Proceedings of OMAE2006
2006 25th International Conference on Offshore Mechanics and Arctic Engineering (OMAE 2006)
Hamburg, Germany, June 4-9, 2006

DRAFT / OMAE2006-92601

ESTIMATION OF WAVE HEIGHT RETURN PERIODS USING A NONSTATIONARY TIME SERIES MODELLING

Christos N. Stefanakos
School of Naval Architecture and Marine Engineering
National Technical University of Athens
9, Heroon Polytechniou str., 15773, Zografos, Athens, Greece
Email: chstef@naval.ntua.gr

also

Department of Shipbuilding
Technological Educational Institute of Athens
Ag. Spyridonos str, 12210, Aegaleo, Athens, Greece
Email: chstef@teiath.gr

Valérie Monbet
Department of Applied Statistics
University of South Brittany
BP 573, 56017 Vannes Cedex, France
Email: valerie.monbet@univ-ubs.fr

also

Hydrodynamics and Metocean
IFREMER/DCB/ERT/HO
BP 70, 29280 Plouzané, France

ABSTRACT

A new method for calculating return periods of various level values from nonstationary time series data is presented. The key-idea of the method is a new definition of the return period, based on the Mean Number of Upcrossings of the level x (MENU method). The whole procedure is numerically implemented and applied to long-term measured time series of significant wave height.

The method is compared with other more classical approaches that take into acount the time dependance for time series of significant wave height. Estimates of the extremal index are given and for each method bootstrap confidence intervals are computed.

The predictions obtained by means of MENU method are lower than the traditional predictions. This is in accordance with the results of other methods that take also into account the dependence structure of the examined time series.

INTRODUCTION

The design of most marine systems, including ships, offshore platforms and offshore wind farms, entails the long-term estimation of extreme values of wind and wave parameters. This calculation is traditionally performed using the so-called Gumbel's approach; a widely used method based on the annual maxima [1]; see also [2]. However, the accuracy of the method is highly dependent on the size of the examined dataset. Moreover, it is very difficult to have large datasets of *in situ* buoy measurements, which is the most reliable source of data.

There are numerous alternatives that use additional data. To mention only some of them, one could refer to: (i) the method of the r-largest maximum values [3, 4], (ii) the Peaks-Over-Threshold (POT) method [5–8], (iii) the enhanced POT method (that takes into account the seasonality of wind and wave data) [9], (iv) the BOLIVAR method [10, 11].

A further step would be to take into account the dependence structure of the time series by modelling it as, e.g., a stationary stochastic process and calculate its extremes based on the theory of these processes [12, 13]. However, existing wave datasets (measurements) has shown that can hardly be considered stationary.

Recently, more advanced models for the representation of time series of wind and wave parameters have been proposed; see, e.g., [14–19]. These more sophisticated time series models permit us to significantly improve stochastic predictions of extreme values. According to these, except for the stochastic character of the series, the dependence structure and the seasonality exhibited are appropriately modelled. Then, the theory of periodically correlated stochastic processes can be used for the extreme-value predictions; see, e.g., Middleton and Thompson [20].

Another interesting method is the so-called MENU method, according to which the H_S return period for prespecified H_S level values is the time period in which the MEan Number of Upcrossings of the level H_S becomes equal to unity. The MENU method has been presented for the first time in [21], where an application with Gaussian synthetic data is presented. A similar approach based on the characteristic function of the Gaussian process is shown in Naess [22]. A full account of MENU method can be found in [23].

In the present work, the MENU method is further developed, exploiting the nonstationary modelling of time series of H_S presented in [14], and the non-Gaussian modelling of the second-order probability structure presented in [24], to calculate return periods from nonstationary time series of H_S.

The results are also compared with results from traditional methods. In general, MENU method gives lower estimates of return values (design values) than the traditional methods do. This is in accordance with the findings of other works that take into account the dependence and seasonality features of the time series [10, 11, 25].

TRADITIONAL METHODS FOR RETURN PERIOD CALCULATIONS

Let us assume that, we have a stochastic sequence of independent, identically distributed (i.i.d.) random variables, corresponding to the sequence of successive maximum values of significant wave height at fixed time intervals (e.g., months, years, etc). This sequence is denoted by

$$X_{\max n} \quad n \qquad X_{\max n} \tag{1}$$

Predicting return periods and the corresponding design values by using data concerning maxima is known as the Gumbel's approach [1]. In this approach, when it is based on annual maxima, the distribution of the population of maxima, denoted by $G(x)$, is known as $n \to \infty$ [26]. In fact, it is the Generalized Extreme Value (GEV) distribution

$$G(x;\lambda,\delta,k) = \begin{cases} \exp\left\{-\left[1 - k\dfrac{x-\lambda}{\delta}\right]^{1/k}\right\} & k \neq 0 \\ \exp\left\{-\exp\left[-\dfrac{x-\lambda}{\delta}\right]\right\} & k = 0 \end{cases} \tag{2}$$

where $\lambda, \delta > 0$, and $-\infty < k < \infty$. Note that, for $k > 0$, the GEV distribution is supported on $(-\infty, \lambda + \delta/k]$, while for $k < 0$, it is supported on $[\lambda + \delta/k, \infty)$.

