

L-moments法を用いた波浪の極値統計解析について

Use of *L*-moments Method for Extreme Statistics of Storm Wave Heights

合田良実¹・久高将信²・河合弘泰³

Yoshimi GODA, Masanobu KUDAKA and Hiroyasu KAWAI

The *L*-moments method is applied for the POT storm wave data at eight stations around Japan, which have the record length of 27 to 38 years. The formulas of the *L*-moments of the Weibull distribution are newly derived for estimation of the shape, scale and location parameters from a sample. A new index called TUD (Twenty-Up Deviation) is developed for judgment of the degree of goodness of data fitting to a candidate distribution. Among eight stations, two stations show the best fitting to the Weibull distribution, four to the General Pareto (GPA) distribution, and two to the General Extreme-Value (GEV) distribution. At two stations fitted to the GPA distribution, the theoretical upper bounds of extreme wave heights are only 7% to 11% greater than the observed maximum wave height.

1. はじめに

波浪の極値統計解析にはこれまでもいろいろな方法が開発され、実用に供されてきた。しかし、近年はHosking (1990)が開発した*L*-moments法が水文統計その他の分野で活用されている。この方法は、同一の地域に所属するとみなされる複数地点の極値資料に対する母分布関数を求める手法(地域頻度解析と称される)の中心となっており、Hosking・Wallis (1997)の著した解説書が広く用いられている。水文分野以外では、外山・水野 (2002)が日本全国のアメダス地点の確率降水量の解析に用いており、高波についてはvan Gelderら (2000)がオランダ沿岸の波浪データをこの手法を使って解析している。

L-moments法では、極値の標本について独自に定義される1～4次の積率ならびに*L*-skewnessと*L*-kurtosisという量を求める。標本にあてはめる分布関数を選ぶと、その形状・尺度・位置母数がこれらの諸量から容易に求められる。具体的な方法はHosking・Wallis (1997)あるいは合田・久高 (2009)を参照されたい。

Hosking・Wallis (1997)は11種類のあてはめ分布関数の*L*-momentsおよび母数推定式を提示している。しかし、高波の極値統計で標準的に使われるワイブル分布に対するものが未提示のため、これに対する理論式を新しく導いた。そして、わが国沿岸のNOWPHASデータベースから波浪観測期間が長期にわたる8地点を選び、*L*-moments法の適用性を検討した。

2. 各種分布関数の*L*-momentsと母数推定式

波浪の極値統計では複数の分布関数を標本にあてはめ、適合度が最も良いものを対象地点の高波の母分布関数と見なすのが通例である。本論文では、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3種類の3母数型分布関数を取り上げる。GEV分布は極値I型とII型を含んでおり、GPAおよびワイブル分布は指数分布を包含している。Hosking・Wallis (1997)は4母数型および5母数型分布も扱っているが、今回は3母数型に限定して検討を進めた。なお、GPA分布については北野ほか (2002)が高波に適用した事例がある。

対象とした3種類の分布関数の関数形、極値の範囲、*L*-skewness (τ_3)と*L*-kurtosis (τ_4)の推定式を表-1、母数推定式を表-2にまとめて示す。また、 τ_4 と τ_3 の関係を図-1に示す。これらの推定式のうち、GEV・GPA分布はHosking・Wallis (1997)に記載されているものであり、

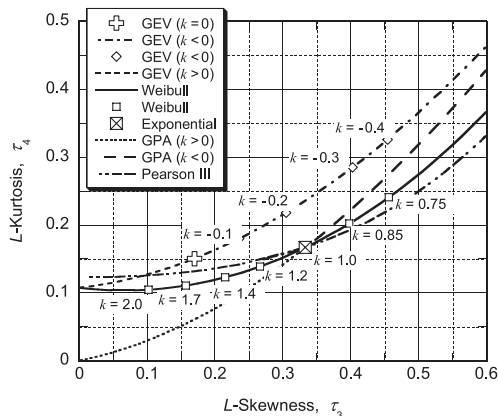


図-1 各種分布関数における *L*-skewness と *L*-kurtosis の関係

1 名誉会員 工博 (株)エコー 顧問
2 正会員 (株)エコー 防災・水工部担当 参与
3 正会員 修(工) (独法) 港湾空港研究所 海洋・水工部
海洋情報研究領域長心得

