

極値波高分布を特徴づける 2 つの指標 —— 裾長さと裾厚度 ——

北野 利一[†]

(受付 2003 年 7 月 30 日; 改訂 2004 年 4 月 21 日)

要 旨

海岸工学で最近に提案された裾長さについて極値理論の視点からその意義を考察し、さらに追加すべき指標として裾厚度を提案する。これらの指標の解釈および提案には、一般化極値分布の尺度・形状母数が Gumbel 確率紙上に描かれる quantile 関数の原点における勾配・曲率であるという特性を利用している。極値波高の母分布の候補として漸近分布の他に 3 母数 Weibull 分布を経験的に用いる長期波浪統計解析において、確率波高に加え、裾長さや裾厚度の 2 つの指標も併せて算定すれば、分布特性の把握や比較が可能となる。ただし、この指標により、得られた分布が最適であるか否かを判定できるものではない。また、極値分布とは異なる分布を極値分布と混在させて検討を行うような他の工学分野でも、これらの指標は適用可能と考える。

キーワード: Weibull 分布, 分布族の比較, quantile 関数, 年最大値資料, 極大値資料。

1. はじめに

沿岸域の構造物を設計するにあたり、来襲波を確率的に扱って沖設計波を算定する。その際、基礎資料として、短期間の地形変化などの影響が無視できる港外の沖に設けられた観測点で 1 時間毎に計測された有義波高の時系列を用いる。その基礎資料をもとに、POT 法 (Peaks-Over-Threshold method) に従い、台風や低気圧による気象擾乱毎の極大有義波高を抽出する。このようにして得られる極大有義波高やその年最大値を標本として、長期波浪統計解析を行う。ここで、有義波高とは、本質的に不規則な海面の高低に対し、ゼロクロス法などによる個別波分解を行って抽出される上位 1/3 の割合に含まれる波高の平均値である (例えば、Goda (2000) を参照)。したがって、有義波高は統計量であり、統計的変動性を伴う。なお、観測誤差を拾いやすい小さな波高も対象とした平均波高や、大きな統計的変動性を伴う観測時間内の最高波高よりも、有義波高は統計的に安定した代表量であることが示されている (合田 (1987), 北野 他 (2003))。

長期波浪統計解析においては、従来、Fisher and Tippett (1928) に基づく極値分布と 3 母数の Weibull 分布が母分布の候補として検討されてきた。しかし、台風・低気圧の来襲する年間頻度が十分に高くないことから、有義波高の時系列の極大値はもとより、有義波高の年最大値であっても、母分布を極値分布とみなすことに疑問がある。そこで、経験的かつ伝統的に海岸工学分野では Weibull 分布が母分布の候補として用いられてきた。

[†] 名古屋工業大学大学院 工学研究科: 〒466-8555 名古屋市昭和区御器所町

Gumbel 分布を極値 I 型とする 3 種類の極値分布は、一般化極値分布として 1 つの分布関数で表現できる。また、定義域の上限が有限である極値 III 型の分布関数は Weibull 分布と極値統計分野では一般に呼ばれている。しかし、上述の長期波浪統計解析に用いる Weibull 分布は、定義域の上限が無限大である通常の Weibull 分布であり、最大値に対する漸近分布を表す極値 III 型分布ではない (Weibull 分布という名称による混乱を避けるために、極値 III 型に相当する分布関数は逆 Weibull 分布と陽に区別した表現で呼ばれることもある)。従って、海岸工学で取り扱う年最大波高に対して、一般化極値分布と Weibull 分布の 2 種類の分布族が用いられていることになる。

極大波高を扱う場合には、従来から、極値分布と Weibull 分布のそれぞれの上裾部が用いられている。年最大波高資料と極大波高資料に対し、同じ 2 種類の分布族を用いるが、その意味は異なることに注意する。特に、任意の閾値で切断した極値分布の上裾部は、一般化パレート分布ではなく、上裾部の漸近分布を表すものではない。しかし、海岸工学での長期波浪統計解析法は、合田 (1988) をもとに数理統計学の知見が十分に応用されないまま独自に発展し今日に至っている。この経緯を思えば、極値分布の上裾部を用いることは、一般化パレート分布に相当するものを意図したと考えられる。なお、海岸工学では、一般化極値分布に対して形状母数が負となるもの (極値 III 型分布に相当) は母分布の候補としない。極値 III 型分布は定義域に上限があるため、過って適合した際に推定される確率波高が過小評価となる理由からか、極値波高の母分布の候補として取り上げられていない。

このように、波浪の極値解析では複数の分布族を用いるため、母分布がいずれの分布族に属するのか、という適合性の検討がこれまで多く議論されてきた。例えば、現在、海岸工学で標準的に用いられている Goda (2000) による解析法では、母数推定に最小自乗法を用いるため、形状母数の違いも含め、母分布関数の最適適合性について相関係数を用いた検討がされている。このような適合性の検討手法の良否はさておき、近隣の複数地点の波浪データを用いて、母分布の特性の違いについての相互比較を検討する際、推定された母分布が、一方は極値分布で、他方が Weibull 分布である場合には、位置母数や尺度母数を単純に比較することはできない。また、1 つの標本を対象に、従来の長期波浪統計解析法で母分布が Weibull 分布と推定される時に、同一の標本を一般化極値分布のみを用いて解析した場合に推定される結果と対比を行う時も同じ問題が生じる。特に、後者については、長期波浪統計解析を行う際の問題として、これまで議論されてきた (例えば、合田 他 (1989) の質議・回答における議論を見よ)。なお、理論的には、Weibull 分布における極値は、漸近的に Gumbel 分布に従う。そのため、現実的な問題として、母分布が極値分布とは異なる Weibull 分布と推定される場合にも、形状母数ゼロ付近となる一般化極値分布と区別がつかない場合が多い。この場合、一般化極値分布とどの程度に類似するのかを数量的に示すことが重要であると考えた。

本研究では、極値分布と他の分布を混在させて解析を行った際に、複数の標本についての母分布の相互比較や、同一の標本に対する解析法の違いによる母分布の推定結果の相互比較を可能とする汎用指標を考察することが目的である。