There is a variety of methods for the estimation of the parameters, most of them well-known and widely used: the maximum likelihood method (MLM) [27], the method of moments, as well as various types of linear tail-weighted least-squares methods (LSM).

In fact, the GEV distribution is a unifying representation of the three extreme-type distributions, which are obtained as follows: (i) FT-I (Gumbel distribution), for k=0, (ii) FT-II (Fréchet distribution), for $k < 0$, λ=1, $\delta = -k$, (iii) FT-III (reversed Weibull distribution), for $k > 0$, $\lambda = -1$, δ=k.

The estimation of return periods for high levels x is based on the formula

$$T_R(x) = \frac{\Delta\tau}{1 - G(x)} \tag{3}$$

where $G(x)$ is the extreme population distribution.

The use of $G(x)$ instead of the initial population distribution $F(x)$ in equation (3), is justified by the fact that the two distributions $G(x)$ and $F(x)$ are right-tailed equivalent; see, e.g., [28]. The selection of the type of $G(x)$ should be based on the behaviour of $G_{emp}(x)$ (empirical distribution) at the right tail only. This task can be processed by adopting a variety of short-cut procedures or by analytical methods either of purely statistical nature or based on a combination of physical observations and statistical arguments [28–30].

In ocean and coastal engineering practice, equation (3) is considered as a milestone for return-period calculations of various wave and wind parameters; see, e.g., [31]. One possible option is to model annual maxima of significant wave height by means of FT-I (Gumbel) distribution [32, 33]; a choice that will be used in the sequel. Figure 1 shows that this choice is convenient for the considered data.

In principle, Gumbel's approach applied to annual maxima forms a sound methodology for predicting long-term extreme values and the corresponding return periods. The most serious problem with this approach is the lack of sufficiently large extreme population data samples that would permit the type of distribution to be safely selected and its parameters to be reliably estimated.

Moreover, the strong hypothesis of the independency of the annual maxima, required for the application of Gumbel's method, can be surpassed since independency is usually achieved in shorter time periods than year. A useful tool to approximate the shorter time period associated with independent successive maxima is the extremal index.

Let $\{X_1, \ldots, X_n\}$ be a strictly stationary stochastic sequence and $\{Z_1, \ldots, Z_n\}$ i.i.d. random variables with the same distribution $F(\cdot)$. Let $M_n^x = \max\{X_1, \ldots, X_n\}$ and $M_n^z = \max\{Z_1, \ldots, Z_n\}$ be the maxima of n values from the two sequences: the stationary and the i.i.d., respectively. For independent variables, the followin well-known formula holds true

$$P(M_n^z \le y) = [P(Z \le y)]^n = F^n(y) \tag{4}$$

However, for dependent data the above relationship does not hold and the distribution of the maximum should be determined by using the complete distribution of the whole sequence. Luckily in many cases there is an approximate relationship that can be compared to eq. (4)

$$P(M_n^x \le y) \approx F^{\theta n}(y) \tag{5}$$

where $\theta \in [0, 1]$ is the so-called extremal index. An exact definition and estimators of extremal index can be found, e.g., in [34, 35].

In sea state time series, the observations X_t and $X_{t+\tau}$ are for sufficiently large time periods τ practically independent, and, on the other hand, the probability of two extreme observations occurring within the same time interval, which is not too long, is small. These qualitative statements can be formulated as precise criteria of time series that have an extremal index of 1, the exact formulation of which will not be given here.

RETURN PERIODS FOR NONSTATIONARY STOCHASTIC PROCESSES

In this section, a new definition of the notion of return period associated with a given level value is introduced, which is valid even if the underlying stochastic process is a general nonstationary one. This definition will be based on the so-called "one-sided barrier problem" or "one-sided first passage problem" [36].

Assume that $\{X(\tau;\beta), \beta \in B\}$ is a nonstationary stochastic process with mean-square differentiable path functions. Let $M(x; t, t+T)$ be the Mean Number of Upcrossings of the level x by the process $\{X(\tau;\beta), \beta \in B\}$ in the time interval $[t, t+T]$.

Definition 1. *When $M(x; t, t+T)$ becomes equal to unity, the time lag (interval) $T = (t+T) - t$ will be called the return period of $X(\tau;\beta)$ associated with the level value x and the starting time t, and it will be denoted by $T_R(x, t)$.*

This definition has been first used in the context of statistical prediction of sea-level extremes in [20]; see also [37].