表-1 極値分布関数と L -skewness (τ_3) および L -kurtosis (τ_4)

分布	分布関数	範囲	L -skewness (τ_3)	L -kurtosis (τ_4)
GEV	$F(x) = \exp\{-\exp[1-k(x-B)/A]^{1/k}\}$	$k > 0: -\infty < x \leq B+A/k$ $k < 0: B-A/k \leq x < \infty$	$2(1-3^k)/(1-2^k)-3$	$[5(1-4^k)-10(1-3^k)+6(1-2^k)]/(1-2^k)$
GPA	$F(x) = 1-[1-k(x-B)/A]^{1/k}$	$k > 0: -\infty < x \leq B+A/k$ $k < 0: B-A/k \leq x < \infty$	$(1-k)/(3+k)$	$(1-k)(2-k)/[(3+k)(4+k)]$
ワイブル	$F(x) = 1-\exp\{-[(x-B)/A]^k\}$	$B \leq x < \infty$	$3-2(1-3^{-1/k})/(1-2^{-1/k})$	$[5(1-4^{1/k})-10(1-3^{1/k})+6(1-2^{1/k})]/(1-2^{1/k})$

表-2 極値分布関数の母数推定式

分布	形状母数, k	尺度母数, A	位置母数, B
GEV	$\approx 7.8590c + 2.9554c^2$ $c = 2/(3 + \tau_3) - \log 2 / \log 3$	$\lambda_2 k / [(1-2^k) \Gamma(1+k)]$	$\lambda_1 - A[1 - \Gamma(1+k)]/k$
GPA	$(1-3\tau_3)/(1+\tau_3)$	$(1+k)(2+k)\lambda_2$	$\lambda_1 - (2+k)\lambda_2$
ワイブル	$285.3 \lambda_3^6 - 658.6 \lambda_3^5 + 622.8 \lambda_3^4 - 317.2 \lambda_3^3$ $+ 98.52 \lambda_3^2 - 21.256 \lambda_3 + 3.5160$	$\lambda_2 k / [(1-2^{1/k}) \Gamma(1+1/k)]$	$\lambda_1 - A \Gamma(1+1/k)$

ワイブル分布に対するものは今回新しく導いたものである。ワイブル分布の形状母数の推定式は、あらかじめ k を入力として τ_3 の値を計算し、その関係に対して多次回帰曲線をあてはめたもので、近似精度は $\pm 0.2\%$ である。

図-1にはGEV分布 ($k < 0$) とワイブル分布で形状母数が特定の値のときの $\tau_4 - \tau_3$ 関係を示す点も記してある。GEV分布の白十字の点 ($k = 0$) はガンベル分布 (FT-I) であり、それから左へ緩やかに下がる線はGEV分布 ($k > 0$) で確率変数 x の値に上限があるケースである。白十字の点から右へ緩やかに上がる線 ($k < 0$) は極値 II 型分布 (FT-II) と同じである。ワイブル分布は $k > 0$ のGEV分布と $k > 0$ のGPA分布の間に位置するが、ワイブル分布は上限値がなく、右裾がどこまでも延びるのが特徴的である。

また、本論文ではPearson III型分布について触れていないが、図-1に示すようにこの分布とワイブル分布は L -moments比の値が隣接している。すなわち、後者はワイブル分布で代替させることが可能であるため、極値統計の標本当てはめの候補から外したものである。

ワイブル分布とGPA分布は指数分布（四角印にクロスの点）を共有している。指数分布の点の左下側では、両者の L -kurtosis の値に差異が見られる。すなわち、ワイブル分布では L -kurtosis の値が余り変化しないのに対し、GPA分布では L -skewness の減少につれて L -kurtosis も減少する。指数分布の点の右上側ではワイブル分布とGPA分布の違いが小さいが、後者のほうが L -kurtosis がやや大きな値を示す。