2. 確率波高

年最大値の分布関数を F とする時、確率波高 x_R と再現期間 R は次式のような関係を満たす。

$$(2.1) \quad R = 1 / \{1 - F(x_R)\}$$

すなわち、確率波高を超える波高の発生頻度が 1 回となる期間の平均が再現期間である。また、

極大値の分布関数の G に対しては、

$$(2.2) \quad \lambda R = 1/\{1 - G(x_R)\}$$

である。ここで、 λ は極大波高の単位時間あたりの平均出現数であり、上式の左辺は再現期間内の高波の総出現数の期待値となっている。式(2.1)および(2.2)は、分布関数の逆関数、すなわち、quantile 関数 F^{-1} および G^{-1} を用いれば、それぞれ、

$$(2.3) \quad x_R = F^{-1}(1 - 1/R)$$

$$(2.4) \quad x_R = G^{-1}(1 - 1/(\lambda R))$$

と書ける。長期波浪統計解析における第一の目的は、50 年あるいは 100 年といった与えられた長期の再現期間に対して、式(2.3)あるいは(2.4)を用いてその確率波高を推定することである。式(2.3)の F^{-1} あるいは式(2.4)の G^{-1} に任意の母分布の quantile 関数を適用しても、それぞれの確率波高の意味はもちろん変わらない。しかしながら、式(2.3)と(2.4)の確率波高は、厳密な定義として、互いに異なる量を表している。このことは、最大波高資料を対象とした確率波高であるのか、極大波高資料を対象としたものかを区別を要することを示唆する。ただし、両者の確率波高は、高波の出現過程が単純ポアソン過程である場合に漸近的に一致する。すなわち、分布関数 F および G が、

$$(2.5) \quad F(x_R) = e^{-\lambda\{1-G(x_R)\}}$$

を満たし、 x_R が十分大きい時、式(2.3)と(2.4)の確率波高の意味は一致する。以上から、確率波高 x_R は、同一な資料を対象とする場合には、異なる分布族に対しても相互に比較可能な汎用指標の 1 つであるといえる。また、最大波高資料と極大波高資料の異なる資料を対象とする場合にも、限定的ではあるが同じことがいえる。

ところで、20 年程度の観測に基づく標本に対し、母分布が異なっても、観測期間より短い期間(例えば、10 年)を再現期間とする確率波高の違いは大きくない。しかし、母分布の分布族が異なれば、観測期間より長い期間となる 50 年あるいは 100 年を再現期間とする確率波高の差異は大きい。そもそも極値統計解析は、母分布の形状の違いにより、このような差異が生まれることを主題に発展してきたものであると言える。したがって、母分布関数の違いを無視して、ある特定の再現期間に固定して得られる確率波高のみに着目し、分布関数や標本の比較を行うことは意味がない。

3. 裾長さ

3.1 導入

合田(2002)は、ある再現期間に対する確率波高だけでなく、分布の裾の広がりも考慮に入れて母分布の特性を比較できるように、裾長度を導入した。すなわち、10 年確率波高 x_{10} に対する 50 年確率波高 x_{50} の比

$$(3.1) \quad \gamma_{50} = \frac{x_{50}}{x_{10}}$$

として、裾長さ γ_{50} を定義した。確率波高 x_{10} および x_{50} は、任意の母分布関数に対して算出できるので、 γ_{50} は汎用指標となりうる。裾長さと確率波高の 1 つを固定すれば、他の確率波高は一意に定まる。図 1 は、これら 2 つの確率波高 x_{10} および x_{50} を固定して、次式の Weibull 分布の累積確率 F と波高 x の関係を示したものである。

$$(3.2) \quad F(x) = 1 - \exp\left\{-\left(\frac{x-b}{a}\right)^k\right\}$$

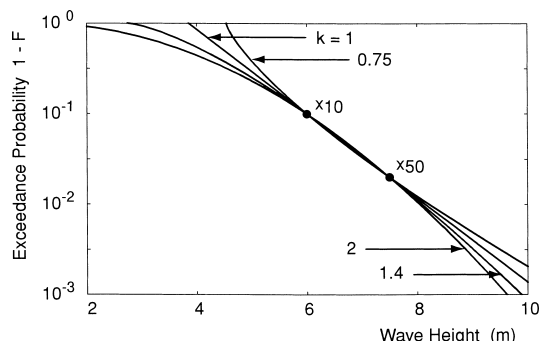


図1. Weibull 分布に対する波高の超過確率(対数確率目盛り).

図中の4種類の形状母数の値 ($k = 0.75, 1, 1.4, 2$) は, Goda (2000) の長期波浪統計解析法で用いることが推奨されている値であり, 10年および50年確率波高は, 例示のために, それぞれ6.0m および7.5m としている.

防波堤設計に係るモンテカルロ・シミュレーション(詳細については, 合田(2001)を参照)によって, 50年間の供用期間中の防波堤に対し, 被災の主因となる総滑動量に影響を与える波高は, 10年確率波高から50年確率波高にかけての波高であることが見出された. 図1を見るとおり, 4種類のWeibull分布について, 2つの確率波高 x_{10} および x_{50} を固定すれば, それら2つの波高の区間では, 超過確率関数 $1-F$ は互いに類似した曲線となっていることがわかる. ただし, 図1は縦軸の超過確率を対数目盛りでとっており, 確率波高 x_{10} および x_{50} の近傍を意図的に拡大していることに注意しなければならない. 合田(2002)では, 裾長さの意義について, 極値理論に基づく考察が行われていないので, 以下では, 極値理論の視点からその意義を考察する.

確率分布の相違をより適切に理解するために, 確率を等間隔目盛りで表したものを図2に示す. 確率波高 x_{10} と x_{50} には含まれる区間で, 異なる形状母数をもつWeibull分布の差異はほとんど見られないが, 図2(a)を見るとおり, そのような区間に含まれる波高の出現確率はわずかである. むしろ, 図2(a)は, 各分布に相違があることを強調していることになる. そこで, 分布確率を累乗して, 10年最大波高および50年最大波高の分布の累積確率と確率波高の関係を図示したものが, 図2(b)および(c)である. 10年最大波高の累積確率と確率波高の関係を示す図2(b)が, 確率としてとりうる主要な範囲(累積確率が0.3~0.8となる範囲)で各種分布関数の差異がわずかであることを示唆し, 図1で意図したことを適切に表現している. また, 図2(c)より, 50年最大波高の累積確率でみれば, 各分布に相違があることが確認できる.