The MENU method

Now, in order to implement Definition 1, we have to calculate the mean number of upcrossings of the level x by the nonstationary process $X(\tau;\beta)$. As it is well known [12], an upcrossing of the level x by the process $X(\tau;\beta)$ occurs when

$$X(\tau;\beta) = x \quad and \quad \frac{dX(\tau;\beta)}{d\tau} > 0 \tag{6}$$

The total number of the upcrossings of the level x within the time interval $[t_1, t_2]$ is given by the equation

$$C_R(x; t_1, t_2; \beta) = \frac{1}{2} \int_{t_1}^{t_2} |\dot{X}(\tau;\beta)| \, \delta(X(\tau;\beta) - x) \, d\tau \tag{7}$$

where $\delta(\cdot)$ is the Dirac delta function. Equation (7) was first derived by Rice [12]. Recalling now that, for any (possibly generalized) function $G(x, y)$

$$\mathsf{E}^{\beta}[G(X(\tau_1;\beta), Y(\tau_2;\beta))] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(x, y) f_{\tau_1 \tau_2}(x, y) \, dx dy \tag{8}$$

where $f_{\tau_1 \tau_2}(x, y)$ is the joint probability density of the random vector $(X(\tau_1;\beta), Y(\tau_2;\beta))$, and applying the ensemble average operator $\mathsf{E}^{\beta}[\cdot]$ in both sides of equation (7), we obtain

$$\mathsf{E}^{\beta}[C_R(x; t_1, t_2; \beta)] = \frac{1}{2} \int_{t_1}^{t_2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\dot{x}| \, \delta(x - x) \, f_{\tau\tau}(x, \dot{x}) \, dx d\dot{x} d\tau \tag{9}$$

Integrating equ. (9) with respect to x, and observing that $\mathsf{E}^{\beta}[C_R(x; t_1, t_2; \beta)]$ is not but the quantity $M(x; t_1, t_2)$ appearing in Definition 1 above, we easily obtain

$$M(x; t_1, t_2) = \int_{t_1}^{t_2} J(x; \tau) \, d\tau \tag{10}$$

where

$$J(x; \tau) = \frac{1}{2} \int_{-\infty}^{\infty} |\dot{x}| \, f_{\tau\tau}(x, \dot{x}) \, d\dot{x} \tag{11}$$

Observing that, $f_{\tau\tau}(x, \dot{x})$ is an even function with respect to the second argument $\dot{x}$ (see, e.g., Levine [38, Vol. 1, pp. 201, 424]), equ. (11) can also be written in the following form

$$J(x;\tau) = \int_0^{\infty} \dot{x} f_{\tau\tau}(x, \dot{x})\, d\dot{x} \tag{12}$$

Note that, relation (10) is totally independent from the specific type of the underlying bivariate distribution considered.

It should be emphasized that $M(x;t_1, t_2)$ is dependent on both time instants t_1 and t_2, since the underlying process $\{X(\tau;\beta), \beta \in B\}$ may be nonstationary.

In accordance with Definition 1, the return period $T_R(x, t)$ associated with the level x and the starting time t is calculated as the unique value T for which

$$M(x;t, t+T) = 1 \tag{13}$$

Let it be noted that, the above definition of return period $T_R(x, t)$ can be applied for any level value x (high or low) and any starting time t. Accordingly, seasonal weather windows can be obtained in this manner.

Implementation of MENU method

Let us now turn to the calculation of the mean number of upcrossings $M(x;t, t+T)$ (and of the associated return periods for various level values x) for a specific form of periodically-correlated stochastic processes. Assume that, a stochastic process $X(\tau;\beta)$ admits of the representation [14]:

$$X(\tau;\beta) = G(\tau) + \sigma(\tau) W(\tau;\beta) \tag{14}$$

where $G(\tau)$ and $\sigma(\tau)$ are deterministic time-dependent periodic functions and $W(\tau;\beta)$ is a zero-mean stationary stochastic process, which will be called hereafter the residual stochastic process.

Especially, the function $G(\tau)$ is defined as

$$G(\tau) = \overline{X}_{mean} + \mu(\tau) \tag{15}$$

where $\overline{X}_{mean}$ is the overall mean value of the process $X(\tau)$, and $\mu(\tau)$ is the seasonal mean value [14]. Note that, sometimes a slowly varying (e.g. linear) trend is used instead of a fixed mean value. However, since the identification of the trend is still questionable, it has been decided to discard it, using instead the overall mean value $\overline{X}_{mean}$. In addition, the function $\sigma(\tau)$ is the seasonal standard deviation of the process.

For the numerical implementation of the MENU method use is only made of the joint probability density function $f^{X\dot{X}}_{\tau\tau}(s_1, s_2)$. In order to evaluate the latter we need, apart from the representation of the process, eq. (14), the representation of the process $\dot{X}(\tau;\beta)$:

$$\dot{X}(\tau;\beta) = \dot{G}(\tau) + \dot{\sigma}(\tau) W(\tau;\beta) + \sigma(\tau) \dot{W}(\tau;\beta) \tag{16}$$

obtained by differentiating eq. (14). The functions $\dot{G}(\tau)$ and $\dot{\sigma}(\tau)$ are the derivatives of $G(\tau)$ and $\sigma(\tau)$ with respect to τ. According to their definition, all these functions $\dot{G}(\tau)$ and $\dot{\sigma}(\tau)$ are periodic.

