Evaluating environmental joint extremes for the offshore industry using the conditional extremes model

Kevin Ewans [a,*], Philip Jonathan [b]

[a] Sarawak Shell Bhd., 50450 Kuala Lumpur, Malaysia
[b] Shell Projects and Technology Thornton, CH1 3SH, UK

ARTICLE INFO

ABSTRACT

Understanding extreme ocean environments and their interaction with fixed and floating structures is critical for the design of offshore and coastal facilities. The joint effect of various ocean variables on extreme responses of offshore structures is fundamental in determining the design loads. For example, it is known that mean values of wave periods tend to increase with increasing storm intensity, and a floating system responds in a complex way to both variables.

Specification of joint extremes in design criteria has often been somewhat ad hoc, being based on fairly arbitrary combinations of extremes of variables estimated independently. Such approaches are even outlined in design guidelines. Mathematically more consistent estimates of the joint occurrence of extreme environmental variables fall into two camps in the offshore industry — response-based and response-independent. Both are outlined here, with emphasis on response-independent methods, particularly those based on the conditional extremes model recently introduced by (Heffernan and Tawn, 2004), which has a solid theoretical motivation. We illustrate an application of the conditional extremes model to joint estimation of extreme storm peak significant wave height and peak period at a northern North Sea location, incorporating storm direction as a model covariate. We also discuss joint estimation of extreme current profiles with depth off the North West Shelf of Australia. Methods such as the conditional extremes model provide valuable additions to the metocean engineer's toolkit.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Offshore structures must be designed for very low probabilities of failure due to storm loading. Design codes stipulate that offshore structures should be designed to exceed specific levels of reliability, expressed in terms of an annual probability of failure or return-period. This requires specification of values of environmental variables with very low probabilities of occurrence. More specifically, since the goal is to determine structural loading due to environmental forcing, it is the combination of environmental phenomena with a given return-period that is sought. For example, most physical systems respond to environmental conditions in a manner that cannot be represented by a single variable — the pitch of a vessel is as much a function of the wave period or wave length as it is of the wave height, and it is necessary to also specify appropriate associated values of period for a given extreme wave height.

The goal is thus to design an offshore facility to withstand extreme environmental conditions that will occur during its lifetime with an appropriate optimum risk level. The level of risk is set by weighing the consequences of failure against the cost of over-designing. Facilities with a 20 to 30 year lifetime generally use 100-year metocean criteria, which with typical implicit and explicit safety factors, lead to annual probabilities of failure of 10^{-3} to 10^{-5}. For example, the load-resistance factor design (ISO, API) have an environmental load factor, $\gamma_E = 1.35$, to use with the loading calculated from an appropriate combination of environmental variables with a return-period of 100 years. The challenge is the choice of the appropriate combination of environmental variables through extreme value analyses.

Estimation of the extremes of single variables is relatively straightforward, given a long time series or time history of that variable that spans many years, and as a consequence, combinations of independently derived variables are often used for estimating environmental forces. One could for example, use the maximum wave height, wind speed, and current speed each with a return period of 100 years to derive the environmental loading with a return period of 100 years, but unless the winds, waves, and currents are perfectly correlated, the probability of this combination of variables is considerably less than 0.01 per annum. Some design codes and guidelines, would suggest taking the 100 year return period of one variable together with the value of an associated variable for a shorter return period. For example, the DNV recommended practice for on-bottom stability of pipelines suggests the combination of the 100-year return condition for waves combined with the 10-year return condition for current or vice-versa, when detailed information about the joint probability of waves and current is not available. Without

---

* Corresponding author. Tel.: +60 85453498.
E-mail address: kevin.ewans@shell.com (K. Ewans).

prior knowledge, the direction of the winds, waves, and currents can even be considered to be the same.

The simple combination of independent variables also glosses over the diverse climates that characterise the World's oceans. For example, the extreme meteorological phenomena in the Gulf of Mexico and the northwest coast of Australia are hurricanes. These are characterised by waves and currents that are driven by the local wind field, and there is a high probability of experiencing extreme winds, waves and currents together. In the Gulf of Guinea, extreme wave events are associated with swells from South Atlantic storms. The swells run normal to coast, while the currents from ocean circulation run along coast, and are independent from the swell. Accordingly, the probability of experiencing extreme waves and extreme currents is low, and the probability that the waves and currents are collinear is even smaller. In the Arabian Gulf, the wave extremes are due to the Shamal, whereas the currents are dominated by tides. As a result, and like the Gulf of Guinea case, the probability of experiencing extreme waves and extreme currents together is relatively low, but unlike the Gulf of Guinea, they are largely inline.

