 50～500 の範囲での平均値を求めた．その結果は**図-7** に示すとおりである．この図には，形状母数 k の偏り量についてプロッティングポジションとして α=0 と α=0.45 の 2 種類を用いた結果を併示しており，3 種類の分布関数の形状母数推計値の偏り量を一括して示した．

形状母数の推計値に関しては，裾長度の影響をほとんど受けないことが数値計算の結果で確認されている．これは，**図-3**(b), **4**(b) に示した L–skewness が裾長度の影響を受けないことに対応している．

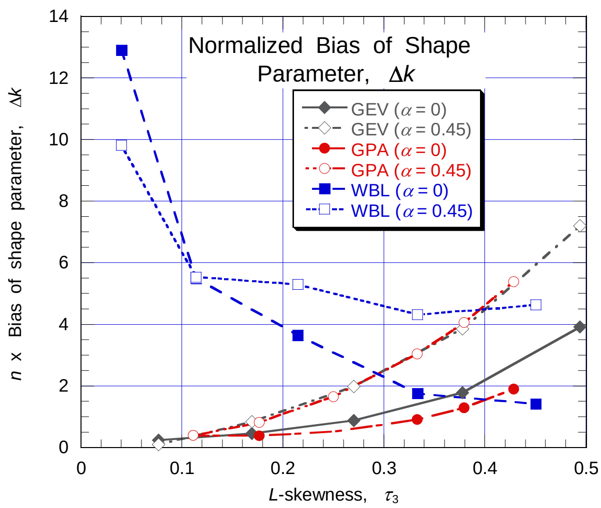


図-7　正規化された形状母数推計値の偏り量

形状母数の偏り量 Δk はすべて正の値である．すなわち，形状母数の推計値は母集団の値よりも常に大きく見積もられる．この傾向は**図-6** でも例示されている．偏り量が正値であるのは，**図-4**(b) において L–skewness の偏り量の大半が負の値であることに関係している．すなわち，**表-2**～**4** で明らかなように，形状母数と L–skewness は逆相関の関係にあるため，後者が小さくなると前者が大きくなるのである．

3 種類の分布関数のうち，一般化極値分布と一般化パレート分布では L-skewness [τ_3] の値の増加，すなわち形状母数の値の減少（正値から負値への変化）につれて形状母数の偏り量が増大する．ワイブル分布では逆に，母集団の形状母数が k = 0.75 から 2.8 へと増大するにつれて，形状母数推計値の偏り量が増加する．

形状母数の推計値の標準偏差については L-モーメント比と同じように，標準偏差の値に標本の大きさ n の 1/2 乗を乗じて正規化し，これを標本の大きさが n = 50～500 の範囲で平均化した．計算の結果は**図-8** のようになった．二つのプロッティングポジションによる結果にはほとんど差がないので，その平均値を対象として 2 次曲線を当てはめ，形状母数の標準偏差の推定式を求めた．この推定式は**表-7** に記載したようなものである．また，**図-8** 中の分布関数ごとの曲線は，この推定式で計算したものである．なお，ワイブル分布に対する当てはめに際しては，k = 2.8 のケースが他の形状母数のケースと大きく外れる結果を示したので，これを除外して推定式を導いた．

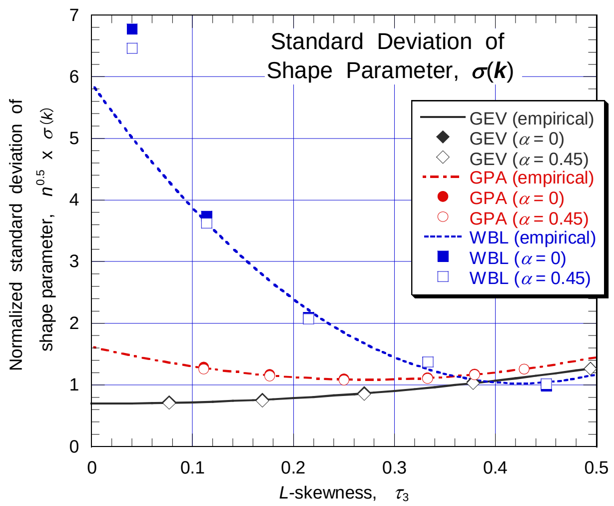


図-8　正規化された形状母数推計値の標準偏差

図-8 においては，一般化極値分布 (GEV) と一般化パレート分布 (GPA) は形状母数推計値の標準偏差が (0.7～1.6) $/n^{0.5}$ 程度であって，L-skewness の値にかかわらず，ほぼ一定である．これに対してワイブル分布では形状母数推計値の標準偏差が (1～

6) $/n^{0.5}$ の範囲で大きく変化している．ただし，*L*-skewness の値の減少に伴って形状母数の値が k = 0.75 から 2.8 に増大するので，変動係数としてみれば変化の幅は 2 倍以下である．

なお，形状母数推計値の標準偏差を指数分布にほぼ相当する k = 0.001 の一般化パレート分布と k = 1.0 のワイブル分布とで比較すると，後者のほうが約 24% 大きい．本来ならば同一の値になるべきものが食い違う理由については，今のところ不明である．

(3) 尺度・位置母数の推計値の偏り量と標準偏差

既に**表-2**～**4** に示したように，尺度母数と位置母数の値は裾長度によって大きく変化する．このため，尺度・位置母数の推計値に関しては真値からの偏りについては位置母数のような一般化が困難であり，分布関数ごとに定性的な記述にとどまらざるを得ない．

一般化極値分布では，尺度母数の推計値 $\hat{A}$ の偏り量は (-2～2) $/n$ の範囲で *L*–skewness の増加につれて負から正の値に変化する傾向があり，一方，裾長度が大きくなると偏り量が小さくなる傾向がある．位置母数の推計値 $\hat{B}$ の偏り量は (-0.4～1.0) $/n$ の範囲にあって *L*–skewness の値にあまり影響されないが，3 通りの裾長度のうちでは γ_{50} = 1.2 のときの偏り量が小さい．標準偏差は，尺度・位置母数とも *L*–skewness の値が大きくなると減少する傾向があるけれども，それらの値は裾長度によって異なる．変化の範囲は (0.1～2.8)$/n^{0.5}$ である．

一般化パレート分布では，尺度母数の推計値の偏り量が (0.5～8.7) $/n$ の範囲にあるが，その変化を裾長度や *L*–skewness に関連づけることが難しい．位置母数の推計値はおおむね負の値であり，最小値は約 –2.3 $/n$ である．尺度母数の推計値の標準偏差は，*L*–skewness の値の増加につれて減少する傾向があり，範囲としては (0.5～7.4) $/n^{0.5}$ である．位置母数の推計値の標準偏差も *L*–skewness の値の増加につれて減少する傾向があり，範囲としては (0.1～1.6) $/n^{0.5}$ である．尺度・位置母数の推計値の標準偏差は，裾長度が大きくなるにつれて増大する傾向が認められる．

ワイブル分布では，尺度母数の推計値の偏り量が *L*–skewness の値の増加につれて減少する傾向があり，変化の範囲は (0.2～13) $/n$ である．位置母数の推計値の偏り量も同じ傾向を示すものの，値としては負値であり， (-0.1～-13) $/n$ の範囲にある．標準偏差は尺度・位置母数ともに *L*–skewness の値の増加につれて減少する傾向を示しており，変化の範囲は尺度母数では (0.5～7.6) $/n^{0.5}$ で，位置母数では (0.1～7.4) $/n^{0.5}$ である．裾長度の影響は一様ではなく，規定しがたい．

このように，尺度母数と位置母数の推計値については数値実験によってあらかじめその統計的挙動を見定めることができない．与えられた標本に対して推計された母数値を使い，モンテカルロ法による数値実験で変動範囲を調べるのが唯一の方法と思われる．

７．確率波高の推計値の偏り率

(1) 当てはめ分布関数を母集団分布に固定した場合の確率波高の偏り率

標本に対して分布関数を当てはめ，形状・尺度・位置母数の値を推計すると，各種の再現期間に対する確率波高を推定することができる．一般には高波極値の母集団が分からないため，複数の分布関数をあてはめの候補とし，最適合の分布をもって母集団分布とみなしている．本研究では，あらかじめ設定した母集団から抽出した標本についてその統計的性質を調べているので，ここでは母集団の分布関数に当てはめて確率波高を推計した場合の偏り率について記述する．

確率波高の偏り率は，分布関数の形状母数，裾長度，標本の大きさ，プロッティングポジション式，および再現期間によって異なる．裾長度が大きな場合には，確率密度分布の右裾が長く伸び，再現期間の長い確率波高として大きな値が出やすくなる．このため，裾長度の増大に伴って確率波高の偏り率が大きくなる．

標本の大きさの影響は，*L*-モーメント比の場合と同様に偏り率が標本の大きさ n にほぼ反比例すると見なすことができる．そこで，プロッティングポジション式毎に偏り率と標本の大きさの積の平均値を求めて整理した結果が**図-9** である．ただし，n の小さい範囲や大きい範囲では偏り率と n^{-1} の比例関係が崩れる場合があり，特に再現期間の長いときに比例関係が崩れやすい．そのため，ここでは N = 50～500 の範囲での平均値を示している．なお，ここでは裾長度が γ_{50} = 1.2 の結果を示しており，裾長度による偏り率の変化については後述する．

プロッティングポジション式としては，$\alpha = 0$ と 0.45 の 2 種類を用いたものを併示している．ここに示しているのは式 (45) で定義される相対比としての偏り率である．たとえば，再現期間 $R = 100$ 年に対する図の縦軸の値が 0.5 であったとし，確率波高が大きさ $n = 100$ の標本から推計されたものであれば，確率波高の推計値は0.5% の偏り率を持つことになる．なお，本研究では高波の平均発生率を $\lambda = 2$ に設定して標本を作製しているので，再現期間 $R = 1000$ 年は非超過確率 $P_R = 0.9995$ に相当する．