Equations (14) and (16) are considered as a linear system (time-dependent transformation) defining $X(\tau;\beta)$ and $\dot{X}(\tau;\beta)$ by means of $W(\tau;\beta)$ and $\dot{W}(\tau;\beta)$ [1]. Thus, the problem is reduced to the evaluation of the time-invariant bivariate pdf $f^{W\dot{W}}_{\tau\tau}(v_1, v_2)$.

The density $f^{W\dot{W}}_{\tau\tau}(v_1, v_2)$ can be obtained by means of various ways; see [24]. In the present work, the process $\dot{W}(\tau)$ is approximated using finite differences, and the density $f^{W\dot{W}}_{\tau\tau}(v_1, v_2)$ is calculated from the density $f^{WW}_{\tau\tau+\Delta\tau}(u_1, u_2)$ (second-order density of $W(\tau)$ at the time instances τ and $\tau+\Delta\tau$), by means of the bivariate linear transformation:

$$\begin{aligned} W(\tau) &= W(\tau) \\ \dot{W}(\tau) &= \frac{W(\tau+\Delta\tau) - W(\tau)}{\Delta\tau} \end{aligned} \tag{17}$$

where $\Delta\tau$ is the (fixed) sampling interval of the stationary residual series $W(\tau)$.

Thus, applying the standard formula of the change of variables, we get [39]

$$f^{W\dot{W}}_{\tau\tau}(v_1, v_2) = \Delta\tau\, f^{WW}_{\tau\tau+\Delta\tau}(v_1, v_1 + \Delta\tau\, v_2) \tag{18}$$

Considering eqs. (14) and (16) as a bivariate time-dependent linear transformation, and applying once again the standard formula of the change of variables, the density $f^{X\dot{X}}_{\tau\tau}(s_1, s_2)$ is expressed in terms of the density $f^{W\dot{W}}_{\tau\tau}(v_1, v_2)$ as follows:

$$f^{X\dot{X}}_{\tau\tau}(s_1, s_2) = \frac{1}{\sigma_\tau^2} f^{W\dot{W}}_{\tau\tau}\left(\frac{s_1 - g_\tau}{\sigma_\tau}, \frac{(s_2 - \dot{g}_\tau)\sigma_\tau - (s_1 - g_\tau)\dot{\sigma}_\tau}{\sigma_\tau^2}\right) \tag{19}$$

where $g_\tau = G(\tau)$, $\sigma_\tau = \sigma(\tau)$, $\dot{g}_\tau = \dot{G}(\tau)$, $\dot{\sigma}_\tau = \dot{\sigma}(\tau)$.

[1] From now on, for brevity and without loss of generality, the chance variable β will be omitted from the argument of the stochastic processes $X(\tau)$, $\dot{X}(\tau)$, $W(\tau)$, $\dot{W}(\tau)$.

Thus, the whole procedure can be considered as a subsequent application of two bivariate transformations of the form

$$w_1 \; \dot{w}_1 \qquad h \; w_1 \; w_2 \qquad w_1 \; \frac{w_2 \quad w_1}{\Delta\tau} \tag{20}$$

$$x \; \dot{x} \qquad w \; \dot{w} \qquad G \; \tau \quad \sigma \; \tau \; w \quad \dot{G} \; \tau \quad \dot{\sigma} \; \tau \; w \quad \sigma \; \tau \; \dot{w} \tag{21}$$

Let us now turn to the calculation of the quantity $M\; x \;; t_1 \; t_2$, by first examining the integrand in the right-hand side of equ. (10). This is a time-dependent function $J\; x \;; \tau$ for fixed its first argument $x = fixed$. The integrand reveals a periodic character, which is due to the periodic (with period T_{year} 1 year) functions included in the density $f^{X\dot{X}}_{\tau\tau}\; s_1 \; s_2$. Taking into account this character, the integral of the right-hand side of equ. (10) is simplified as follows

$$M\; x \;; t_1 \; t_2 \qquad n \int_{t_1}^{t_1 \quad T_{year}} J\; x \;; \tau \; d\tau \quad \Re \; t_1 \; t_2; T_{year} \tag{22}$$

where $n \quad t_2 \quad t_1 \quad T_{year}$ (x integer part of x).

The residual part $\Re \; t_1 \; t_2; T$ becomes zero, if we select $t_2 = nT_{year}$.

Finally, according to Definition 1, the time interval $T_R \quad T$, which corresponds to the value $M\; x \quad t \; t \quad T$ =1, will be called the return period of $X \; \tau$ associated with the level value x and the starting point t. The above described calculations are repeated for various fixed level values x of $X \; \tau$. In this way, we obtain the well-known return-value diagram.