3.2 Gumbel 確率紙

Gumbel 確率紙を用いれば, 図2に示したような3種類の異なる期間に対する最大値分布を用いた議論が容易にできることを示そう.

Gumbel 確率紙上の分布関数 F の quantile q_P の曲線は,

$$(3.3) \quad y_P = -\log(-\log P)$$

と表される標準 Gumbel 分布の quantile y_P を横軸にとり,

$$(3.4) \quad q_P = F^{-1}(P)$$

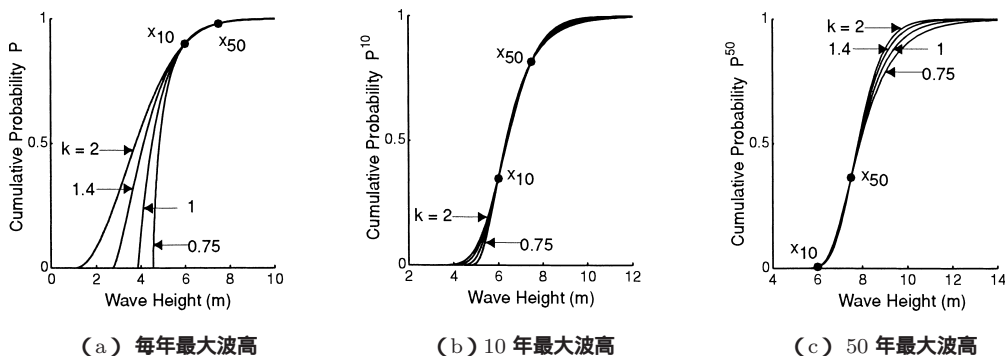


図 2. Weibull 分布に対する波高の超過確率(等確率間隔目盛り).

を縦軸に描いたものである．したがって，式(3.4)における分布関数 F を次式に表される位置母数 μ および尺度母数 σ とする Gumbel 分布

$$(3.5) \quad F(x) = \exp \left\{ - \exp \left(- \frac{x - \mu}{\sigma} \right) \right\}$$

とすれば， q_P と y_P には，

$$(3.6) \quad q_P = \mu + \sigma y_P$$

という直線関係がある．また，式(3.4)における分布関数 F を位置および尺度母数を μ および σ とする一般化極値(GEV)分布

$$(3.7) \quad F(x) = \exp \left\{ - \left(1 + \xi \frac{x - \mu}{\sigma} \right)^{-1/\xi} \right\}$$

とした時には， q_P と y_P の関係は，次式ようになる．

$$(3.8) \quad q_P = \mu + \sigma \frac{\exp(\xi y_P) - 1}{\xi}$$

Gumbel 確率紙上における GEV 分布の quantile 関数は，横軸 y_P の原点を移動すれば，移動量に対応した期間 L における最大値分布の quantile 関数となる．このとき，その位置および尺度母数がもとの分布の特性量とどのような関係があるかを示そう．以下では具体的に，横軸 y_P の原点を次式のように移動させ，10 年確率波高に対する横軸の値がゼロとなるように，改めて軸 y'_P を設ける．

$$(3.9) \quad y'_P = y_P - y_0; \quad y_0 = y_P \Big|_{P=1-1/10}$$

この時，GEV 分布の quantile q_P は次式で表され，式(3.8)と同じ表現となっていることがわかる．

$$(3.10) \quad q_P = x_{10} + \sigma_{10} \frac{\exp(\xi y'_P) - 1}{\xi}$$

上式における x_{10} および σ_{10} は，

$$(3.11) \quad x_{10} = \mu + \sigma \frac{\exp(\xi y_0) - 1}{\xi}; \quad \sigma_{10} = \sigma \exp(\xi y_0)$$

である．ここで注意すべきことは， x_{10} が再現期間 10 年とする確率波高となることを意図して，原点を移動している点である．また，年最大値の分布が GEV 分布に従い，その位置および尺度母数を μ および σ とすれば， L 年最大値分布は同一の形状母数の値をもつ GEV 分布に一致し，その位置母数 μ_L および尺度母数 σ_L は

$$(3.12) \quad \mu_L = \mu + \sigma \frac{L^\xi - 1}{\xi}, \quad \sigma_L = \sigma L^\xi$$

となる．式(3.12)における L を

$$(3.13) \quad L = \exp(y_0) \approx 9.49$$

と置き換えることにより，約 9.49 年最大波高の分布関数の位置母数は 10 年確率波高に一致し ($\mu_L = x_{10}$)，尺度母数も対応するものに一致する ($\sigma_L = \sigma_{10}$)．また，約 9.49 年最大波高の quantile 関数は，Gumbel 確率紙において，毎年最大波高の quantile 関数の横軸の原点を 10 年確率波高の横軸の値 y_{10} に移動したものと同一となるのがわかる．

以上のことを再現期間を 10 年間とするかわりに任意の再現期間で議論することは容易である．したがって，GEV 分布の quantile 関数を Gumbel 確率紙上に描き，その quantile 曲線上にある特定の点に着目した時，その点を位置母数とする期間最大値分布となる．図 3 に示す GEV の quantile 関数は，それぞれ形状母数は異なる値をとるが，10 年確率波高 x_{10} と 50 年確率波高 x_{50} が一致し，それら 2 つの確率波高にはさまれる領域では，quantile 関数に大きな差異は見られない．これは，約 9.49 年最大波高の分布関数が類似することを意味する．また，10 年確率波高 x_{10} 未満の領域や 50 年確率波高 x_{50} の領域では，quantile 関数の違いが見られるため，毎年最大値分布や 50 年最大波高の分布関数は類似したものではないといえる．ある 1 つの quantile 曲線に対して，横軸の原点を移動した時，その曲線がある期間内の最大値分布の quantile 曲線となることを厳密にいうためには GEV 分布であることが必要である．しかし，複数の quantile 曲線が類似したものかどうかを判断する程度であれば，GEV 分布であることは必要ない．むしろ，GEV 分布ではない分布と GEV 分布との比較や，GEV 分布ではない分布どうしの比較を行う際には Gumbel 確率紙にそれらの quantile 関数を描いて類似や相違を