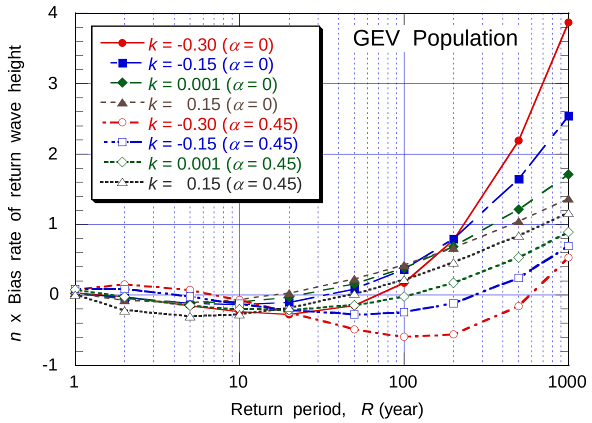
確率波高推計値の偏り率は 3 種類の分布関数のいずれにおいても，再現期間が 50 年 (非超過確率 $P_R = 0.99$) 以下であれば ±0.5/n 程度以下であり，ほぼ不偏性を満足しているといえる．しかし，再現期間が 50 年を超えるようになると偏り率が急速に増大し，再現期間 $R = 500$ 年 (非超過確率 $P_R = 0.999$) では偏り率が 2/n 以上となるケースも見られる．

分布関数の中では，ワイブル分布の偏り率が最小であり，一般化パレート分布の偏り率が最も大きい．また，$\alpha = 0$ のプロッティングポジション式は偏り率が正の値のことが多いのに比べて，$\alpha = 0.45$ のプロッティングポジション式は偏り率が負のことが多い．

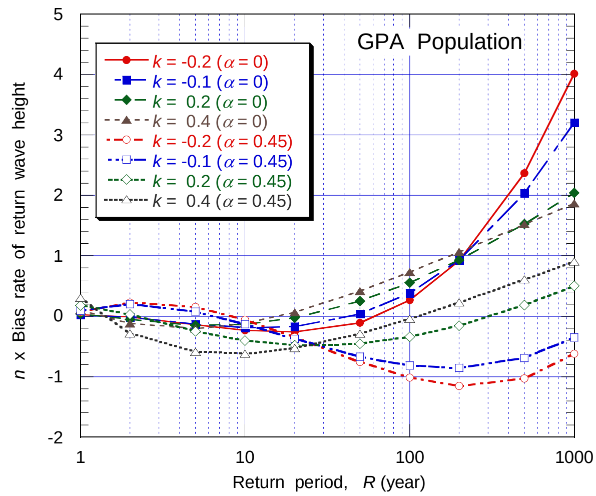
偏り率の絶対値で比較すると，再現期間の長い範囲で $\alpha = 0$ のプロッティングポジション式のほうの偏り率が大きくなっている．これまでのL-モーメント比や形状母数の推計値の場合には $\alpha = 0$ のほうが小さな偏りを示していたことを考えると，確率波高の推計値に関して二つのプロッティングポジション式の優劣が逆転するのはやや不思議な気もする．

もっとも，Hosking · Wallis (1995) が $\alpha = 0$ と $\alpha = 0.35$ のプロッティングポジションを比較検討したときには，地域頻度解析などで分布関数を適正に当てはめるためには $\alpha = 0$ を使うべきであると結論する一方で，確率波高の推計が主目的であれば $\alpha = 0.35$ の使用も許容できるとしている．今回新しく提案した $\alpha = 0.45$ のプロッティングポジションは $\alpha = 0.35$ の方式よりも推計値の偏りがはるかに小さいので，これを確率波高の推定に使用するのは当を得たものと思われる．

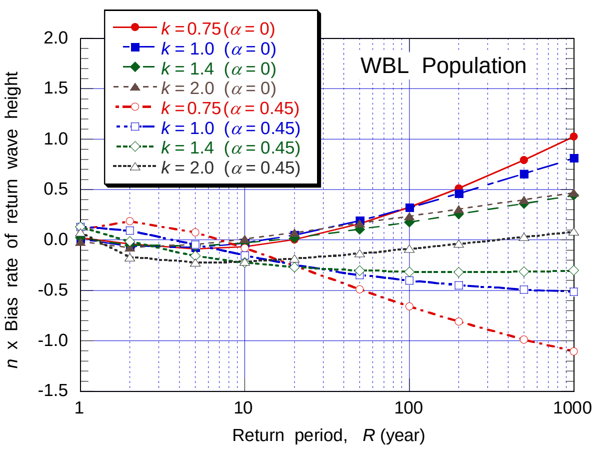
裾長度の影響を 3 種類の分布関数のそれぞれから 1 ケースについて調べた結果が**図-10** である．一般化極値分布はガンベル分布に相当する $k = 0.001$，一般化パレート分布は $k = 0.2$，ワイブル分布は指



(a)　一般化極値分布 (GEV)



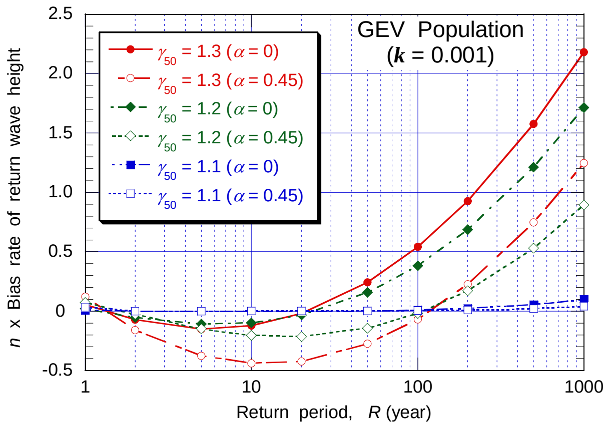
(b)　一般化パレート分布 (GPA)



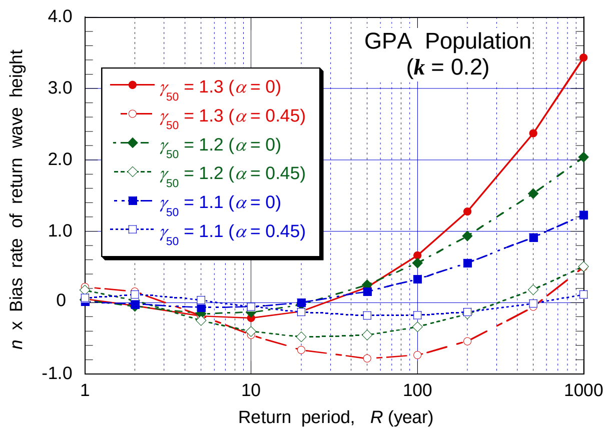
(c)　ワイブル分布

図-9　正規化された確率波高推計値の偏り率（$\gamma_{50} = 1.2$ の場合）

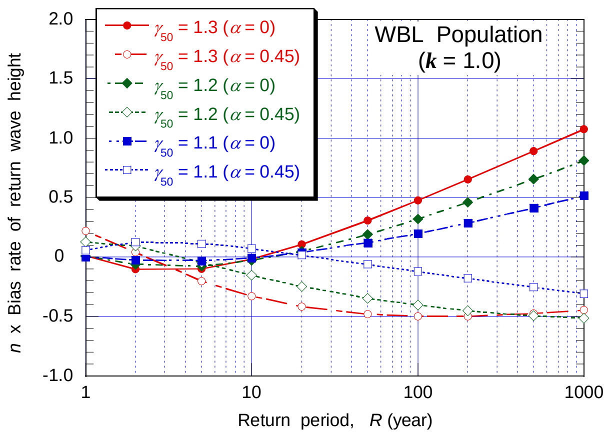
数分布である $k = 1.0$ である．図で明らかなように，裾長度が大きくなるにつれて確率波高の偏り率の絶対値が増大している．



(a)　一般化極値分布 ($k = 0.001$)



(b)　一般化パレート分布 ($k = 0.2$)



(c)　ワイブル分布 ($k = 1.0$)

図-10　正規化された確率波高推計値の偏り率（裾長度の影響の比較）

また，プロッティングポジションの影響をみると，$\alpha = 0.45$ よりも $\alpha = 0$ のほうが偏り率の絶対値が大きいことが確認される．特に，裾長度が $\gamma_{50} = 1.3$ の場合に差異が明瞭である．波高の極値統計の最終目的は，与えられた標本に基づいて所定の再現期間における確率波高を推計することであり，その際には偏り率ができるだけ小さいことが望まれる．その意味で，波高の極値統計解析ではプロッティングポジションとして新しく提案する $\alpha = 0.45$ の定数を使用するのが適切と考えられる．

(2) 最適合分布関数が母関数に的中する割合

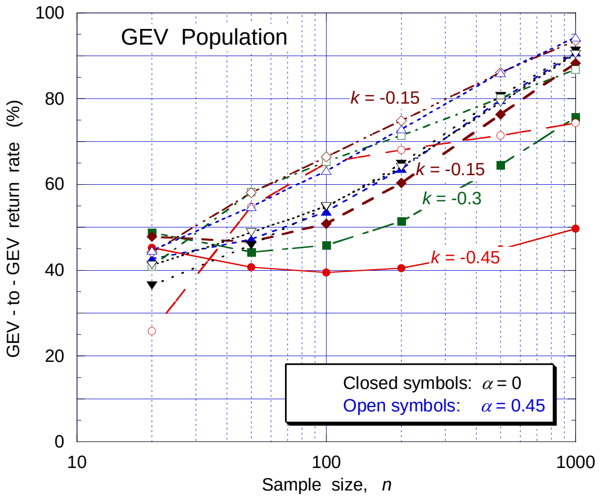
前項では当てはめ分布関数を母集団分布に固定した場合の確率波高の偏り率を記述した．実際の極値統計解析では，標本がどのような母集団から得られたものかが未知であり，複数の分布関数のうちで標本にもっとも適合すると考えられるものを母集団とみなしている．

今回の数値シミュレーションでは一般化極値分布，一般化パレート分布，およびワイブル分布が高波極値の母集団分布の候補であると考え，その中から**最小残差法**に基づく最適合の分布を当てはめ候補として取り上げている．すなわち，3 (3) で述べたように，標本のデータを昇順に並べ替えた値 x_i とその推計値 $\hat{x}_i$ の間の相関係数 $r(x_i, \hat{x}_i)$ を計算し，その上限値である 1 からの残差 $\Delta r = 1 - r(x_i, \hat{x}_i)$ が最小となるものを最適合と判断している．