APPLICATION TO TIME SERIES OF SIGNIFICANT WAVE HEIGHT

In the present section, the MENU method will be applied to time series of significant wave height to obtain return period estimates. The results will be compared with results of the following methods: a) the *standard (annual) Gumbel's method*, where the return periods are extrapolated for annual maxima, and b) the *weekly Gumbel's method*, where weekly maxima are considered instead of annual ones.

For this purpose, we will use two datasets: one WAM (hindcast) datapoint with coordinates (45^{o}N, 12^{o}W) in the central Atlantic Ocean and one from buoy measurements (51.85^{o}N, 155.92^{o}W, NOAA's buoy 46003) in the north Pacific Ocean. Hindcast data are widely used in design problems because they are available in a worldwide grid. However, it is well known that hindcast time series are particularly smooth and may underestimate significant wave height. This is the reason we shall also compare the method using bouy time series. On the other hand, *in situ* measurements have missing values that have been assumed to be missing at random and they have not been filled in.

In MENU method, the joint probability density function $f^{WW}_{\tau\,\tau \;\; \Delta\tau}\; u_1 \; u_2$, i.e. the second-order density of $W \; \tau$ at the time instances τ and ($\tau \quad \Delta\tau$) is modeled by a Plackett copula model [24]. Other copulas have been tested, in particular extreme value copula, but none of them fits so well to the data as Plackett does. The two alternatives of Gumbel's method are directly applied to the time series of significant wave height, without any particular time series modeling.

For each method, the return periods are estimated and associated 90% bootstrap confidence intervals are computed. In the case of the Gumbel's methods, the distribution fitted to the data is first used to generate a large number of samples. Then, a Gumbel distribution is fitted to these artificial samples and finally confidence intervals are deduced. Since MENU method is considered to be an exact method, confidence intervals can be omitted. If confidence intervals are to be added, artificial time series have to be generated as follows. Firstly the seasonal component of observed time series is removed using the model (14) in order to obtain the stationary process $W \; \tau$. Then $W \; \tau$ is transformed to a Gaussian process $Wg \; \tau$ using g-transform of [40] and a large number of artificial Gaussian time series with the same spectrum as $Wg \; \tau$ is generated. Finally inverse transformations are applied. It may be verified that the generated time series and the reference ones have the same statistical properties as in [41].

Hindcast data

As a first example, we present results computed using the hindcast time series of significant wave height. Before going further, the distribution of Gumbel's model (scale=1.3,location=13.6) is compared with the distribution of the GEV model (shape=-0.02, scale=1.1, location=13.7), using the time series of annual maxima. In Figure 1, a well-known result is shown on gumbel paper, i.e. the confidence intervals of GEV distribution are larger than the Gumbel's ones.

Then, the extremal index is estimated for the stationarized time series and plotted on Figure 2. The Ferro and Segers estimator is used. In practice the extremal index is compute for eaxh year and the mean is plotted. One can observe that if weekly maxima are considered for the estimation of return periods, the extremal index is equal to unity independently of the level chosen for the time series of weekly maxima. As in Ferro and Segers, we also deduce the run length (number of points in each independant cluster) with bootstrap confidence intervals (Fig. 3). We observe that the run length is close to 1 for all levels of the time series of the weekly maxima.

In the sequel, the results of MENU method are compared with the ones of the two Gumbel's methods; see Figure 4. We observe that the estimation of the return period is not signifi-

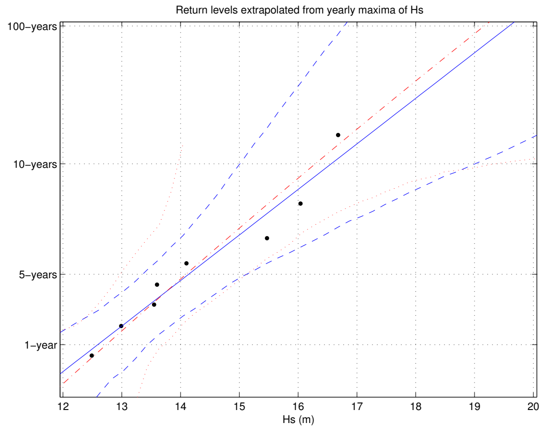

Figure 1. Distribution (on Gumbel paper) for the hindcast data. Dots : observed annual maxima, Plain line: Gumbel distribution (dashed line : corresponding 90% confidence interval), Dashed-dotted: GEV (dotted line : corresponding 90% confidence interval)