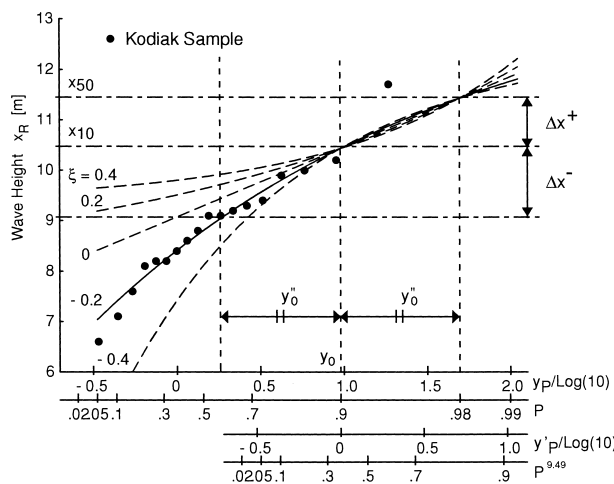


図 3. Gumbel 確率紙における GEV 分布の Quantile 関数．

検討すればよい。GEV 分布ではない Weibull 分布の互いの類似性について図 2 のような 3 種類の図を用いた議論は、Gumbel 確率紙上に描けば 1 種類の図のみで議論できるわけである。

Gumbel 確率紙には、もう 1 つの重要な特性がある。式(3.6)および(3.8)において、

$$(3.14) \quad q_P = \mu \Leftrightarrow y_P = 1 \Leftrightarrow P = e^{-1} \approx 0.368$$

となるので、累積確率が e^{-1} となる quantile q_P は、位置母数 μ の値に一致する。さらに、点 $q_P = \mu$ における quantile q_P の勾配 dq_P/dy_P は

$$(3.15) \quad \left. \frac{dq_P}{dy_P} \right|_{q_P=\mu} = \sigma$$

であり、尺度母数 σ の値に一致する。さらに、式(3.15)の関係を軸 y'_P で見れば、式(3.10)より、

$$(3.16) \quad \left. \frac{dq_P}{dy'_P} \right|_{q_P=x_{10}} = \sigma_{10}$$

となる。上式の左辺は、次式で示す Δy および Δx という離散量による比 $\Delta x/\Delta y$ で表すことができる。

$$(3.17) \quad \begin{cases} \Delta y = \frac{\exp(\xi y''_P) - 1}{\xi}; & y''_P = y_{P=1-1/50} - y_{P=1-1/10} \\ \Delta x = x_{50} - x_{10} \end{cases}$$

式(3.17)を式(3.16)に代入し、変形することにより、式(3.1)で定義される裾長さ γ_{50} は

$$(3.18) \quad \gamma_{50} = \frac{\sigma_{10}}{x_{10}} \Delta y + 1$$

と表されることがわかる。したがって、合田(2002)により提案された裾長さ γ_{50} は、約 9.49 年最大波高の確率分布を GEV 分布とみなした時の位置母数 x_{10} と尺度母数 σ_{10} を用いた無次元量で表されることがわかる。このことは、毎年最大波高の母分布が厳密に GEV 分布とみなせない場合であっても、10 年確率波高の近傍の quantile 関数の勾配を、GEV 分布の尺度母数とみなすことにもとづく。以上の考察により、裾長さ γ_{50} は、任意の母分布関数に対して同義の意味をもち、それゆえ、異なる分布族の分布関数に対しても、相互に比較可能な汎用指標となっている。

3.3 一般化パレート分布を用いた裾長さの推定

毎年最大値に対する母分布については、Gumbel 確率紙における GEV の quantile 関数の性質を用いて考察し、10 年確率波高 x_{10} とそれが位置母数となる約 9.49 年最大値分布の尺度母数 σ_{10} により裾長さ γ_{50} が式(3.18)で表されることを示した。これに対し、有義波高の時系列を POT 法により抽出した極大波高の母分布については、指数確率紙における一般化パレート分布(GPD)の quantile 関数の性質を利用し、10 年確率波高 x_{10} とそれを閾値とする極大波高分布の特性量により裾長さ γ_{50} が表されることを示そう。

閾値を u 、尺度母数を σ とする一般化パレート分布(GPD)の累積分布関数 G は、次式のよう表される。

$$(3.19) \quad G(x) = 1 - \left(1 + \xi \frac{x - u}{\sigma} \right)^{-1/\xi}$$

閾値 x_1 を越える極大波高が 1 年に平均的に 1 回の割合で出現するならば、閾値 x_1 は 1 年確率波高とみなせる。なお、期間最大波高と極大波高のそれぞれの確率波高の意味には相違点があ

ることは、2 節で述べたとおりである。ここで、閾値 x_1 を越える極大波高の分布を GPD (尺度母数 σ_1 および形状母数 ξ) とすれば、その quantile q_P は、以下のように式 (3.8) と同じ表現式で表される。

$$(3.20) \quad q_P = x_1 + \sigma_1 \frac{\exp(\xi y_P) - 1}{\xi}$$

ただし、上式の y_P は次式に示す標準指数分布の quantile である。

$$(3.21) \quad y_P = -\log(1 - P)$$

式 (3.9) で表されるように変数 y_P を移動させて新たな変数 y'_P を用いれば、式 (3.20) に示す GPD の quantile q_P も式 (3.10) として書き改められる。この場合の定数 x_{10} および σ_{10} は、

$$(3.22) \quad x_{10} = x_1 + \sigma_1 \frac{\exp(\xi y_0) - 1}{\xi}; \quad \sigma_{10} = \sigma_1 \exp(\xi y_0)$$

と表される。これは、式 (3.11) における GEV の位置母数 μ および尺度母数 σ を GPD の閾値 x_1 および尺度母数 σ_1 に書き改めたものである。さらに、式 (3.10) として表される極大波高の quantile q_P に対し、式 (3.16) および (3.17) と同じ操作を行えば、極大波高に対する裾長さ γ_{50} も式 (3.18) で表される。ところで、

$$(3.23) \quad R = \exp(y_0)|_{y_0 = -\log(1/10)} = 10$$

とおき、式 (3.22) の 1 つめの式の両辺から極大波高 $x (> x_{10})$ をそれぞれ引いて整理すれば、

$$(3.24) \quad 1 - H\left(\frac{x - x_{10}}{\sigma_{10}}\right) = R \left\{ 1 - H\left(\frac{x - x_1}{\sigma_1}\right) \right\}; \quad H(y) = 1 - (1 + \xi y)^{-1/\xi}$$