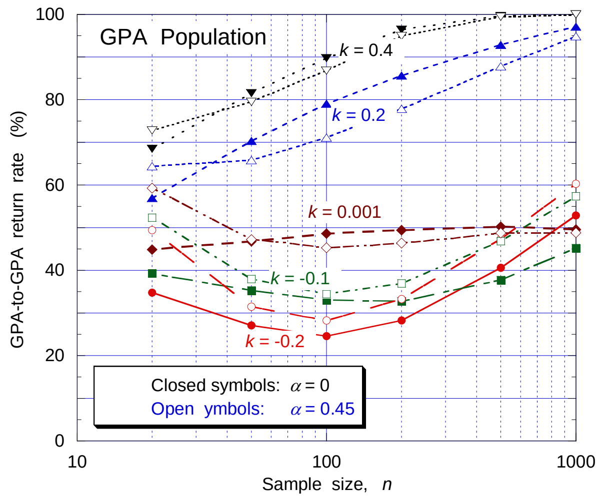
ここで，一つの母集団から得られた極大値の標本に対して，*L*-モーメント法によって最適分布と認定された分布関数が元の母集団分布に合致する割合をここでは当てはめ的中率を呼ぶことにする．この的中率は，標本が大きくなるにつれて向上することが予想され，このことは今回の数値計算の結果でも確認されている．

当てはめ的中率が標本の大きさによって変化する状況を**図-11** に示す．裾長度を 3 通りに変えて計算した結果には有意な差が見られなかったので，当てはめ的中率としては 3 通りの裾長度の結果の平均値を用いている．また，標本の大きさが $n = 10$ と小さい場合には的中率が低くなるため，図には表示していない．

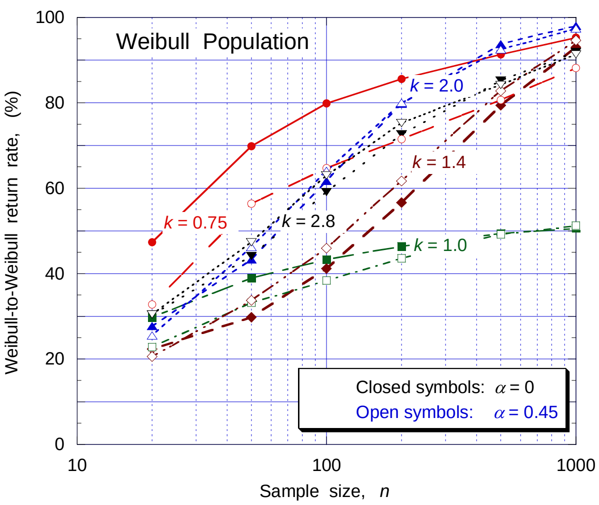
当てはめ的中率は標本が大きくなるにつれて向上する．例外は一般化パレート分布で形状母数の値が負の場合であり，標本の大きさが $n = 100$ の辺りで的中率が 30% 前後に落ち込んでいる．この付近では，一般化パレート分布よりもワイブル分布が最適合と判断される割合が多くなっている．**図-1** の *L*-kurtosis と *L*-skewness の相関図に見られるように，形状母数の値が負のときの一般化パレート



(a) 一般化極値分布 (GEV)



(b) 一般化パレート分布 (GPA)



(c) ワイブル分布

図-11 母集団への当てはめ的中率と標本の大きさの関係

分布はワイブル分布と近接した位置を占めており，標本の大きさが十分でないときには，母集団の識別がむずかしいのであろう．

また，一般化極値分布で形状母数が k = -0.45 のケースでは，プロッティングポジションとして α = 0 を用いたときには，標本が n = 1000 と非常に大きくなっても当てはめ的中率が 50% にとどまる．このときには一般化パレート分布を最適とする割合が 46% に達している．ただし，プロッティングポジションとして α = 0.45 を用いると，標本が大きくなるにつれて一般化極値分布を最適合とする割合が増大し，n = 1000 では的中率が 74% となる．

なお，**図-11**(b) の一般化パレート分布の k = 0.001 のケースと**図-11**(c) のワイブル分布の k = 1.0 のケースは，ともに指数分布を代表している．両ケースとも，標本が大きくなっても的中率が 50% にとどまるのは，この場合には二つの分布関数を識別することが理論的に不可能なためである．

著者は以前に極値統計解析の信頼度手法について，最小 2 乗法を用いた場合の結果を検討した（合田, 1988）．そのときには，ワイブル分布で形状母数を 4 通りに固定したものとガンベル分布を組み合わせた 5 種類の 2 母数型分布を当てはめ対象として適合率をしらべた．今回のように 3 母数型分布の中からの選択でないため状況が異なるけれども，標本の大きさが 100 程度では的中率が 50%～80% であり，ガンベル分布の場合には的中率 53% であって，残りは k = 1.4 あるいは k = 2.0 のワイブル分布のほうによく適合した．

分布関数の当てはめ的中率は最適合の判定基準に依存するので，今回用いた**最小残差法**以外の基準を使うと当てはめ的中率としてやや異なる数値が得られるであろう．しかしながら，標本の大きさが 100 程度未満では，標本の背後に隠れている母分布関数を的確に見つけだすのはかなり難しいといえよう．

(6) 母集団が未知の場合の確率波高推計値の偏り率

実際の極値統計解析では母集団が未知であり，複数の当てはめ候補の中から最適合の分布関数を選んだとしても，それが真の母集団を表しているという保証はない．

今回の数値実験では，高波極値の母集団である可能性の高い 3 種類の分布関数，すなわち一般化

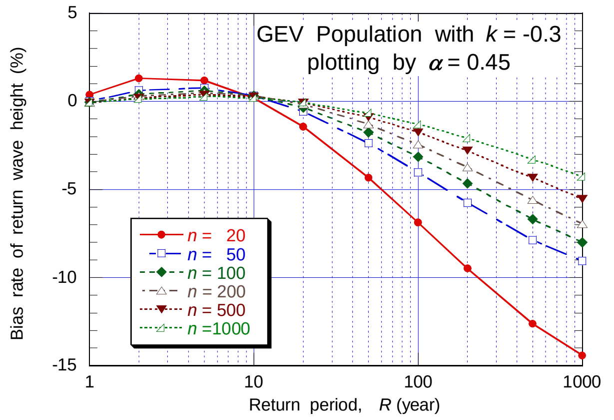
極値分布，一般化パレート分布，およびワイブル分布を母集団の代表として選んでいる．そして，標本毎に最適合分布関数を選び，4 (3) に述べた 10 通りの再現期間について確率波高を推計し，N = 50,000 回の繰り返し計算で得られた標本毎の確率波高推計値の平均値ならびに標準偏差を算定した．そして，推計値の平均について偏り率を計算している．なお，この際に50,000個の標本に当てはめられた分布関数の形状・尺度・位置母数はそれぞれに異なる．

これまでの検討結果から明らかなように，式 (1) のプロッティングポジション式において定数を α = 0 と α = 0.45 に設定したときの推計値の信頼度には有意な差が見られない．しかし，α = 0 の使用には一つの問題が残る．すなわち，式 (1) に $i = n$ を代入すると $P_n = 1$ となり，プロッティングポジションが定められない．このため，データ中の最大値を図面にプロットして当てはめた分布関数と比較することができない．また，最適合関数を定めるための相関係数 $r(x_i, \hat{x}_i)$ を $i = 1 \sim n$ に適用できないため，相関係数の計算は $i = 1 \sim (n - 1)$ に限定して最適合の判定を行なわざるを得ない．

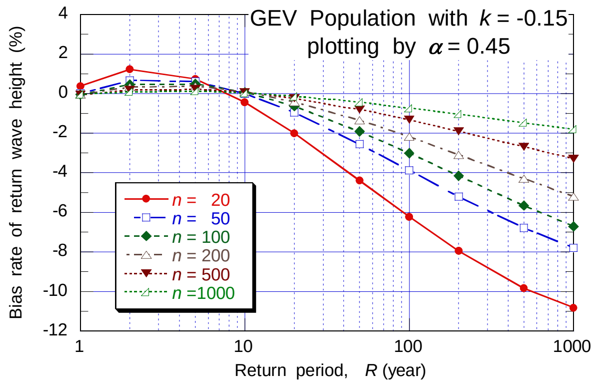
このようなところから，高波の極値統計解析に L-モーメント法を適用する際には，プロッティングポジションとして α = 0.45, β = 0 を使用するのが適切と考えられる．したがって，以下に述べる母集団未知の場合の確率波高の推計値の偏り率ならびに推計値の変動係数の検討においては，すべてこのプロッティングポジション式を用いて計算した結果を使用する．

母集団が未知の場合の確率波高の推計値の偏り率については，裾長度が γ_{50} = 1.2 の場合の計算結果を分布関数毎に**図-12**～**14** に示す．

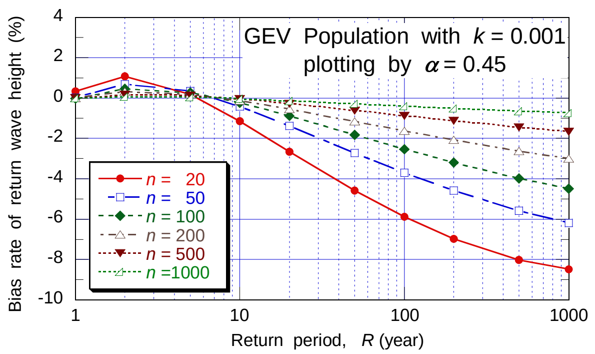
図-12 は一般化極値分布の母集団から抽出された標本に当てはめられた最適合分布から推計された確率波高の偏り率を百分率で表示したもので，形状母数の値毎に 4 枚のグラフとして，標本の大きさ別にプロットしてある．これまでの図表のように標本の大きさ n乗じて正規化したものではない．たとえば，一つの標本に対して形状母数が k = 0 の一般化極値分布が最適合と判定されたとき，標本の大きさが n = 20 であれば 100 年確率波高の推計値が -6% の偏り率，標本の大きさが n = 100 であれば 100 年確率波高の推計値が -2.3% の偏り率を持つと読み取られる．