このように、 L -skewness と L -kurtosis の二つのパラメータを用いることによって、極値分布関数の差異を判別できる可能性が考えられる。図-1に示したのは母集団についての結果であり、データ数に限りのある単一の標本の場合には統計的変動性の影響によって母集団ほどの差別

化はむずかしいと考えられる。

3. 使用した高波資料の概要

先に筆頭著者はNOWPHASの1989年までのデータベースを使用し、日本沿岸の高波の極値分布関数について考察した（合田2008, pp. 374-375）。それから既に10年を経過し、波浪データベースも拡充されたところから、長期間にわたる波浪データが蓄積されている8地点を選び、POT法で抽出した高波資料の解析を行った。これらの高波資料の概要を表-3に示す。

観測年数は1年単位の欠測期間を除外したものであるが、月単位の欠測期間の調整は行っていない。また、合田・小長谷（1998）が行ったような、高波のピーク時の部分欠測も補填していない。また、気象原因別（台風系と低気圧系など）のデータ区分もおこなっていない。したがって、この資料だけでは確率波高の推定作業を進めることはできない。

表中のデータ数は、波高閾値が最も低いときのものであり、それから0.5mごとに閾値を高めていったので、解析を行ったデータ数は閾値ごとに異なる。

表-3 検討に用いた高波資料の概要

地点名	観測期間	年数	データ数	波高閾値 (m)	$(H_{1/3})_{\max}$ (m)
酒田	1970-2008	38	923	3.0 (0.5) 5.0	10.65
金沢	1970-2008	37	722	3.0 (0.5) 5.0	8.14
むつ小川原	1971-2008	37	551	2.5 (0.5) 5.0	9.56
鹿島	1972-2008	36	598	2.5 (0.5) 5.0	7.50
波浮	1973-2008	35	640	3.0 (0.5) 5.0	8.48
志布志	1980-2008	28	482	1.5 (0.5) 4.0	10.30
中城湾	1975-2008	27	599	2.0 (0.5) 5.0	12.08
那覇	1973-2008	35	656	2.5 (0.5) 4.5	9.24

4. L -moments法における最適分布の判定

標本に対する分布関数の適合度の良さを判定する方法は、あてはめ方法によって異なる。最小2乗法の場合に

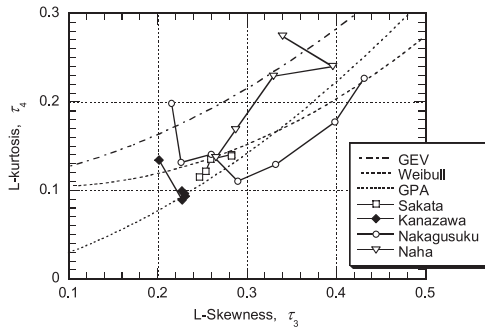


図-2 波高閾値を変えたときの沿岸波浪の L -skewness と L -kurtosis の関係

は、確率変量とその基準化変量との間の相関係数が高いほど適合度が良いとみなす。最尤法では尤度を判定基準に使うけれども、最適関数判定の感度がそれほど良いものではない。

L -moments法の場合には、まず L -skewness の値によって分布関数ごとの形状母数が算定される。そうすると、その分布関数の L -kurtosis の理論値が計算されるので、その値と標本から求めた値との差を判定基準とする方式が一つ考えられる。図-2は、沿岸8地点のうちの4地点について波高閾値を変えることによって、 L -skewness と L -kurtosis の関係がどのように変わるかも示したものである。

酒田港では $H_c = 3.0$ m のデータが左下に位置し、閾値を上げるにつれて表示点がワイブル分布の曲線に沿って右上に移動する。金沢港では $H_c = 5.0$ m のデータが一つだけ左上に離れており、他の閾値のデータは表示点が GPA の曲線上の一カ所に固まっている。中城湾の資料では、 $H_c = 2.0$ m のデータが右上に位置し、閾値を上げるにつれて表示点がワイブル分布の曲線の下方を左下へ向けて大きく移動し、 $H_c = 5.0$ m のときには GEV の関係曲線の左上へ飛び跳ねる。那覇港の場合には、 $H_c = 3.0$ m のデータが左下に位置し、閾値を上げるにつれて表示点が GEV の曲線へ向かって大きく右上に移っていく。