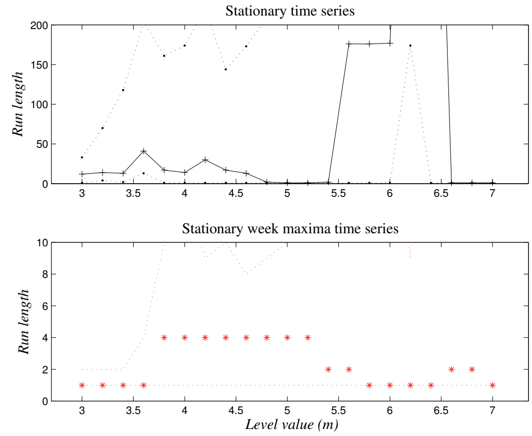

Figure 3. Run length for the hindcast data ; Top : original stationary time series (-+) with .95% confidence interval (:.); Bottom : weekly maxima (-*) with .95% confidence interval (:)

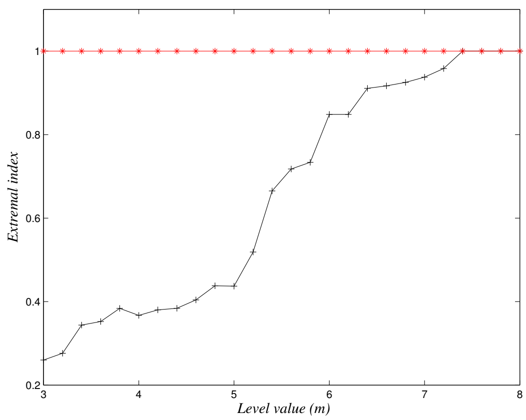

Figure 2. Extremal index for the hindcast data, original stationary time series (-+), weekly maxima (-*)

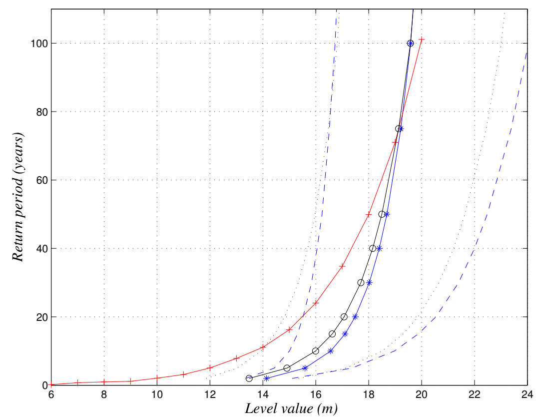

Figure 4. Return values for hindcast data estimated from Gumbel's approach with annual maxima (blue -*), Gumbel's approach with weekly maxima (black o-) and MENU method (red +-). (Confidence intervals are plotted for Gumbel's estimators as follows: annual maxima dashed blue line, weekly maxima dotted black line).

cantly different. The confidence intervals of the weekly Gumbel's model (scale=1.5,location=6.2) are narrower than the ones of the annual model. Moreover, for lower levels of significant wave height the MENU method gives larger estimates for return periods than the two Gumbel's approaches. One possible reason could be that hindcast time series is particularly smooth (as it is often the case in hindcast datasets) and it tends to lead to larger duration of exceedance time than in buoy measurements. For higher values of significant wave height the MENU method gives lower return periods, although the results from all three methods are very close.

Buoy data

The return values estimators are now obtained using data from NOAA's buoy 46003. The data preprocessing showed that the time series consists of 21 years with a 3-hourly time interval and about 25% missing values. The extremal index is estimated for the stationarized time series and plotted in Figure 5. We can conclude again that the use of weekly maxima leads to constant extemal index equal to unity.

Then, the return periods are computed and compared in Figure 6. One can remark that the obtained estimations are quite

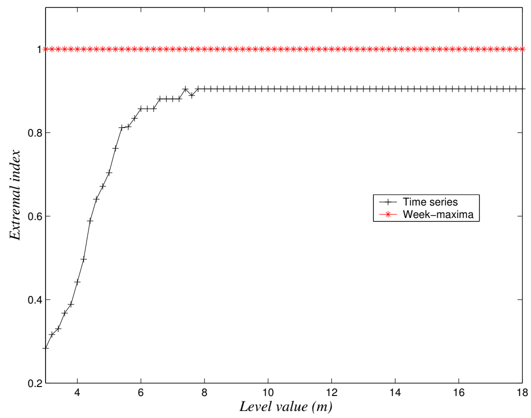

Figure 5. Extremal index for NOAA buoy data, original stationary time series (-+), weekly maxima (-*)

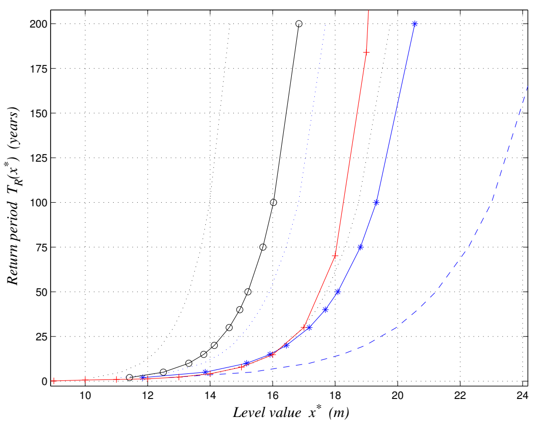

Figure 6. Return values for NOAA buoy data estimated from Gumbel's approach with annual maxima (blue -*), Gumbels approach with week maxima (black o-) and MENU method (red +-); for Gumbel's estimations confidence intervals are plotted, annual maxima : dashed blue line, weekly maxima : dotted black line.