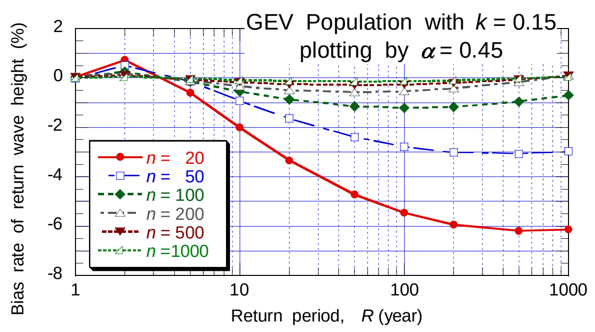
(a) k = -0.3 のケース



(b) k = -0.15 のケース

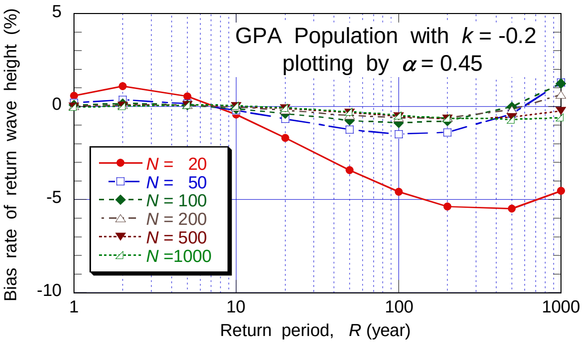


(c) k = 0.001 のケース

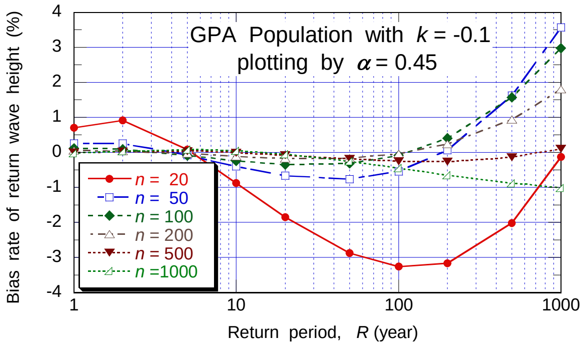


(d) k = 0.15のケース

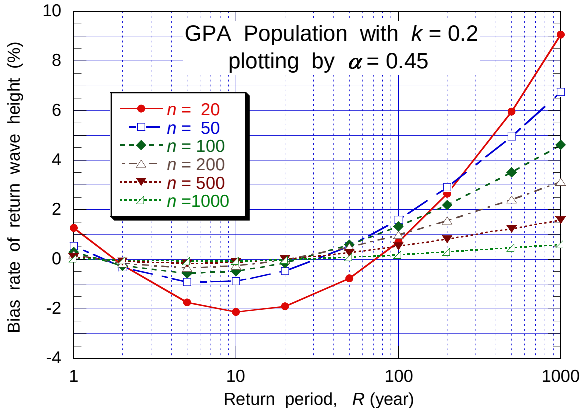
図-12 母集団が未知の場合に一般化極値分布の標本への最適合分布から推計された確率波高の偏り率 (γ_{50} = 1,2, α = 0.45)



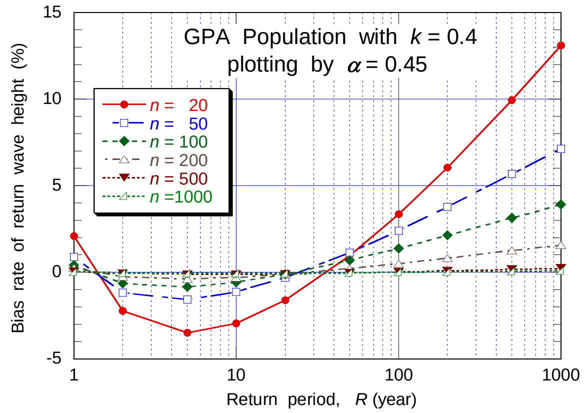
(a) k = -0.2 のケース



(b) k = -0.1 のケース

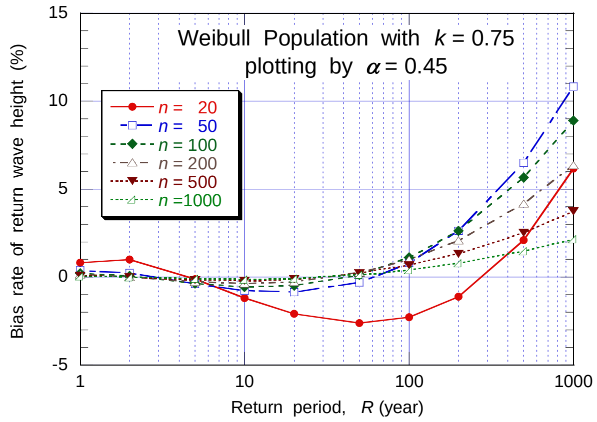


(c) k = 0.2 のケース

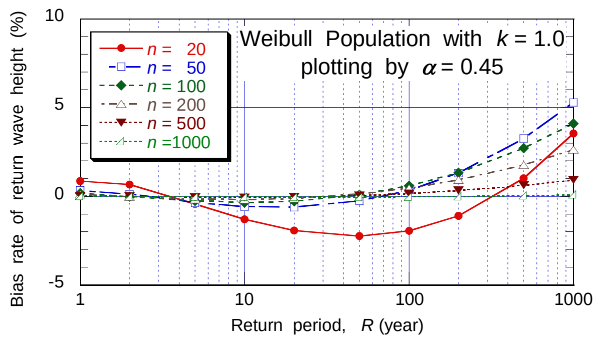


(d) k = 0.4のケース

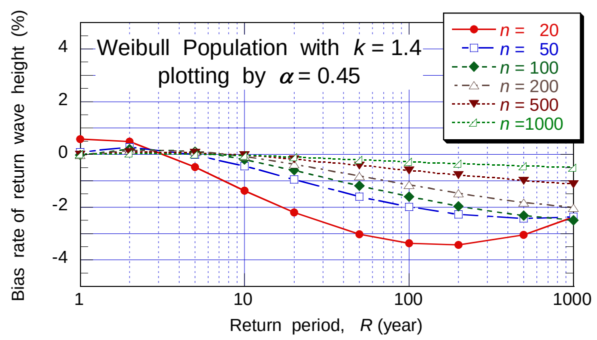
図-13 母集団が未知の場合に一般化パレート分布の標本への最適合分布から推計された確率波高の偏り率 (γ_{50} = 1,2, α = 0.45)



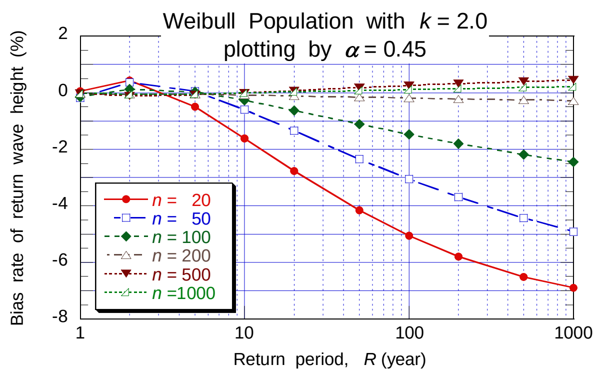
(a) k = 0.75 のケース



(b) k = 1.0 のケース



(c) k = 1.4 のケース



(d) k = 2.0 のケース

図-14 母集団が未知の場合にワイブル分布の標本への最適合分布から推計された確率波高の偏り率 (γ_{50} = 1,2, α = 0.45)

一般化極値分布から抽出された標本の場合には，確率波高の偏り率は負の値であるのが大半であるが，一般化パレート分布から抽出された標本の場合には**図-13** に示すように偏り率が正の値のことが多い．

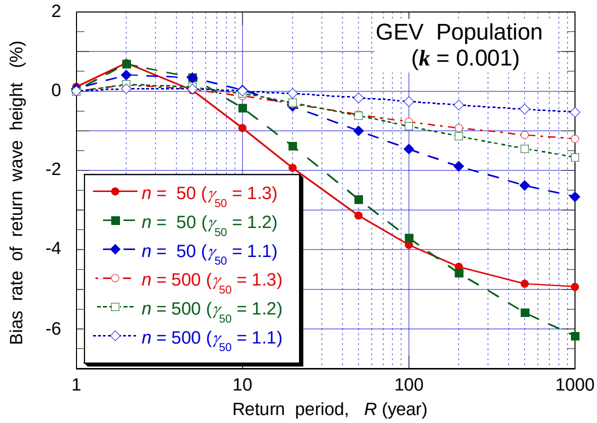
ワイブル分布から抽出された標本について，母集団が未知であるとして最適合の分布関数から確率波高を推計した場合の偏り率は，**図-14** に示すとおりである．偏り率の絶対値は，一般化極値分布および一般化パレート分布の場合よりもやや小さめであるものの，標本の大きさならびに再現期間の長さによる確率波高推計値の偏り率変化の様相は類似している．

こうした母集団が未知の場合の確率波高の偏り率は裾長度によっても影響される．**図-15** は特定のケースについて裾長度の影響を示したもので，(a) は一般化極値分布の $k = 0.001$ のケース，(b) は一般化パレート分布の $k = 0.2$ のケース，(c) はワイブル分布の $k = 1.0$ のケースである．図の混乱を避けるため，標本の大きさが $n = 50$ および 500 の場合を示している．

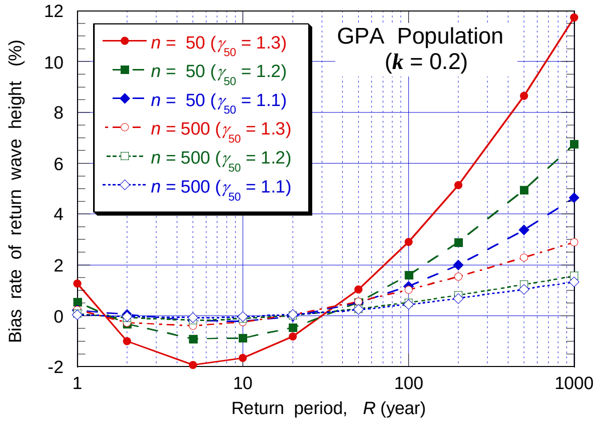
図-15 に示した例では，裾長度が $\gamma_{50} = 1.1$ から 1.3 へと大きくなるにつれて，確率波高の偏り率が増大している．特に，標本の大きさが $n = 50$ のときに裾長度の影響が顕著に認められる．ただし，**図-15**(a) の一般化極値分布の場合には裾長度が $\gamma_{50} = 1.2$ のときよりも $\gamma_{50} = 1.3$ のときの偏り率がやや小さめになっている．

