論文の内容の要旨

論文題目 On the Extreme Value Estimation of Significant Wave Height

（有義波高の極値推定について）

氏　　名 和田良太

Strength design of ocean structure requires knowledge of the maximum load it may encounter over its lifetime. Fifty or hundred year return period significant wave height is used to estimate the maximum wave load. Quantification of the extreme wave height and its confidence interval is a critical issue. This dissertation is about the extreme value estimation of significant wave heights.

Long duration of extreme data is required for precise extreme value estimation. We reviewed the characteristics of extreme wave data. Observation data and simulation data of extreme wave is always associated with data uncertainty, especially when we attempt to obtain long duration of data.

Extreme value estimation is often conducted with limitation on the number of data. For ocean regions prone to be hit by hurricanes or typhoons, such as ocean region around Japan, extreme waves are results of totally different physical process compared to normal conditions. Extreme data are independent from the “bulk” observations, and the rareness of events makes the available data for analysis even smaller. Smaller data size brings larger statistical uncertainty. Monte Carlo experiment of estimating extreme values from limited numbers of random sample data was conducted. For example, 100-year return value estimated from 30 year data of annual maximum indicated a wide variance.

We reviewed various methods of extreme value estimation, such as the maximum likelihood method, the method based on moment and so forth. Maximum likelihood estimation uses theoretical methods for confidence interval, but the asymptotic conditions are only valid when the data size is large. Resampling techniques are practical, but the estimated confidence intervals tend to be narrow. Some methods use empirical methods to estimate the confidence interval. The adequacy of the estimated confidence interval is a question.

We closely investigated Goda’s method (1988) of extreme value estimation, which is widely used in Japan. The comparison of Goda’s method and estimation based on likelihood inference, unveiled the shortcomings of Goda’s method. Goda’s method relies on empiricism, and fails to evaluate the statistical uncertainty considering various shapes of extreme value distribution.

Data uncertainty and statistical uncertainty are indispensable in extreme value estimation of significant wave height estimation. Previous estimation methods have focused on finding the best fitting parameter set, and had paid less attention to the estimation of confidence interval. The goal of this study is to quantify uncertainties that are associated with extreme value estimation of significant wave heights.

We propose an estimation method based on Bayesian statistics with group likelihood and flat prior, defined as Eq. (1) ~ Eq. (3). By utilizing flat prior, likelihood directly corresponds to the probability density function of the estimation. This is why we call our method Likelihood Weighted Method (LWM).

$$P(X|\theta) = L_G(\theta) = \prod_{i=1}^{n} \left\{ F\left(x_i + \Delta_i\right) - F\left(x_i - \Delta_i\right) \right\} \tag{1}$$

$$P(\theta) = p_\theta \tag{2}$$

$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{\int_\theta P(X|\theta)P(\theta)d\theta} = \frac{P(X|\theta)p_\theta}{p_\theta \int_\theta P(X|\theta)d\theta} = \frac{L_G(\theta)}{\int_\theta L_G(\theta)d\theta} \tag{3}$$

F : *cumulative density function of extreme value distribution,*

θ : *parameters of extreme value distribution,*

$P(\theta|X)$: *posterior density of* θ, $P(\theta)$: *prior density of* θ,

p_θ : *flat prior,* $L_G(\theta)$: *group likelihood,* $X = (x_1, x_2, , , x_n)$: *extreme data*

The use of group likelihood and flat prior makes LWM a unique and practical estimation method. We utilize group likelihood (Eq. (1)) proposed by Barnard (1967) and Giesbrechet & Kempthorne (1976). Their ideas are to utilize the round off errors or precisions that are always associated with the observation data. Group likelihood has two advantages. Method based on likelihood has limitation in the estimation due to non-regularity of extreme value distribution. The use of group likelihood avoids singularity and has no restriction. By defining round off errors for each data, group likelihood can explicitly introduce data uncertainty from observation and simulation in extreme value estimation.

The use of flat prior (Eq. (2)) makes the numerical implication of LWM simple and efficient. When using non-uniform priors, the integration is usually intractable, and numerical integration by Markov Chain Monte Carlo method (MCMC) is inevitable. MCMC required expertise and limited the use of Bayesian inference to academia. In LWM, numerical integration can be conducted with uniformly discretized parameters, and a simple rectangle method with mid-point approximation can be used for quadrature.