波高の閾値によって適合分布関数が増えることはある程度考えられるけれども、図-2に示す中城湾や那覇港のケースは変化が極めて大きく、単一の極値資料に対して L -skewness と L -kurtosis の関係だけに基いて最適の分布関数を選ぶ方式には疑問を抱かせる。実際に表-1の高波資料について、標本から得られた L -kurtosis の値とその理論値との差を分布関数ごとに比較し、最低値を示す分布関数を調べてみると、波高の閾値を上げるたびに異なる最適分布が得られてしまう。

たとえば、図-3は中城湾の高波データについて波高の閾値を $H_c = 2.0$ (1.0) 5.0 m と4通りに変えたときの波高データと、確率統計量として推定された波高値とを比較し

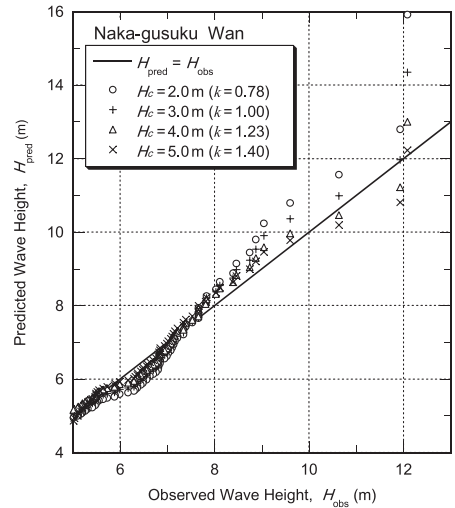


図-3 中城湾のデータに波高閾値を変えてワイブル分布を当てはめたときの順序統計量と確率統計量の比較 (Q-Q プロット)

た図である。データは波高の大きいほうから順に並べられており、こうした表示法はQuantile-Quantile (Q-Q) プロットと呼ばれる。ここでは当てはめ分布関数をワイブル分布に固定しているが、波高閾値によって形状母数が変化している。なお、このプロットに際してはプロット・ポジションとして Hosking・Wallis (1997) が紹介している式を降順第 m 位に対して変更した $P_m = 1 - (m - 0.65)/n$ を用いている。ただし、 n は標本の大きさ (データ数) である。

図-3から明らかなように、閾値の設定によって確率統計量が観測された波高の順序統計量と一致する度合いが相当に変化する。設計波高を選定するために極値統計解析を行う場合には、主として波高の上位のデータに対する極値分布関数の適合度に関心が持たれる。そこで、まず極値データの観測値の順序統計量と確率統計量の相対偏差 (%) を取り上げ、上位 20 個の 2 乗平均偏差 TUD (Twenty-Up Deviation) を適合度の指標として用いることを考えた。すなわち、

$$TUD(\%) = \sqrt{\frac{1}{20} \sum_{m=1}^{20} \left(\frac{H_{pred,m} - H_{obs,m}}{H_{obs,m}} \right)^2} \times 100 \quad \cdots \cdots (1)$$

20 個を取り上げたことに特別な理論的根拠はない。各種分布関数を当てはめたときの適合度を比較し、最もよく適合するものを選ぶ判断基準として使うためのものである。上位 10 個についても試行してみたところ、TUD の絶対値はやや大きくなったものの、分布関数ごとの適合度の順位は変わらなかった。