different. For instance, concerning the 100-year return period, the annual Gumbel's method (scale=2.1, location=7.3) yields a return value approximately equal to 19 m, the weekly Gumbel's method (scale=1.3, location=2.7) to 16 m and the MENU method 18.3 m. Note also that, only MENU estimates lie in the intersection of both confidence intervals.

CONCLUSIONS

In this paper, a new method for calculating return periods of various level values from nonstationary time series is applied to long-term data of significant wave height. The key idea of the method is a new definition of the return period concept based on the MEan Number of Upcrossings of a level value x (MENU method).

In the present work, the MENU method is applied for the first time to non-Gaussian data, and return-period calculations for various level values are produced by means of real data. The present application takes fully advantage of the appropriate modelling of the second-order probability structure of the stochastic process developed in [24].

Results are compared with results obtained from two variants of Gumbel's approach (annual and weekly maxima). All comparison results show that predictions based on MENU method are in agreement with predictions from traditional methods.

ACKNOWLEDGMENT

This work has been financially supported by the Archimedes program (TEI of Athens), co-funded by the European Union (75%) and the Greek Government (25%) under the framework of the Educational and Initial Vocational Training Program - Archimedes.

REFERENCES

[1] Gumbel, E., 1958. *Statistics of extremes*. Columbia University Press, New York.

[2] Ochi, M., 1998. *Ocean waves: The stochastic approach*. Cambridge University Press, Cambridge.

[3] Tawn, J., 1988. "An extreme value theory model for dependent observations". *Journal of Hydrology,* **101** , pp. 227–250.

[4] Guedes Soares, C., and Scotto, M., 2004. "Application of the *r* largest-order statistics for long-term predictions of significant wave height". *Coastal Engineering,* **51** , pp. 387–394.

[5] van Vledder, G., and Zitman, T., 1992. "Design waves: Statistics and engineering practice". In Proceedings of the Second International Offshore and Polar Engineering Conference, vol. III, pp. 170–178.

[6] van Vledder, G., Goda, Y., Hawkes, P., Mansard, E., Martin, M., Mathiesen, M., Peltier, E., and Thompson, E., 1994. "Case studies of extreme wave analysis - a comparative analysis". In Ocean Wave Measurement and Analysis, O. Magoon and J. Hemsley, Eds., American Society of Civil Engineers & National Oceanic and Atmospheric Administration, pp. 977–991.

[7] Ferreira, J., and Guedes Soares, C., 1998. "An application of the Peaks Over Threshold method to predict extremes of significant wave height". *Journal of Offshore Mechanics and Arctic Engineering,* **120** , pp. 165–176.

[8] Naess, A., 1998. "Statistical extrapolation of extreme value data based on the Peaks Over Threshold method". *Journal of Offshore Mechanics and Arctic Engineering,* **120** , pp. 91–96.

[9] Morton, I., Bowers, J., and Mould, G., 1997. "Estimating return period wave heights and wind speeds using a seasonal point process model". *Coastal Engineering,* **31** , pp. 305–326.

[10] Boukhanovsky, A., Lopatoukhin, L., and Rozhkov, V., 1997. "Approach, experience, software and examples of wind wave climate investigations". In Proceedings of the International Conference RAO'97.

[11] Rozhkov, V., Boukhanovsky, A., and Lopatoukhin, L., 1999. "Method for calculation of extreme metocean events". In Proceedings of the International Medcoast Conference on Wind and Wave Climate in the Mediterranean and the Black Sea.

[12] Rice, S., 1944/1945. "Mathematical analysis of random noise". *Bell System Techical Journal,* **23/24** , pp. 282–332/46–156. (Selected Papers on Noise and Stochastic Processes, N. Wax (ed.), Dover, New York, 1954).

[13] Leadbetter, M., Lindgren, G., and Rootzen, H., 1983. *Extremes and Related Properties of Random Sequences and Series*. Springer-Verlag, New York.

[14] Athanassoulis, G., and Stefanakos, C., 1995. "A nonstationary stochastic model for long-term time series of significant wave height". Journal of Geophysical Research, **100** (C8) , pp. 16149–16162.

[15] Guedes Soares, C., and Ferreira, J., 1996. "Representation of non-stationary time series of significant wave height". *Probabilistic Engineering Mechanics,* **11** , pp. 139–148.