という関係が得られる。式 (3.24) の右辺は、1 年確率波高 x_1 を越える波高分布の上裾部 (超過確率が 1/10 から 1 までの範囲) を抽出した超過確率分布を意味する。そのため、式 (3.24) の左辺は、10 年間に平均的に 1 回の割合で出現する 10 年確率波高 x_{10} を閾値とし、その閾値を越える極大波高の超過確率である。したがって、極大波高に対する裾長さ γ_{50} は、10 年確率波高 x_{10} とその値を閾値とする GEV の尺度母数 σ_{10} を用いて表される無次元量である。なお、2 節で述べたとおり、期間最大波高と極大波高とは扱いが異なるので、裾長さ γ_{50} が同じ式で表されていても、その厳密な意味は異なる。

式 (3.18) に含まれる x_{10}/σ_{10} は、GPD における閾値と尺度母数の比 u/σ と次のような関係がある。

$$(3.25) \quad 1 - \xi \frac{x_{10}}{\sigma_{10}} = \left(\frac{\tau_u}{10}\right)^\xi \left(1 - \xi \frac{u}{\sigma}\right)$$

ここで、 τ_u は極大波高が閾値 u を越える平均周期であり、単位時間あたりの出現数 λ_u の逆数である。この式は、標本からの直接的な推定量 $\hat{\sigma}, \hat{\xi}$ および $\hat{\tau}_u$ から裾長さの推定量 $\hat{\gamma}_{50}$ を算定する際に必要となる。ところで、式 (3.25) の導出にあたっては、次のように考える。式 (3.24) の成立と同じ理由から、

$$(3.26) \quad 1 - H\left(\frac{x - x_1}{\sigma_1}\right) = \tau_u^{-1} \left\{ 1 - H\left(\frac{x - u}{\sigma}\right) \right\}, \quad x > x_1 > u$$

が成立する。そこで、式 (3.24) および (3.26) における変数 x にゼロを代入し、 $(1 - x_1/\sigma_1)$ を消去して整理すればよい。閾値 u が正であるなら変数 x にゼロを代入できないが、その理由は

H を分布関数とみているからであり、ここでの計算にはその理由はあてはまらない)。さらに、式(3.22)における尺度母数の関係から、

$$(3.27) \quad \frac{\sigma_{10}}{\sigma_1} = \exp(\xi y_0) = 10^\xi; \quad \frac{\sigma_1}{\sigma} = \tau_u^{-\xi}$$

となるので、式(3.25)は次式に示す修正尺度母数 σ_* の定義式と同義であることがわかる。

$$(3.28) \quad \sigma_* = \sigma - \xi u = \sigma_{10} - \xi x_{10}$$

すなわち、閾値が変化した時、尺度母数 σ の値は変化するのに対し、修正尺度母数 σ_* の値は変化しないことを示すものである。このような特性量が定義できることは、GPD の重要な性質の 1 つである。

以上の結果を用いれば、極大波高資料に対しても、GPD による裾長さ γ_{50} の推定が可能となる。

4. 裾厚度

図 3 は、2 つの確率波高 x_{10} および x_{50} を固定して、形状母数 ξ の値が異なる GEV 分布の quantile 関数を描いたものである。位置および尺度母数の汎用指標として、確率波高 x_{10} および裾長さ γ_{50} を固定しているので、確率波高 x_{10} と x_{50} には含まれる波高については、分布関数の違いによる差異は小さい。なお、図中の点は、後で具体例として示す Kodiak 沖の波浪資料の年最大値を GEV 分布を用いて母数推定を行い、プロットしたものである。GEV 分布の形状母数 ξ の推定値(95% 信頼率)は、

$$(4.1) \quad \hat{\xi} = -0.202 \pm 0.272$$

であり、quantile 関数の推定曲線は図 3 における $\xi = -0.2$ のものにほぼ相当する。確率波高 x_{10} より小さい波高に対して、 $\xi \neq 0.2$ の quantile 関数はいずれも波高の標本値と異なる傾向を示しており、互いの違いが明確に現れている。このことは、図 3(a)における異なる分布関数による差異が顕著に示されることに呼応する。

確率波高 x_{10} の右側で決まる確率波高 x_{50} と x_{10} の差 Δx^+ に対する勾配 $\Delta x^+/\Delta y$ を用いて、裾長さ γ_{50} は定義されている。これに対し、確率波高 x_{10} の左側で決まる勾配 $\Delta x^-/\Delta y$ は、右側の勾配 $\Delta x^+/\Delta y$ に一致しない($\xi = 0$ の例外を除く)。このために、標本から推定される値とは異なる形状母数の quantile 関数は、標本値から乖離する。このような分布関数による差異を考慮に入れるために、形状母数と密接に関連する、もう 1 つの指標が必要である。

式(3.8)に対する q_P の 2 階の微分係数を、その 1 階の微分係数で無次元化すれば、

$$(4.2) \quad \left(\frac{d^2 q_P}{dy_P^2} / \frac{dq_P}{dy_P} \right) \Big|_{q_P=\mu} = \xi$$

となる。したがって、GEV 分布の形状母数 ξ は、位置母数となる波高における quantile 関数 ($q_P = \mu$) の曲率であるといえる。さらに、GEV 分布では式(4.2)において $q_P = \mu$ と限定する必要はなく、全ての波高における quantile 関数の曲率となることは明らかである。また、式(4.2)における微係数の比は、

$$(4.3) \quad \frac{d^2 q_P}{dy_P^2} / \frac{dq_P}{dy_P} = \frac{d}{dy_P} \log \frac{dq_P}{dy_P}$$

と変形できる。さらに、式(3.9)により横軸の原点を移動し、

$$(4.4) \quad d \log(dq_P) = \log \frac{\Delta x^+}{\Delta x^-}, \quad dy_P = y_P'' = y_P|_{P=1-1/50} - y_P|_{P=1-1/10}$$