中城湾の高波極値データについて TUD による適合度を

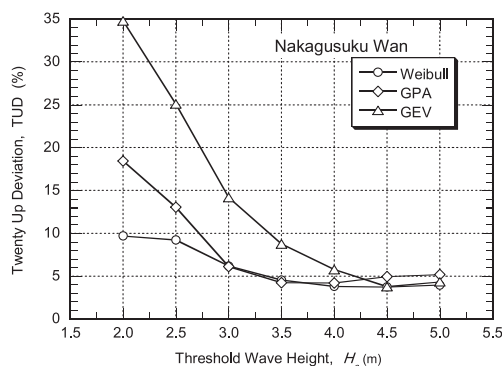
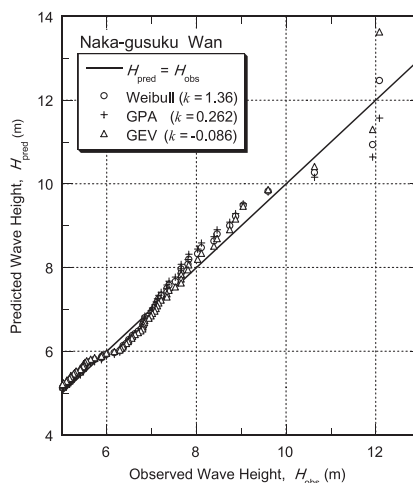


図-4 中城湾の高波データに対する TUD 値

図-5 中城湾の高波 ($H_c = 4.5$ m) に対する極値分布関数

計算すると図-4のようになり、高波抽出の閾値を上げるに従って適合度が高まることが分かった。先に図-3に示した中城湾の L -skewness と L -kurtosis の関係の変化では、閾値が低いときは適合関数が GPA、やがてワイブル分布、最後に GVE のように見えるけれども、図-4のように明確ではなく、また図-3 の方式では適合度を定量的に評価できない。

高波抽出の閾値によって適合度が変化するの、中城湾の場合には台風系と低気圧系という母集団の異なる高波が混在しているためである。合田ら (1998) の場合には抽出された高波の全てについて気象原因を確認し、台風系とそれ以外の高波に区分して解析したけれども、今回はそうした吟味を行っていない。波高の閾値が低い場合には波高が比較的小さい非台風系のデータが数多く含まれ、閾値を上げるに従って台風系の高波のみが解析対象となったものであろう。図-4によれば、閾値を $H_c \geq 4.0$ m とすることによって、TUD の値が安定するので、そうしたデータを対象として中城湾の高波の確率波高を計算すればよいと思われる。

5. 各地点の高波の極値分布特性

本節では、表-1に示した8地点の高波資料に対して L -moments 法で極値分布関数を当てはめた結果について述べる。図-5は中城湾の高波データであって、図-4の結果を参照して高波抽出の閾値を $H_c = 4.5$ m と設定した場合を

示している。この高波データでは観測値の第1位が12.08 m、第2位が11.93 m と近接しており、このため確率統計量の推定結果が前者に対しては過大、後者に対しては過小な結果となっている。

3種類の分布関数は適合度到大差がなく、どの関数も適用可能である。GEV 分布は順序統計量の第1位に対してやや過大な値を与えるけれども、第2位以下に対しては他の二つの分布関数よりも適合度がよい。

他の7地点に対しても同じように POT 法適用の波高閾値 H_c を TUD が最小になる波高として選定し、3種類の極値分布関数の中で最適分布を選定した結果を表-4に示す。この表から明らかのように、GEV、GPA、ワイブル分布が最適となるのがそれぞれ2、4、2地点と分散している。なお、データ数 n が表-3よりも少ないのは、閾値以下のデータを切り捨てたことによる。このうち、GPA 分布が最適となった金沢港と志布志港に対するあてはめ結果を図-6、7に示す。

両港とも、GPA 分布への適合は良好であるものの、この分布は極値に理論的上限值が存在する。表-4に記すように、金沢港では $H_{upper} = 8.73$ m と観測第1位の $H_{max} = 8.14$ m の7%増しでしかない。志布志港でも上限波高は $H_{upper} = 11.48$ m と $H_{max} = 10.30$ m の11%増しである。