極値統計解析の実務においては，今回の数値実験のように与えられた標本の母集団を知ることができない．したがって，厳密にいえば**図-12〜14** に提示した情報が使えるわけではない．また，裾長度の影響も定式化がむずかしい．しかし，波高の極値統計解析に限っていえば，高波の極値分布関数の裾長度は平均的に1.2程度であるので，まずは**図-12〜14** の偏り率を参考にすることができよう．すなわち，標本に対して最適合と判定された分布関数に応じてこれらの図表から偏り率を読み取り，確率波高の推計値を補正するのが良いと思われる．

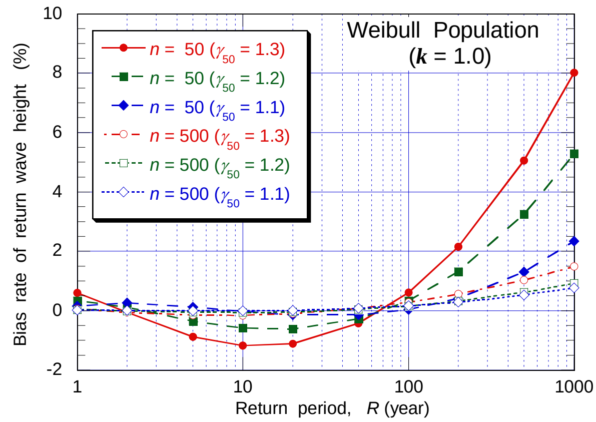
なお，**図-12〜15** では横軸を再現期間としているが，これは平均発生率を $\lambda = 2$ に設定したときの非超過確率から換算したものである．非超過確率は式 (37) で算定するので，たとえば平均発生率が $\lambda = 20$ であれば，そのときの再現期間100年はこれらの図表の再現期間1000年に相当する．



(a) 一般化極値分布 ($k = 0.001$)



(b) 一般化パレート分布 ($k = 0.2$)



(c) ワイブル分布 ($k = 1.0$)

図-15 母集団が未知の場合の確率波高の偏り率に及ぼす裾長度の影響

図-12〜15 での結果でも分かるように，確率波高の偏り率はかなり大きなものである．偏り率を下げるためには，十分に大きな標本を使うことが不可欠である．少なくてもデータ数が 100 以上の標本を使うことが望ましいといえる．

８．確率波高の推計値の変動係数

極値統計においては，推定された確率波高の絶対値だけでなく，その信頼区間も必要な情報である．今回の数値計算では 4(3) に述べたように，10 通りの再現期間毎に 50,000 個の標本から求めた確率波高の平均値と標準偏差を算出している．この結果に基づいて，確率波高の変動係数を次のように定義した．

$$CV[H_R]=\frac{\sigma[H_R]}{\mu[H_R]}=f(k,n,\gamma_{50},y_R) \qquad (47)$$

ここに，$\sigma[H_R]$ と $\mu[H_R]$は再現期間 R 年の確率波高の標準偏差と平均値，y_R は R 年に対する基準化変量である．式 (47) は，変動係数が形状母数 k，標本の大きさ n，裾長度 γ_{50}，および基準化変量 y_R の関数であることを表している．この変動係数の大きさについては，プロッテイングポジション式による違いがほとんど認められなかった．このため，以下においては式 (1) に $\alpha=0.45, \beta=0$ を代入した新提案のプロッティングポジション式を用いて得られた結果について記述する．

式 (47) の右辺に導入されている標本の大きさの影響については，標準偏差あるいは 2 乗平均誤差と同じように $\sqrt{n}$ にほぼ逆比例することが認められた．また，裾長度の影響に関しては大略の傾向として $(\gamma_{50}-1)^{0.7}$ に比例することが認められた．そこで，変動係数に対して次のような基準化変量の 2 次式を当てはめることを試みた．

$$CV[H_R]=(ay_R^2-by_R+c)(\gamma_{50}-1)^{0.7}/\sqrt{n} \qquad (48)$$

そして，母集団の分布関数の形状母数毎に定数 a, b, c を算定し，さらにその結果を形状母数の関数として表示することを試みた．これによって，分布関数毎に**表-8** のような経験式を得ることができた．

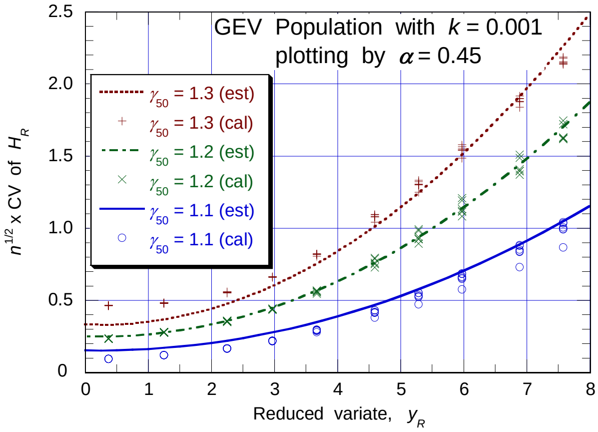
表-8　確率波高の変動係数の近似推定式の係数表

分布関数	係数の推定式
一般化極値分布 (GEV)	$a = 2.7037k^3 + 2.4351k^2 + 0.7325k + 0.0830$ $b = 10.953k^3 + 10.509k^2 + 3.1359k + 0.0372$ $c = 11.311k^3 + 12.526k^2 + 5.6088k + 0.7719$
一般化パレート分布 (GPA)	$a = \exp[-5.3844k^3 - 0.6089k^2 + 7.7316k - 1.9805]$ $b = -22.438k^3 + 24.678k^2 + 6.1046k + 0.1062$ $c = -28.287k^3 + 18.860k^2 + 6.9797k + 0.6776$
ワイブル分布	$a = 0.9742(\ln k)^2 + 0.6949(\ln k) + 0.1517$ $b = 1.9049(\ln k)^2 + 1.8142(\ln k) + 0.3677$ $c = 0.6349(\ln k)^2 + 1.8561(\ln k) + 0.9095$

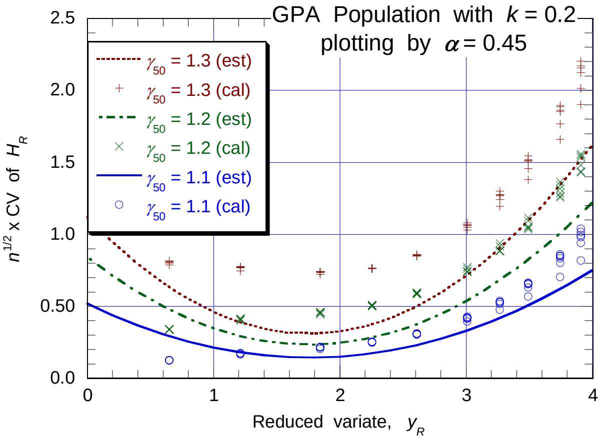
これらの経験式による確率波高推計値の変動係数の予測値を数値実験で得られたデータと比較した結果の例を**図-16** に示す．ここでは，一般化極値分布 ($k=0.001$)，一般化パレート分布 ($k=0.2$)，およびワイブル分布 ($k=1.0$) のケースを示している．横軸は式 (34)～(36) で定義される基準化変量であり，連続した曲線が式 (48) に**表-8** で計算される係数値を代入して求めた予測値である．記号で示しているデータは，標本の大きさが $n=20$～1000 の 6 通りの計算結果であり，3 種類の裾長度ごとに記号を変えて示している．

まず，**図-16**(a) の一般化極値分布の場合には，数値実験によるデータは式 (48) の予測値の周辺に散らばっており，予測式はほぼ妥当である．なお，同一の基準化変量の値でデータが上下に散らばっているのは，変動係数が標本の大きさの –1/2 乗に比例するとの仮定からのずれによるものである．

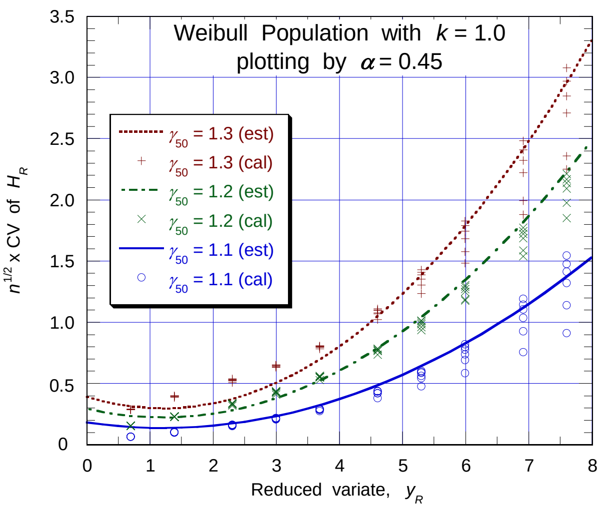
次の**図-16**(b) の一般化パレート分布の場合には，予測値は数値計算データをかなり下回っており，適合度はあまりよくない．この分布関数では，形状母数が大きくなったときに式(48) のような 2 次曲線を当てはめると，基準化変量が小さくなるにつれて変動係数が増加する傾向が過大に表現されてしまう．当てはめる近似曲線を変えることによって適合度が向上するかもしれないけれども，3 種類の分布関数のすべてについて広範囲の形状母数に共通して適用できるような近似式を見いだすことはむずかしい．このケースは，近似式の適用限界を示す資料として参照すべきであろう．それでも，数値計算データは裾長度の増大に応じて変動係数も大きくなっており，その傾向は予測値でも同様である．すなわち，式(48) に組み込まれている $(\gamma_{50}-1)^{0.7}$ の項は裾長度の効果を適切にほぼ表現しているといえる．



(a) 一般化極値分布 (k=0.001)



(b) 一般化パレート分布 (k=0.2)



(c) ワイブル分布 (k=1.0)

図-16 母集団が未知の場合に標本への最適合分布から推計された確率波高の変動係数

最後の**図-16**(c)のワイブル分布の場合は，**図-16**(a)の一般化極値分布と同じ程度に数値実験データが予測曲線の周りに散らばっている．裾長度がγ_{50}=1.1のデータは標本の大きさによるばらつきが大きい．これは，標本の大きさの影響を$n^{-1/2}$で近似したことによる誤差である．