と離散化する．

結果として，確率波高 x_{10} での近傍での quantile 関数の曲率を意味するように，裾厚度 ξ_{10} を次式で定義することができる．

$$(4.5) \quad \xi_{10} = \log_b \frac{\Delta x^+}{\Delta x^-} = \log \frac{\Delta x^+}{\Delta x^-} / \log b$$

ただし，比 $\Delta x^+/\Delta x^-$ および対数の底 b は，

$$(4.6) \quad \frac{\Delta x^+}{\Delta x^-} = \frac{x_{50} - x_{10}}{x_{10} - x_{2.37}}; \quad b = \exp(y_P'') \approx 5.22$$

である．なお，母分布が GEV 分布である場合には，裾厚度 ξ_{10} は GEV 分布の形状母数に一致する．また，式(4.6)の比 $\Delta x^+/\Delta x^-$ は位置および尺度母数に依存しないことから，裾厚度 ξ_{10} は分布関数の形状母数のみにより決まることは明白である．

極大波高資料の母分布に対しても，3.3 節での考察に基づき，式(2.4)から得られる確率波高と式(3.19)で表される標準指数分布の quantile 関数を用いて， y_P'' が等しくなるように $\Delta x^+ (= x_{50} - x_{10})$ に対応する Δx^- を定めることにより，式(4.6)の代わりに

$$(4.7) \quad \frac{\Delta x^+}{\Delta x^-} = \frac{x_{50} - x_{10}}{x_{10} - x_2}; \quad b = 5$$

を用いればよいことがわかる．その結果，母分布が GPD となる場合には，裾厚度 ξ_{10} は GPD の形状母数 ξ に一致する．また，非毎年資料の母分布に対する ξ_{10} も，毎年資料の母分布のものと同様に，位置および尺度母数に依存せず，形状母数により決まる．しかし，裾厚度 ξ_{10} は 10 年確率波高という特定の再現期間の再現値の近傍で定義される量であるので，非毎年資料の母分布の場合，平均発生周期(あるいは，年間発生率)に依存する可能性がある．ただし，母分布が GPD となる場合に限り，裾厚度 ξ_{10} はその形状母数 ξ に一致するので， ξ_{10} は平均発生周期に依存しない．

以上の考察から，裾厚度 ξ_{10} は，任意の母分布関数に対して算出可能かつ相互に比較可能な汎用指標である．なお，本研究では，推定された母分布の特性比較のための指標として裾厚度 ξ_{10} を導入したが， ξ_{10} は極値分布(特に，形状母数の値によっては最尤推定に意味のない場合を含む極値 III 型分布)の形状母数の推定に用いられる統計量(Reiss and Thomas(2001), p. 121 を参照)

$$(4.8) \quad \hat{\xi} = 2 \log(\hat{r}) / \log \left(\frac{\log P_0}{\log P_1} \right); \quad \hat{r} = \frac{x_{[nP_2]:n} - x_{[nP_1]:n}}{x_{[nP_1]:n} - x_{[nP_0]:n}}$$

に相当するものである．ただし，式(4.8)における P_i ($i = 0, 1, 2$) は，

$$(4.9) \quad (\log P_1)^2 = (\log P_2)(\log P_0); \quad 0 < P_0 < P_1 < P_2 < 1$$

を満足する値で， $x_{j:n}$ は標本サイズ n における昇順に並べた j 番目の順序統計量である．

毎年最大値資料(AMS)および極大値資料(POT)の母分布関数を Weibull 分布とする場合の裾厚度 ξ_{10} の値を表 1 に示す．形状母数 k は Goda(2000)で提案されている 4 つの値に加え，

表 1. Weibull 分布に対する裾厚度 ξ_{10} の値．

k (Weibull)		0.75	1	1.4	2	(3)
AMS		0.21	0.07	-0.05	-0.15	(-0.22)
POT:	$\lambda = 2$	0.12	0.00	-0.10	-0.18	(-0.24)
	4	0.09	0.00	-0.08	-0.14	(-0.19)
	8	0.08	0.00	-0.07	-0.12	(-0.16)

$k = 3$ の場合も参考として付した。Weibull 分布の吸引域 (Domain of Attraction) は Gumbel 分布 ($\xi = 0$ の GEV) であることから、年間発生率 λ の増加とともに、POT の場合の ξ_{10} の値はゼロに漸近する傾向が認められる。また、 $k > 1$ の Weibull 分布に対しては ξ_{10} が負である。このことは、極値 III 型分布を母分布の候補としない従来の極値波浪解析法において、極値分布ではない Weibull 分布を母分布のもう 1 つの候補として含めることにより、 x_{10} の近傍での形状が極値 III 型分布に類似する分布関数も含めた検討となっていることがわかる。

5. 適用例

裾長さおよび裾厚度の適用例として、Goda (2000) が検討した Kodiak 沖における 1956 ~ 1975 年の 20 年間における極大波高資料 (6 m 以上の 78 波高) を用いる (資料の詳細は、van Vledder et al. (1993) を参照)。なお、1957 年には 6 m 以上の高波が発生しなかった (6 m 未満の波高は不明) ことから、図 3 における年最大値は、1958 ~ 1975 年の 18 年間の年最大波高を用いている。Goda (2000) は、Gumbel 分布と 4 種類の Fréchet 分布 (以上は、形状母数 $\xi = 0, 0.1, 0.2, 0.3, 0.4$ の GEV 分布に相当) と 4 種類の Weibull 分布の計 9 種類の分布関数およびそれらの上裾部を用いて、最小自乗法により母数を推定し、相関係数に基づく独自の方法で最適を求めている。

本論文では、GPD を母分布として用いて最尤推定を行う。図 4 ~ 7 は、閾値を変化させて得られる裾長さ γ_{50} 、裾厚度 ξ_{10} (GPD の場合には、形状母数 ξ に相当)、修正尺度母数 σ_* および次式に示される超過平均値 ε の各々の推定値である。

$$(5.1) \quad \varepsilon = E(x - u)_+; \quad s_+ = \max(0, s)$$

なお、図の作成に関しては、Coles (2001) で紹介されているプログラムを参考にして作成した。閾値による各々の推定値の変化を見るとおり、適切な閾値は $u = 8.0$ m と判断できる。その結