[16] Cunha, C., and Guedes Soares, C., 1999. "On the choice of data transformation for modelling time series of significant wave height". *Ocean Engineering,* **26** , pp. 489–506.

[17] Stefanakos, C., and Athanassoulis, G., 2001. "A unified methodology for the analysis, completion and simulation of nonstationary time series with missing-values, with application to wave data". *Applied Ocean Research,* **23/4** , pp. 207–220.

[18] Stefanakos, C., and Athanassoulis, G., 2003. "Bivariate stochastic simulation based on nonstationary time series modelling". In 13th International Offshore and Polar Engineering Conference, ISOPE'2003, pp. 46–50.

[19] Stefanakos, C., and Belibassakis, K., 2005. "Nonstationary stochastic modelling of multivariate long-term wind and wave data". In 24th International Conference on Offshore Mechanics and Arctic Engineering, OMAE'2005.

[20] Middleton, J., and Thompson, K., 1986. "Return periods of extreme sea levels from short records". Journal of Geophysical Research, **91** (C10) , pp. 11707–11716.

[21] Athanassoulis, G., Soukissian, T., and Stefanakos, C., 1995. "Long-term variability and its impact to the extreme-value prediction from time series of significant wave height". In Fourth International Workshop on Wave Hindcasting and Forecasting: Climate Variability and Extremes, V. Swail, Ed., Atmospheric Environment Service, pp. 343–358.

[22] Naess, A., 2001. "Crossing rate statistics of quadratic transformations of Gaussian processes". *Probabilistic Engineering Mechanics,* **16** , pp. 209–217.

[23] Stefanakos, C., and Athanassoulis, G., 2006. "Extreme value predictions based on nonstationary time series of wave data". Environmetrics, **17** (1) , pp. 25–46.

[24] Stefanakos, C., 1999. *Nonstationary stochastic modelling of time series with applications to environmental data*. PhD thesis, National Technical University of Athens, Dept. of Naval Architecture and Marine Engineering, Athens, Greece, June.

[25] Sobey, R., and Orloff, L., 1995. "Triple annual maximum series in wave climate analyses". *Coastal Engineering,* **26** , pp. 135–151.

[26] Galambos, J., 1978. *The Asymptotic Theory of Extreme Order Statistics*. John Wiley and Sons, New York.

[27] Prescott, P., and Walden, A., 1980. "Maximum-likelihood estimation of the parameters of the three-parameter generalized extreme-value distribution". *Biometrika,* **67** , pp. 723–724.

[28] Castillo, E., 1988. *Extreme Value Theory in Engineering*. Academic Press, Boston.

[29] Muir, L., and El-Shaarawi, A., 1986. "On the calculation of extreme wave heights: A review". Ocean Engineering, **13** (1) , pp. 93–118.

[30] Castillo, E., and Sarabia, J., 1992. "Engineering analysis of extreme value data: Selection of models". Journal of Waterway, Port, Coastal, and Ocean Engineering, **118** (2) , pp. 129–146.

[31] Ochi, M., 1982. "Stochastic analysis and probabilistic prediction of random seas". *Advances in Hydroscience,* **13** , pp. 217–375.

[32] Carter, D., and Challenor, P., 1983. "Methods of fitting the Fisher-Tippett type I extreme value distribution". Ocean Engineering, **10** (3) , pp. 191–199.

[33] Labeyrie, J., 1991. "Time scales and statistical uncertainties in the prediction of extreme environmental conditions". *Reliability Engineering and System Safety,* **32** , pp. 243–266.

[34] Ferro, C., and Segers, J., 2003. "Inference for clusters of extreme values". *Journal of the Royal Statistical Society, Series B,* **65** , pp. 545–556.

[35] Beirlant, J., Goegebeur, Y., Teugels, J., and Segers, J., 2004. *Statistics of extremes*. John Wiley and Sons, Chichester.

[36] Abrahams, J., 1986. "A survey of recent progress on level crossing problems for random processes". In *Communications and Networks. A survey of recent advances*, I. Blake and H. Poor, Eds. Springer-Verlag, pp. 6–25.

[37] Hamon, B., and Middleton, J., 1989. "Return periods of extreme sea levels: The exceedance probability method". Int. Hydrogr. Rev., **66** (2) , pp. 165–177.
[38] Levine, B., 1973. *Fondements Théoriques de la Radiotechnique Statistique*. Editions MIR, Moscou.
[39] Larson, H., and Shubert, B., 1979. *Probabilistic models in engineering sciences.Vol I: Random variables and stochastic processes*. John Wiley & Sons, New York.
[40] Rychlik, I., Johannesson, P., and Leadbetter, M., 1997. "Modeling and statistical analysis of ocean wave data using transformed Gaussian processes". *Marine Structures,* **10** , pp. 13–47.
[41] Monbet, V., Ailliot, P., and Prevosto, M., 2005. "Survey of stochastic models for wind and sea-state time series". *Submitted to Probabilistic Engineering Mechanics* .