ここに示した**表-8**の係数を用いた式(48)の推定式は，確率波高推計値の信頼区間の概略値を求めるためのものであるので，**図-16**に見られるような推定式とデータの差異は，実用的には許容される範囲と言えよう．実際の極値統計解析では，与えられた標本に対して最適合の分布関数を定め，所定の再現期間に対する確率波高を推計する．それに対する信頼区間はその概略値を式(48)と**表-8**を用いて算定し，さらに詳しくは推計された分布関数の母数を使ってモンテカルロ法による数値シミュレーションを実施して信頼区間を見いだすことが望ましい．

ここで，確率波高の推計値の信頼区間の推定の試算例を示す．いま，大きさがN = 120で平均発生率がλ = 2の極大値波高の標本について統計解析を行い，形状母数がk = 1.2のワイブル分布が最適合と判定されたとする．尺度・位置母数はそれぞれA =1.69 m, B =4.47 mであった．この分布が母集団であるとみなして，再現期間R = 100年に対する確率波高とその信頼区間を試算してみる．

まず，この分布関数における確率波高と裾長度を求めておく必要がある．再現期間10, 50, および100年に対する非超過確率は，式(37)によってそれぞれP_{10} = 1 – 1/(2×10) = 0.95, P_{50} = 1 – 1/(2×50) = 0.99, P_{100} = 1 – 1/(2×100) = 0.995である．したがって，基準化変量は式(36)でy_{10} = $\{-\ln(1-0.95)\}^{1/1.2}$ = 2.495, y_{50} = $\{-\ln(1-0.99)\}^{1/1.2}$ = 3.570, y_{100} = $\{-\ln(1-0.995)\}^{1/1.2}$ = 4.013 となるので，確率波高はそれぞれH_{10} = 4.47 + 1.69×2.495 = 8.687 m, H_{50} = 4.47 + 1.69×3.570 = 10.503 m, H_{100} = 4.47 + 1.69×4.013 = 11.25 mである．これによって，裾長度はγ_{50} = H_{50}/H_{10} = 10.503/8.687 = 1.209 と算定される．

確率波高の推計値の変動係数を求めるための定数は，**表-8**中の推定式によって次のように計算される．

$$\begin{cases} a = 0.974(\ln 1.2)^2 + 0.695\ln 1.2 + 0.152 = 0.311 \\ b = 1.905(\ln 1.2)^2 + 1.814\ln 1.2 + 0.368 = 0.762 \\ c = 0.635(\ln 1.2)^2 + 1.856\ln 1.2 + 0.910 = 1.269 \end{cases}$$

したがって，変動係数は次のように計算される．

$$CV[H_R] = (0.311 \times 4.013^2 - 0.762 \times 4.013 + 1.269) \times (1.209 - 1.0)^{0.7} / \sqrt{120} = 0.0982$$

すなわち，100年確率波高の標準偏差は11.25×0.0982 = 1.105 mである．90%信頼区間として標準偏差の1.64倍を用いると，この条件における100年確率波高H_{100} = 11.25 mの信頼区間が9.44〜13.06 mと推定される．

９．*L*−モーメント法のプロッティングポジション等に関する既往研究との比較

(1) Unbiased Positionについて

最初に**まえがき**で紹介したように，式(1)のプロッティングポジション式で定数として$\alpha=0, \beta=0$の組合せを推奨したのはHosking · Wallis (1997)の解説書(33-34頁)である．その根拠としているのはHosking · Wallis (1995)の論文である．

この論文では，Landwehrほか(1979b)が用いた$\alpha=0.35, \beta=0$の組合せを“plotting position estimator”(以下ではPPと略記)と呼び，$\alpha=0, \beta=0$の組合せを“unbiased estimator”と呼称している．後者は厳密にはunbiasedでないけれども，*L*-モーメントの計算式の構成に基づいてunbiasedとして引用するのが便利であるとしている．

Hosking · Wallis (1995)がPPが適切でないとしている一つの理由は，PPを使うと標本のデータ座標軸を移動（各データにある一定値を加除）したときにt_2, t_3, t_4などの*L*-モーメントおよび*L*-モーメント比の値が変化する，すなわち不変性が保持されないということである．しかし，データの座標軸を移動することの意義が不明であり，それに対する*L*-モーメントの不変性を保持すべき理由も理解しにくい．波浪の極値統計では4(1)で説明したように，代表波高の絶対値によって尺度母数や位置母数が規定されるので，データの座標軸の移動という操作はまず考えられない．

第2の理由として，カッパ分布から抽出したときのτ_3, τ_4の*L*-skewness, *L*-kurtosisの平均値の偏り（母集団の真値との差）について数値実験を行ったところ，$n=20$の小標本ではPPのほうがUnbiasedよりもかなり大きな偏り量を得たこと挙げている．この数値実験では，母集団の諸元を$\lambda_1=1, \lambda_2=0.2$に固定し，$\tau_3$は0〜0.5の間を0.05間隔，$\tau_4$は0〜0.3の間を0.05間隔で変化させている．カッパ分布は，尺度・位置母数の他に分布形状に係わる二つの母数を持つ4母数分布であり，τ_3, τ_4の変化に応じて分布形状も大幅に変化していたと想定される．

今回の数値実験においても，*L*-モーメント比の推定に関して$\alpha=0, \beta=0$のUnbiased Positionが極めて優れた不偏性を示しており，Hosking · Wallis (1995)の結論を裏付けている．

Hosking · Wallis (1995)は確率波高の偏り率についても幾つかの検討結果を示している．形状母数$k=-0.2$のGEV母集団($A=0.231$, $B=0.81$)から抽出された$n=50$の標本に関し，非超過確率0.999に対する確率波高(quantile)の偏り率がPPでは3.5%，unbiasedでは4.5%の例を示している．（今回の数値実験は母数値が異なるので直接に比較できないけれども，同じオーダーの偏り率が得られている．）また，高水流量の頻度解析データの事例として地点数21，地点毎のデータ数が10〜30のケースについてPPとunbiasedの偏り率を比較している．その結果の表から計算すると，非超過確率0.99および0.999においてPPの偏り率が0.6%と1.2%，unbiasedによる偏り率が1.1%と1.5%であり，PPの偏り率のほうが低くなっている．

こうした計算結果からHosking · Wallis (1995)は，単一地点の標本について確率波高などを推定する際にはPPに若干の優位性があると認めている．今回新しく提案した$\alpha=0.45, \beta=0$のプロッティングポジション式は7(1)で示したように，Unbiased Positionよりも確率波高の推計値の偏り率が小さい．また，*L*-モーメント比や形状母数の推計においても，Unbiased Positionよりも偏り率がやや大きいものの，実用的には無視できる量である．

(2) GEV分布の偏り量の漸近理論値問題

L-モーメント法の母体であるProbability Weighted Moments (PWM：確率重み付き積率法)では，母数推計値の偏り量と分散についての漸近理論値が利用でき，Hosking et al. (1985)が紹介している．理論値はグラフの形で掲載されているので，図から読み取った概略値で検討する．

形状母数の偏り量(Bias)は，それに標本の大きさnを乗じた値が母集団の形状母数の値が大きくなるにつれて減少する．[n×Bias]の値は，$k=-0.3$, -0.15, 0, および0.15 (t3 = 0.38, 0.27, 0.17, および0.08)でそれぞれ約2.0, 0.9, 0.4, 0.2である．これは今回の数値実験で得られた**図-7**とほとんど同じ値であり，数値シミュレーションの信頼度を保証するものである．

また，形状母数の偏り量の分散については，標本の大きさを乗じた値が$k=-0.3$, -0.15, 0, 0.15でそれぞれ約1.43, 0.75, 0.55, 0.48である．これを標準偏差と標本の大きさの平方根の積に換算すると，それぞれ1.2, 0.87, 0.74, 0.69となり，**図-8**の結果とほぼ一致する．

なお，Hoskingほか(1985)の論文では一般化極値分布において尺度・位置母数をそれぞれ$A=1, B=0$に設定し，形状母数を幾つかに選んだときの確率波高の推計値の偏り量や分散の漸近理論値が示されている．さらに，尺度・位置母数の推計値の偏り量や分散についての数値実験結果も示している．

ただし，尺度・位置母数は次元量であり，たとえば$k=0$のときには$H_{mean}=0.5772$，平均発生率を$\lambda=1$とすると$H_{10}=1.195, H_{50}=1.592$であり，裾長度は$\gamma_{50}=1.33$と高

波の極値としては例を見ない大きな値である．PWM法やL-モーメント法を扱った研究では$A=1, B=0$にに設定しても一般性を失わないとする論文 (たとえば，Landwehrほか[1979], Hoskingほか[1985]) をみかけるけれども，裾長度の観点からはそうした取り扱いには一般性を認めることはできない．したがって，Hoskingほか(1985) の論文に記載されている確率波高および分布関数の尺度・位置母数の推計値の偏り量について比較検討することは困難である．

１０．高波の極値統計解析に*L*-モーメント法を適用する際の手順について

(1) 単一地点の資料の解析

L-モーメント法による極値統計解析の手順については3. で述べたけれども，高波資料の標本に対する解析が行われ，所定の再現期間に対する確率波高が推計された後の処理について補足する．

これまでに述べてきたように，極値統計解析で得られた推計値は常に真値からのなにがしかの偏りを伴っており，またある幅の分散を持つ．高波の母集団が既知であったとしても真値そのものは把握不能である．まして母集団自体が不確定な現時点においては，本研究で取り扱ったように，一般化極値分布，一般化パレート分布，およびワイブル分布を母集団の候補として考え，そのうちで標本に最も適合した分布関数を母集団とみなすことになる．

この場合の確率波高の偏り率は**図-12**～**14**の図表を参照し，分布関数毎に形状母数と標本の大きさに応じた適切な内挿によって推定する．これらの図表は裾長度が$\gamma_{50}=1.2$に対するものなので，裾長度がこれから外れる場合には一つの参考値として扱わざるを得ない．