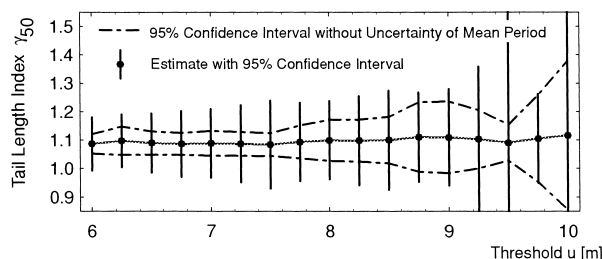


図 4. Kodiak 沖の高波分布に対する裾長さ γ_{50} の推定。

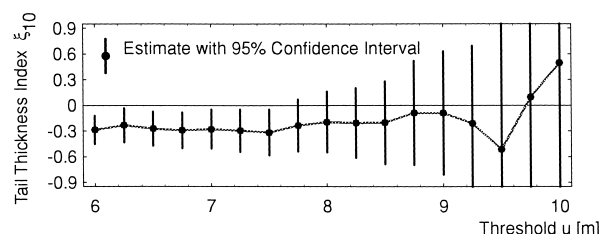


図 5. Kodiak 沖の高波分布に対する裾厚度 ξ_{10} の推定。

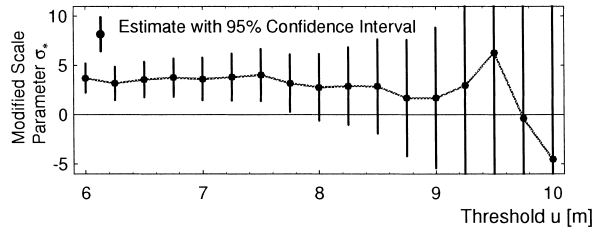


図 6. Kodiak 沖の高波資料に対する GPD の修正尺度母数 σ_* の推定 .

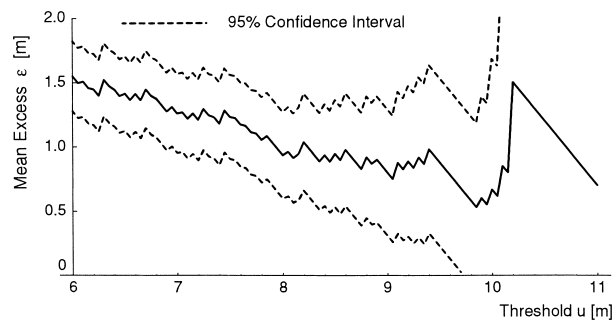


図 7. Kodiak 沖の高波資料に対する超過平均 ϵ の推定 .

果, 確率波高 x_{10} および x_{50} , 裾長さ γ_{50} および裾厚度 ξ_{50} の推定結果(95% 信頼率)は次のように得られた .

$$(5.2) \quad \begin{cases} \hat{x}_{10} = 10.38 \pm 0.81 \text{ m}; & \hat{x}_{50} = 11.41 \pm 1.30 \text{ m} \\ \hat{\gamma}_{50} = 1.100 \pm 0.138; & \hat{\xi}_{10} = -0.196 \pm 0.356 \end{cases}$$

なお, GPD を用いた推定では, 平均周期 τ_u の推定誤差が含まれる($\hat{\tau}_u = 0.833 \pm 0.170$ 年). 確率波高のような再現値の推定誤差に対しては, 平均周期 τ_u の推定誤差が無視できる程度に小さい(Coles(2001)). しかし, 図 4 で示されるように, 点線は平均周期の推定誤差を含めずに算出した裾長さ γ_{50} の信頼区間を示し, 総ての推定誤差に基づく信頼区間と比較すれば, 平均周期の推定誤差が無視できないことがわかる. また, 裾厚度 ξ_{10} については, GPD および GEV 分布の場合には, その形状母数に相当するが, 点推定値は互いにほぼ等しい. この例においては, GPD による $\hat{\xi}_{10}$ ($= \hat{\xi}$) の信頼区間が, GEV 分布によるもの(式(4.1)の値)に比較して小さいとはいえない.

Goda(2000)では, Kodiak 沖の極大波高資料に対して, 位置, 尺度および形状母数が

$$(5.3) \quad \hat{a} = 1.8621 \text{ m}; \quad \hat{b} = 5.805 \text{ m}; \quad \hat{k} = 1.4$$

の Weibull 分布であると推定しており, その場合の裾長さおよび裾厚度は,

$$(5.4) \quad \hat{\gamma}_{50} = 1.133; \quad \hat{\xi}_{10} = -0.081$$

となる. 式(5.4)での $\hat{\xi}_{10}$ は上述の式(5.2)で示される $\hat{\xi}_{10}$ よりゼロに近く, やや Gumbel 分布よりの母分布として推定している. Goda(2000)の結果によるモデルと標本の Q-Q 図と GPD

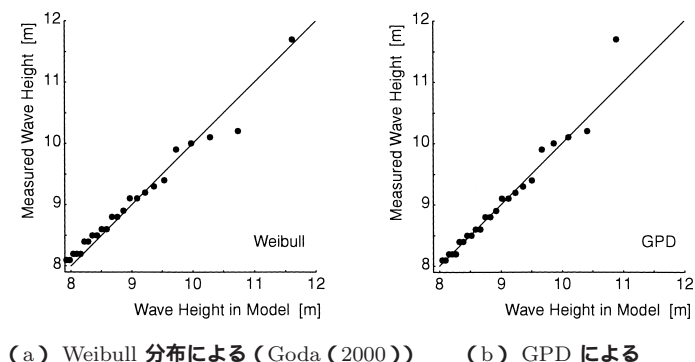


図 8. 推定母分布関数による極大波高の Q-Q 図 .