表-4 沿岸8地点の高波資料に対する最適の極値分布関数のあてはめ結果

地点名	H_c (m)	n	最適分布	k	A (m)	B (m)	$H_{pred,1}$ (m)	$H_{obs,1}$ (m)	H_{upper} (m)
酒田	5.0	244	ワイブル	1.149	1.187	4.92	11.01	10.65	—
金沢	5.0	162	GPA	0.3294	1.236	4.98	8.24	8.14	8.73
むつ小川原	5.0	63	GEV	-0.201	0.513	5.49	10.18	9.56	—
鹿島	3.5	214	GPA	0.1265	1.110	3.47	8.02	7.50	12.24
波浮	5.0	66	GPA	0.2811	1.547	4.87	9.11	8.48	10.38
志布志	4.0	42	GPA	0.3704	2.868	3.74	10.17	10.30	11.48
中城湾	4.5	88	ワイブル	1.356	2.331	4.24	12.47	12.08	—
那覇	4.5	51	GEV	-0.250	0.460	4.86	9.40	9.24	—

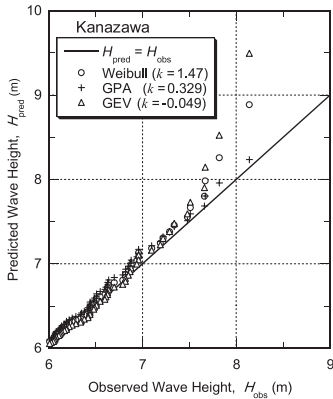


図-6 金沢港の高波 ($H_c = 5.0$ m) に対する極値分布関数のあてはめ結果

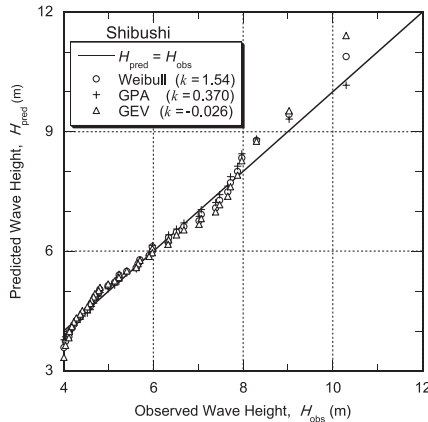


図-7 志布志港の高波 ($H_c = 4.0$ m) に対する極値分布関数のあてはめ結果

特に金沢港の場合には、日本海の波浪特性から見て8.7mを超える波高が発生しないとは考えにくい。筆者らは設計波高選定の立場から、こうした波高に上限値を持つGPA分布を高波の極値解析に用いることに疑念を抱く。適合度が若干劣っていても、ワイブル分布を適用すべきではなかろうか。

今回は単一地点ごとの解析であったが、いずれ波浪特性が類似の複数地点を対象として極値統計解析をおこない、共通の母分布関数を探索していきたいと考えている。

6. まとめ

本論文は、高波の極値統計解析にL-moments法を適用するための一つの試論であり、ここまでの段階で明らかになった事項を列挙すると以下のようになる。

- 1) L-moments法でこれまで考慮されていなかったワイブル分布を当てはめ候補の極値分布関数に含めることができた。
- 2) 標本から計算されるL-skewnessに基づいてワイブル分布、一般化極値 (GEV) 分布、および一般化パレート (GPA) 分布の形状・尺度・位置母数の3母数を求める方法が明確化された。
- 3) 標本に対する極値分布関数の適合度判定の指標として、順序統計量の上位20個の2乗平均偏差を用いるTUD (Twenty-Up Deviation) を新たに考案した。
- 4) 日本沿岸で長期間にわたる波浪観測データが利用できる8地点についてL-moments法による極値統計解析を試みたところ、ワイブル分布を最適とするものが2地点、GPA分布を最適とするものが4地点、GEV分布を最適とするものが2地点であった。
- 5) GPA分布は適合する地点が多いものの、確率統計量が上限値で頭打ちとなる危険性が例証された。

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 (2008) : 「耐波工学－港湾・海岸構造物の耐波設計－」, 鹿島出版会, 430p.
- 合田良実・久高将信 (2009) : 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・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.
- 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.
- 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 (1997): "Regional Frequency Analysis," Cambridge Univ. Press, 224p.
- van Gelder, P. H. A. J. M., J. De Ronde, and N. W. Neykov (2000): Regional frequency analysis of extreme wave heights: trading space for time, Coastal Engineering 2000 (Proc. 26th ICCE, Sydney), ASCE, pp. 1099-1112.