確率波高の推計値の信頼区間については，まず裾長度を式 (38) によって算定する．そして，式 (48) と**表-8**を用いて確率波高の変動係数を計算し，それから標準偏差を推計する．確率波高の推計値が正規分布に従うと仮定することにより，90%信頼区間であれば推計値の上下に標準偏差の1.64倍を取った幅として定められる．標準偏差の1倍であれば，68% 信頼区間となる．当然ながら，この信頼区間は推計値の偏り量だけ補正したものである．なお，信頼区間の試算例は8. で紹介している．

(2) 複数地点の資料に対する地域頻度解析

地域頻度解析は，均質とみなされる地域内の複数地点の資料を共通に取り扱い，その地域の極値の母集団推定の信頼度を高めるとともに，全体としてのデータ数を使うことによって推計値の信頼区間の幅を狭めることを目的としている．Hosking · Wallis (1997) による地域頻度解析の手順は次の通りである．

1) 地理的な位置その他を勘案し，対象とする地域を選定してその地域内の極値データを収集する．
2) 各地点のデータをその平均値で除し，無次元化する．
3) 無次元化された標本について，単一地点と同じ手順で*L*-モーメント法による解析を行う．
4) 各地点の*L*-CV, *L*-skewness, および*L*-kurtosisの値について各地点のデータ数を重みとする加重平均を行い，地域としての3種類の*L*-モーメント比を算出する．
5) 各地点の3種類の*L*-モーメント比の地域全体の値（重心）からの距離を計算し，その距離が著しく大きければその地点は不整合であるとして当該地域から除外する．これはDiscordance Test (不整合性テスト)と名付けられており，その作業指針はHosking · Wallis(1997) が示している．
6) さらに，各地点の3種類の*L*-モーメント比を用いて地域内の標本の均質性をチェックし，地域の均質性が確保できないときには幾つかの地点を除外あるいは入れ替えて組み直す．これは Heterogeneity Test (不均質性テスト) と名付けられている．
7) 均質性が確認された地域について，地域全体の3種類の*L*-モーメント比を使って地域共通の分布関数を選定する．
8) 地点毎の極値分布関数は，形状母数は地域共通の値を用い，尺度母数と位置母数は地域共通の母数値に各地点の平均値を乗じて換算する．
9) 地点毎に換算された分布関数の母数を使い，各種の再現期間に対する確率波高を算定する．
10) 確率波高の信頼区間等については，地域共通の分布関数を使って数値シミュレーションを行い，変動範囲を調べて推定する．

(3) 新たな不均質性指標の提案

Hosking · Wallis (1997) は不均質性指標として3種類の指標を提示している．

$$H=\frac{(V-\mu_V)}{\sigma_V} \tag{49}$$

上式中のVは次式で定義される量であり，μ_Vとσ_Vはカッパ分布を当てはめて行う数値シミュレーションによって求められるVの平均値と標準偏差である．

$$V=\left\{\frac{\sum_{i=1_i}^{N} n_i\left(t^{(i)}-t^{R}\right)^2}{\sum_{i=1_i}^{N} n_i}\right\}^{1/2} \tag{49}$$

ここに，Nは地点数，n_i はi番目の地点のデータ数，$t^{(i)}$はその地点のL-モーメント比に係わる統計量，t^{R}はその統計量に関してカッパ分布当てはめのシミュレーションで得られた地域平均の値である．

不均質性指標は，3種類の統計量について検討する．第1は，2次のL-モーメントと1次のL-モーメントの比であるL-CV [t] である．すなわち，ある地点のtの値が他よりも著しくかけ離れていると，その地域は不均質と見なされる．第2の統計量としては，L-CVとL-skewnessの相関プロット図において，ある地点が全体の重心からどれだけ離れているかを検討する．第3の統計量としては，L-kurtosisとL-skewnessの相関プロット図における重心からの距離を用いる．なお，第2と第3の不均質指標については，Vとして2次元座標上の距離を用いて算定する．

地域が均質であるか否かを判断する不均質性指標の閾値としてHosking · Wallis (1997) はHが1以下であれば均質性に問題はなく，Hが2を超えると不均質である公算が大きいとしている．

表-9　解析対象地点のデータ個数とL−モーメント比

地点名	データ個数, n	L-CV	L-skew	L-kurt
留　萌	554	0.0964	0.2477	0.1183
瀬　棚	330	0.0927	0.2749	0.1393
深　浦	408	0.1007	0.2780	0.1434
秋　田	314	0.1067	0.2663	0.1323
酒　田	518	0.1096	0.2589	0.1299
新潟沖	252	0.1061	0.2973	0.1333
輪　島	298	0.0826	0.2514	0.1306
金　沢	387	0.0961	0.2241	0.1002
福　井	248	0.1118	0.2574	0.1237
鳥　取	320	0.0898	0.2952	0.1534
浜　田	400	0.0917	0.2772	0.1326
全地域	4029	0.09856	0.26428	0.12978

著者らが先に行った日本海沿岸の高波の地域頻度解析 (合田ら, 2010) では，3通りの不均質性指標のうちでL-CVに係わるH(1) が最もシビアであり，**表-9**に示す留萌から浜田までの11地点全体ではH(1) = 5.27となって，これらを一つの地域に含めることは不可との判断が下された．また，輪島から浜田までの 5 地点をまとめて取り扱っても不均質性指標がH(1) = 4.49と高い値であり，一つの地域に含めることができないと判断された．しかしながら，11地点のL-CVには輪島の0.0826から福井の0.1118までの狭い範囲に入っており，これでこの地域が不均質とみなされるのかについては疑念が抱かれる．

Hosking · Wallis (1997) の不均質性指標は4母数型の極値分布であるカッパ分布の数値シミュレーションに基づいているけれども，カッパ分布を導入する意義は明瞭でない．

むしろ，今回の数値実験によってL-モーメント比の変動係数の推定式が**表-6**のようにとりまとめられたので，これを用いて各地点のL-モーメント比の標準偏差を算定する．そして，各地点のL-モーメント比の地域平均値からの差をこの標準偏差で除したものを偏心率と名付け，地点毎の偏心率の分散の度合いが正規分布に近ければその地域を均質とみなし，偏心率の分散が正規分布から外れていれば不均質とみなすことを提案する．

偏心率は，L-CV, L-skewness, およびL-kurtosisについては次式で計算する．

$$\psi_{k,i}=[\zeta_{k,i}-\mu(\zeta_k)]/\sigma(\zeta_k) \quad : k=1,2,3;\quad i=1,2,\ldots,N \tag{51}$$

ここに，$\zeta_1, \zeta_2, \zeta_3$はそれぞれ$L$-CV[$t$], L-skewness[t_3], およびL-kurtosis[t_4]を代表する．さらに，2次元座標上の距離に関する偏心率として次の二つを定義する．

$$\psi_{4,i}=\left[\frac{[\zeta_{1,i}-\mu(\zeta_1)]^2+[\zeta_{2,i}-\mu(\zeta_2)]^2}{\sigma^2(\zeta_1)+\sigma^2(\zeta_2)}\right]^{1/2} \tag{52}$$

$$\psi_{5,i}=\left[\frac{[\zeta_{2,i}-\mu(\zeta_2)]^2+[\zeta_{3,i}-\mu(\zeta_3)]^2}{\sigma^2(\zeta_2)+\sigma^2(\zeta_3)}\right]^{1/2} \tag{53}$$

式(52) のψ_4はL-CVとL-skewnessの相関プロット図上の偏心率，ψ_5はL-kurtosisとL-skewnessの相関プロット図上の偏心率を表す.

こうした各地点の偏心率の分散の度合いを正規分布と比べるには，偏心率の最大値を用いるのが簡便である．すなわち，

$$\Psi_k=\max\{|\psi_{k,i}|: i=1,2,\ldots,n\}: k=1,2,\cdots,5 \tag{54}$$

正規分布であれば，偏心率の最大値が1.6を超える確率は11%，2.0を超える確率は4.6%であるので，5種類の偏心率について正規分布と比較しながら均質性を調べることができる．

表-9の日本海沿岸11地点の高波資料について，母集団がワイブル分布であるとしてL-モーメント比に関する偏心率を計算すると，**表-10**のような結果が得られる．日本海全域を一つの地域とみなした場合には，**表-10**に示すようにL-CVに関する偏心率Ψ_1が輪島で最大値3.52となり，やはり全域を一つの地域とみなすことが無理であることが示される．

表-10　日本海沿岸の高波の均質性指標（全域共通）

地点名	不均質性指標としての偏心率				
	Ψ_1 (*L*-CV)	Ψ_2 (*L*-skew)	Ψ_3 (L-krt)	Ψ_4 (CV-skw)	Ψ_5 (krt-skw)
留　萌	-0.677	-0.834	-0.801	0.823	0.823
瀬　棚	-1.399	0.420	0.438	0.428	0.428
深　浦	0.522	0.604	0.675	0.636	0.636
秋　田	1.682	0.079	0.117	0.096	0.096
酒　田	2.927	-0.266	0.008	0.208	0.208
新潟沖	1.307	1.159	0.150	0.897	0.897
輪　島	-4.314	-0.475	0.038	0.372	0.372
金　沢	-0.696	-1.666	-2.014	1.767	1.767
福　井	2.387	-0.234	-0.273	0.249	0.249
鳥　取	-2.018	1.224	0.980	1.114	1.114
浜　田	-1.824	0.561	0.147	0.444	0.444
最大値	4.314	1.666	2.014	1.642	1.767

表-11　日本海沿岸の高波の均質性指標（南北２分割）

地点名	不均質性指標としての偏心率				
	Ψ_1 (*L*-CV)	Ψ_2 (*L*-skew)	Ψ_3 (L-krt)	Ψ_4 (CV-skw)	Ψ_5 (krt-skw)
留　萌	-1.745	-0.961	-0.919	0.988	0.947
瀬　棚	-2.203	0.319	0.360	0.477	0.337
深　浦	-0.294	0.492	0.591	0.487	0.538
秋　田	0.986	-0.019	0.037	0.180	0.027
酒　田	2.042	-0.390	-0.096	0.534	0.311
新潟沖	0.729	1.070	0.078	1.058	0.825
北最大値	2.203	1.070	0.919	1.058	0.825
輪　島	-3.010	-0.340	0.151	0.528	0.281
金　沢	0.657	-1.514	-1.848	1.501	1.611
福　井	3.253	-0.110	-0.164	0.611	0.133
鳥　取	-0.905	1.368	1.081	1.357	1.240
浜　田	-0.546	0.721	0.277	0.717	0.585
南最大値	3.253	1.514	1.848	1.501	1.611