によるものを図 8 に示す . Goda (2000) は最小自乗法による推定であることから , Q-Q 図におけるモデル値は標準サイズ \tilde{N} における降順 m 番目の順序統計量に対し , プロットング公式

$$(5.5) \quad 1 - \hat{F}_m = \frac{m - (0.20 + 0.27/\sqrt{\hat{k}})}{\tilde{N} + (0.20 + 0.23/\sqrt{\hat{k}})}$$

で得られる基準化変量(合田(1988))を用いている . また , GPD では閾値 u を 8.0 m としているので , Goda (2000) による Q-Q 図は 8.0 m 以上を抜き出している . 図 8(a) より , 最大波高の実測値とモデル値がほぼ同程度であることが見てとれる . 一般に , 高波資料の最大値の変動は他に比べて大きいにも関わらず , Goda (2000) の結果は最小自乗法による推定であることから , 推定に与える高波資料の最大値の寄与が過大になっている疑いがある . このことが原因となって , Goda (2000) の結果による $\hat{\xi}_{10}$ の値が , GPD による値よりゼロ側にくる , すなわち Goda (2000) により推定される母分布が Gumbel 分布に近くなっていると考えられる .

6. むすび

本研究では , 一般化極値分布における尺度および形状母数に相当する裾長さおよび裾厚度を考察した . 極値分布ではない分布も母分布関数の候補に混在させて検討する波浪の長期統計解析を行う分野では , 再現値である確率波高だけでなく , 裾長さおよび裾厚度を算定すれば , 異なる分布族間での分布関数の比較が可能となることを示した . また , 高波の経験分布として従来から用いられている Weibull 分布では , その裾厚度の値が正負のいずれもとりうることを示した . このことにより , 極値 III 型分布を母分布の候補としない従来の極値波浪解析法において , 極値分布ではない Weibull 分布を母分布のもう 1 つの候補とすることから , 極値 III 型分布に相当する分布関数も含めた検討となっていることがわかった . 裾長さおよび裾厚度の具体的な資料に対する適用例も示した .

裾長さと裾厚度の呼称についてふれておきたい . 裾長さは , 合田 (2002) によりすでに名づけられており , 本論文でもその呼称を踏襲している . ただし , その英語名を , Goda (2001a) は spread parameter としている (なお , それに先立ち , Goda (2001b) は , skewness parameter と名づけたが , 一般的な統計用語としての skewness と混乱するために変更した経緯がある) が , 本論文では , tail length index と呼ぶことを提案したい . なぜなら , 裾長さとは GEV の尺度母数を汎用化したものであり , 毎年最大値分布の裾部に着目して得られる , 約 9.49 年最大値分

布の範囲の代表長さ(の無次元量)である。そのため、長さという概念がふさわしいと考える。したがって、裾の形状に直接依存する量ではないことを意図した。また、母数(parameter)ではなく、確率波高と同様、母数を用いて算定される量であるので、指標(index)としている。他方、裾厚度(tail thickness index)は GEV の形状母数を汎用化したものであり、裾部の quantile が占める比率(厚み)を見ていると解釈して名づけた。

謝 辞

本論文の推敲にあたり、京都大学防災研究所 間瀬肇 助教授にはアドバイスを頂いた。また、査読者の刺激的な意見は本論文を修正するにあたり非常に参考になった。この場を借りて感謝の意をあらわす。

参 考 文 献

- Coles, S. (2001) *An Introduction to Statistical Modeling of Extreme Values*, Springer, London (see also <http://www.maths.bris.ac.uk/~masgc/ismev/summary.html>)
- Fisher, R. A. and Tippett, L. H. C. (1928) Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Proceedings of the Cambridge Philosophical Society*, 24, 180–190.
- 合田良実(1987) 数値シミュレーションによる波浪の標準スペクトルと統計的性質, 第 34 回海岸工学講演会論文集, 131–135.
- 合田良実(1988) 極値統計におけるプロットング公式ならびに推定値の信頼区間に関する数値的検討, 港湾技術研究所報告, 27(1), 31–92.
- Goda, Y. (2000) *Random Seas and Design of Maritime Structures*, World Scientific, London.
- Goda, Y. (2001a) Performance-based design of caisson breakwaters with new approach to extreme wave statistics, *Coastal Engineering Journal*, 43(4), 286–316.
- Goda, Y. (2001b) Extreme wave statistics for reliability-based design of caisson breakwater, *Proceedings of the International Workshop on Advanced Design Methods of Maritime Structures in the 21st Century, Yokosuka, Japan*, Paper No. K-1.
- 合田良実(2001) 防波堤の信頼性設計における極値分布の選択とその影響, 海洋開発論文集, 17, 1–5.
- 合田良実(2002) 設計波高に係わる極値統計分布の裾長さパラメータとその意義, 海岸工学論文集, 49, 171–175.
- 合田良実, 宝 馨, 高棹琢馬(1989) 「水文頻度解析における確率分布モデルの評価基準」に対する討議・同回答, 土木学会論文集, No. 405/II-11, 265–272.
- 北野利一, 二宮太一, 喜岡 涉, 間瀬 肇(2003) 有義波高の統計的変動性——波別解析法に自由度を導入した理論解——, 海岸工学論文集, 50, 161–165.
- Reiss, R.-D. and Thomas, M. (2000) *Statistical Analysis of Extreme Values*, 2nd ed., Birkhäuser, Berlin.
- van Vledder, G., Goda, Y., Hawkes, P., Mansard, E., Martin, M. H., Mathiesen, M., Peltier, E. and Thompson, E. (1993) Case studies of extreme wave analysis: A comparative analysis, *Proceedings of 2nd International Symposium on Ocean Wave Measurement and Analysis, New Orleans*, 978–992, American Society of Civil Engineers.

Two Indices for Tail Part Distribution of Population of Extremal Wave Heights: Tail Length and Tail Thickness

Toshikazu Kitano

(Nagoya Institute of Technology)

Both Fisher-Tippett's extreme value distributions and the Weibull distribution are tested as candidate population distributions in extreme wave analysis in the field of coastal engineering. The local maxima or even the annual maxima of significant wave heights don't necessarily promised to follow a theoretical extreme value distribution. This paper introduces two dominant indices, tail length and tail thickness, associated with two or three return values for the purpose of comparing different distributions. These indices are intended to correspond to the scale and shape parameters of the generalized extreme value distribution, whose tail is regarded as that of the target population function at a specific return period. By using a return value together with the tail length and thickness indices, the difference between distributions can be easily examined quantitatively. A practical example of analysis using these indices is demonstrated by a hindcasting wave data set. These indices are also applicable in other fields of extreme analysis by employing not only the extreme value distribution but also the others as the candidates of population distribution.