The statistical performance of LWM has been evaluated with Monte Carlo experiment. LWM was shown to be an efficient estimation method for various parameters of General Extreme Value distribution, General Pareto Distribution and 3 parameter Weibull distribution. The computation cost of LWM was compared with MCMC (Figure 1). LWM was shown to give accurate estimations with 1/200 computation cost. The evaluation results suggest that LWM is a practical method to estimate extreme values and their confidence intervals.

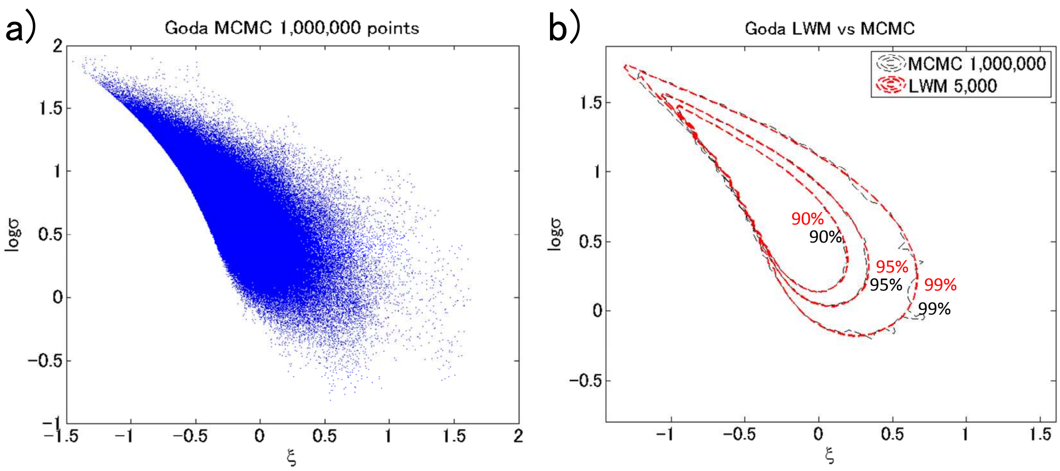

Figure 1 a) $P(X;\theta)$ estimated for extreme wave data from Goda (1988) by Markov Chain Monte Carlo method with 1,000,000 calculation steps. b) Red dashed lines indicate the confidence intervals of 90%, 95%, and 99% by LWM with 5,000 calculation steps. The black dashed lines show the same for MCMC with 1,000,000 calculation steps. The two estimations agree well.

Previously, extreme value estimation has focused on the best estimated extreme value. We propose *expected extreme value distribution* (EEV) to give a better description of the estimation and its confidence interval. EEV is a probability-weighted distribution of all the estimated extreme value distribution. The uncertainty of the estimation is inherited in the expected extreme values that are derived from EEV. The characteristics of EEV have been evaluated with Monte Carlo experiment and through application to observed wave data. For example, in the estimation of extreme wave height in Gulf of Mexico, the expected extreme value gave a better description of the observed largest wave height than the best estimate extreme value.

The results of LWM indicate a large epistemic uncertainty of the characteristic wave height. For hurricane-dominated ocean regions, large epistemic uncertainty in extreme value estimation of significant wave height is inevitable.

In common design rules, the estimated loads are multiplied by safety factors to account for the uncertainty in the estimation. In spite of the large epistemic uncertainty, only the aleatory uncertainty of the environmental condition is explicitly considered in the safety factor calibration. The safety factors are set to have redundancy considering various elements of uncertainty in estimating extreme load, and the adequacy of the safety factors has been empirically justified.

Some ocean structures have experienced exceedance of the estimated extreme load. These extreme events could have been anticipated if epistemic uncertainty has been considered in design rules. The impact of considering larger uncertainty in the significant wave height needs further research. At any rate, the quantification of uncertainties in extreme value estimation of significant wave height would contribute to a more sophisticated structure design.

[References]

[1] Bitner-Gregersen, Elzbieta M. "Uncertainties in data for the offshore environment." Structural Safety 7.1: 11-34, (1990).

[2] Goda Y., “On the methodology of selecting design wave height”, Proceedings of the 21st International Conference on Coastal Engineering, Malaga, Spain p. 899–913, (1988).

[3] Barnard, George A. "The use of the likelihood function in statistical practice." Proceedings of the Fifth Berkeley Symposium. Vol. 1. 1967.

[4] Giesbrecht, F., and O. Kempthorne. "Maximum likelihood estimation in the three-parameter lognormal distribution." Journal of the Royal Statistical Society. Series B (Methodological) (1976): 257-264.