このため，日本海沿岸の11地点を留萌から新潟東までの北地域6地点，輪島から浜田までの南地域5地点に分割して偏心率を計算すると，**表-11**のような結果が得られる．これによれば，北地域に関してはΨ_1～Ψ_5のどの最大偏心率も1.8以下となり，均質性に問題はないと判断される．ただし，南地域に関してはΨ_1が金沢で2.65となり，均質性についての疑問が残る．ただし，Ψ_2～Ψ_5に関してはいずれもも1.8以下であるので，南地域を一体として解析することも許容されると思われる．

式(51)～(54)による不均質性指標の一つの利点は，*L*-モーメント比の値が地域内の他地点からかけ離れた地点を直ちに特定できることである．Hosking · Wallis (1997)の不均質性指標の場合には，全体として均質性に問題があっても，その是正は地点の入れ替えを試行錯誤によって行わなければならず，問題地点の抽出に手間取ることが多い．その点，新しく提案する不均質指標としての偏心率であれば，地域内の値手の組み替えが容易に行える．

3種類の*L*-モーメント比の中では，*L*-skewnessが母集団の形状母数を規定するだけでなく，**表-6～8** にまとめた各種の変動係数・標準偏差の独立変数として機能している．その意味で，これに関する最大偏心率Ψ_2を最も重視すべきではないかと思われる．今後はさらに多くの極値データに適用し，最大偏心率の許容限界について検討したいと考えている．

１０．まとめと今後の課題

今回のモンテカルロ法による一連の数値シミュレーションによって明らかになった諸点をまとめると次のようになる．

1) *L*-モーメント法においてもプロッティングポジションの選定は重要な課題であり，不適切なものを用いると分布関数の母数や確率波高に大きな偏り (Bias) をもたらす．
2) Hosking · Wallis (1977) が“unbiased position”と称して推奨するプロッティングポジション式は，*L*-モーメントや分布関数の母数の推計値の偏りがきわめて小さいことが確認された．
3) ガンベル分布と正規分布に対して最小2乗法を用いるときに有効であったGringorten式とBlom式は*L*-モーメント法に用いると不偏性を満足しなかった．したがって，この2式は*L*-モーメント法には推奨できない．
4) 今回の数値実験で導入した新しいプロッティングポジション式(α=0.45, β=0) は，偏りが僅少であり，“unbiased position”と同様に**不偏性**をほぼ満足する．また，確率波高推計値の偏りが“unbiased position”よりもやや小さいので，*L*-モーメント法のプロッティングポジション式としてこれを推奨する．
5) *L*-モーメント法であっても標本が小さいと母集団以外の分布関数に当てはめられる割合が大きくなり，また母数値や確率波高の信頼区間の幅が非常に大きくなる．標本のデータ個数20では不十分であり，できれば100以上が望ましい．
6) 極値分布関数が一般化極値分布 (GEV)，一般化パレート分布 (GPA)，およびワイブル分布である母集団から抽出されたそれぞれの標本に対し，*L*-モーメント比と分布関数の形状母数の推計値に対する偏り率の算定図表が作成され，さらにそれらの標準偏差を推定するための経験式が求められた．
7) 3種類の極値分布関数に対し，それぞれを母集団として設定したときの確率波高の推計値の偏り率の図表が作成された．

8) 母集団として3種類の極値分布関数の中から最適等分布を選定する場合の当てはめ的中率の図表が作成され，標本の大きさが 100 程度未満では当てはめ的中率が高くないことが確認された

9) 母集団が未知の場合の確率波高の変動係数に関する経験式が導かれた．この経験式は，裾長度，標本の大きさ，基準化変量の関数であり，数値実験によるデータの変動係数をかなりの確度で推定可能である．

10) 地域頻度解析を行うための地域分割に際して検討すべき不均質性指標に関して，*L*-モーメント比の変動係数の経験式を用いて推定される標準偏差に基づく新しい不均質性指標を提案し，日本海沿岸の高波資料についてその適用事例を説明した．

今回の数値実験では，3種類の分布関数の中から最適合のものを選ぶ指標として極値データとその推計値との間の相関係数を用い，相関係数が最大のものを最適合とする方式を用いた．母集団から抽出された標本に最も適合する分布が母集団の分布関数に的中する割合は，標本が大きくなるにつれて向上する．しかし，一般化パレート分布で形状母数の値が負の値の場合に，標本の大きさが $n = 100$ とかなり大きいときでも的中率が30%前後にとどまる場合があり，最適合の判断基準をさらに工夫する必要があろう．

また，今回新しく提案した地域内の複数地点の不均質性指標については今後さらに多くの事例について検討を進め，その適用性を向上させたい．

自然現象の極値については未解明の問題が数多くあり，今後も多様な研究を進めていくことが肝要である．

なお，この数値実験を実行するに当たっては，プログラムのバグ解消のために（株）エコー参事の久高将信氏に幾度となくご助力頂いた．ここに記して謝意を表する次第である．

参考文献

北野利一・間瀬　肇・喜岡　渉・矢野陽一郎 (2002)：一般化パレート分布による極値波浪解析－拡張形状母数の推定－，海岸工学論文集，第49巻，pp. 161-165.

合田良実 (1988)：極値統計におけるプロッティング公式ならびに推計値の信頼区間に関する数値的検討，港湾技術研究所報告，第27巻 第1号，pp. 31-92.

合田良実 (2002)：設計波高に係わる極値統計分布の裾長度パラメータとその意義，海岸工学論文集，第49巻，pp. 171-175.

合田良実 (2008)：「耐波工学 – 港湾・海岸構造物の耐波設計–」，鹿島出版会，第13章，pp. 347-354.

合田良実・小野澤昌己 (1990)：極値 II 型分布の特性と推計値の信頼区間，土木学会論文集，No.417/II-13(ノート), pp. 289-292.

合田良実・小長谷　修・永井紀彦 (1998)：極値波浪統計の母分布関数に関する実証的研究，海岸工学論文集，第45巻，pp. 211-215.

合田良実・竹下直樹・永井紀彦 (1999)：太平洋南岸の極値波高の母分布関数について，海洋開発論文集，Vol. 15，土木学会海洋開発委員会，pp. 327-331.

合田良実・久高将信 (2009a)：高波の極値統計解析に対する*L*-moments 法の適用と分布関数の選択について，ECOH/YG 技術論文 No. 10, 2009年5月7日，12p.

合田良実・久高将信・河合弘泰 (2009b)：*L*-moments 法を用いた波浪の極値統計解析について，海岸工学論文集，第56巻，pp.161-165.

合田良実・久高将信 (2010)：*L*-moment法による日本海沿岸の高波の地域頻度解析について，ECOH/YG 技術論文 No. 12, 2009年4月12日，13p.

外山奈央子・水野　量 (2002)：*L*-momentsを用いた地域頻度解析による全国アメダス地点における確率降水量の推定，気象庁研究時報，54巻 5－6号合併号，pp. 55-100.

Blom, G. (1965): *Statistical Estimates and Transformed Beta–Variables*, John Wiley & Sons, New York, Chapter 12.

Boes, D. C., J.-H. Heo, and J. D. Salas (1989): Regional flood quantile estimation for a Weibull model, *Water Resources Res.*, Vol. 25, No. 5, pp. 979-990.

Coles, S. (2001): *An Introduction to Statistical Modeling of Extreme Values*, Springer, 208p.

Chowdhury, J. U., J. R. Stedinger, and L.-H. Lu (1991): Goodness-of-fit tests for regional generalized extreme value flood distributions, *Water Resources Res.*, Vol. 27, No. 7, pp. 1765-1776.

Greenwood, J A. , J. M. Landwehr, N. C. Matalas, and J. R. Wallis (1978): Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form, *Water Resources Res.*, Vol. 15, No. 5, pp. 1049-1054.

Gringorten, I. I. (1963): A plotting rule for extreme probability paper, *J. Geophys. Res.*, Vol. 68, No. 3, pp. 813-814.

Hosking, J. R. M. (1990): *L*-moments: Analysis and estimation of distributions using linear combinations of order statistics, *J.Roy. Statistical Soc*., *Series B*, **52**, pp. 105-24.

Hosking, J. R. M. and J. R. Wallis (1987): Parameter and quantile estimation for the generalized pareto distribution, *Technometrics*, Vol. 29, No. 3, pp. 339-349.

Hosking, J. R. M. and J. R. Wallis (1995): A comparison of unbiased and plotting-position estimators of *L* moments, *Water Resources Res.*, Vol. 31, No. 8, pp. 1049-1064.

Hosking, J. R. M. and J. R. Wallis (1997): "*Regional Frequency Analysis*," Cambridge Univ. Press, 224p.

Hosking, J. R. M., J. R. Wallis, and E. F. Wood (1985): Estimation of the generalized extreme-value distribution by the method of probability-weighted moments, *Technometrics*, Vol. 27, No. 3, pp. 251-261.

Landwehr, J. M., N. C. Matalas, and J. R. Wallis (1979a): Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles, *Water Resources Res.*, Vol. 15, No. 5, pp. 1055-1064.

Landwehr, J. M., N. C. Matalas, and J. R. Wallis (1979b): Estimation of parameters and quantiles of Wakeby distributions, *Water Resources Res.*, Vol. 15, No. 5, pp. 1361-1379. (Correction, *Water Resources Res*. Vol. 15, p. 1672.)

Ma, Q.-S., Y.-B. Li, and J. Li (2006): Regional frequency analysis of significant wave heights based on *L*-moments, *China Ocean Engineering*, Vol. 20, No. 1, pp. 85-98.

Sinclair, C. D. and M. I. Ahmad (1988): Location-invariant plotting positions for PWM estimation of the parameters of the GEV distribution, *J. Hydrology*, Vol. 99, pp. 271-280.

van Gelder, P. H. A. J. M., J. De Ronde, N. W. Neykov, and P. Neytchev (2000): Regional frequency analysis of extreme wave heights: trading space for time, *Coastal Engineering 2000* (*Proc. 26th ICCE*, Sydney), ASCE, pp. 1099-1112.