It is therefore clear that simple and relatively arbitrary combinations of independent criteria will result in joint criteria with an unknown probability, and further a given choice of combination will result in joint criteria with different probabilities for different oceans, when in fact the desired outcome are conditions that will give facilities designed to the same level of reliability. Accordingly, joint criteria with known probability of occurrence are required.

There is a considerable literature on the theory and modelling of multivariate extremes. The interested reader might consult Beirlant et al. (2004) for a general statistical introduction. Estimation of joint extremes is of interest to many environmental fields, particularly hydrology (see, e.g. Hawkes, 2008; Hawkes et al., 2002; Keef et al., 2013). The approaches used by the offshore industry to calculate joint extreme environmental conditions essentially fall into two camps — response-based and response-independent. The response-based approach relies on the specification of a response model giving load as a function of environment and permitting a back calculation of the environmental variables once an extreme load has been established. The response-independent or environmental approach involves developing joint criteria for the environmental variables alone associated with rare return periods. An outline of the response-based approach is given in Section 2. The main section of this paper is therefore Section 3, which gives a review of contemporary methods for calculating joint extreme environmental variables. Discussion and conclusions are presented in Sections 4 and 5 respectively.

2. Response-based methods

Response-based methods involve calculating a key response or several key responses via a response function in which the variables are environmental variables. For example, Tromans and Vanderschuren (1995) describe generic response functions for the mud-line base shear and over-turning moment of steel jacket structures. Their response function is given in terms of a sum of terms involving variables of the winds, waves, and currents. The coefficients of the terms are determined by calibration of large number of conditions with a given wave kinematics and current profile on a one meter diameter vertical column.

With a given response function, a long-term data set of environmental variables can be converted into an equivalently long-term data set of responses, allowing an extreme value analysis of the response variable to be undertaken. Estimates of the extremes of the response variable can be made to a given annual probability of exceedance or return-period, and this value can be used in the original response function to back-calculate the environmental variables, to establish an appropriate design set of environmental variables for detailed engineering design. It should be noted that the response variable calculated from the response function need not be an actual engineering response or load, but it must have the same statistical behaviour with the environmental variables as an actual engineering response or load.

The back-calculation of the environmental variables from the response function is not trivial. In its simplest form the back-calculation involves establishing an optimum combination of environmental variables, based on relationships established from the data and assumptions that these relationships will also apply in the extreme. Usually, one of the variables, such as the wave height in the case of the steel jacket response functions, is assumed to be dominant and the value of this variable is set at the return-period of interest. The other variables are then determined from their respective relationships for this value of the dominant variable. The optimum set of variables when substituted in the response function would give the extreme value of the response variable. The optimum choice of variables can also be determined by extending a response-independent method with the addition of the response variable. Distributions of the values of the environmental variables conditional on an extreme response variable can then be established, and an appropriate choice such as the most probable of each variable can be made. A consequence of the response-based approach, and in particular to having the probability distribution for the response or load variable, is that it is possible in principle to calculate the reliability of a structure against failure due to the environmental loading. We assume that the structural strength or resistance, R, of a structure can be characterised by a probability density function f_R. For any given value x of structural resistance, the structure fails if the environmental load E exceeds x. Writing the probability $\Pr(E > x)$ as $\overline{F}_E(x)$, it follows that the probability p_F of structural failure is

$$p_F = \int \overline{F}_E(x) f_R(x) dx.$$

The estimation of reliability is central to the FORM and SORM methods (e.g. Winterstein and Engebretsen, 1998). For a set of environmental variables X, a safety margin function $g(X)$ is defined such that when $g(X) \leq 0$, the structure will fail, otherwise it is safe. In the case of the steel jacket structure discussed above, $g(X) = R - E$; i.e., the structure will fail when the environmental load is greater than the resistance. The probability of failure is then determined from

$$p_F = \int_{g(x) \leq 0} f(x) dx$$

where $f(x)$ is now the joint probability density function of the set of environmental variables. The integral is difficult to solve since both $f(x)$ and the integration boundary, $g(x) = 0$, the failure surface, are multidimensional and usually nonlinear. The problem is simplified by re-expressing the set $X(= \{X_1, X_2, \ldots\})$ as a set of (independent) conditional random variables $\left\{ \tilde{X}_1, \tilde{X}_2, \ldots \right\}$. For example, for appropriately ordered variables we can write

$$\tilde{X}_1 = X_1, \quad \text{then} \quad \tilde{X}_2 = X_2 \big| X_1 \quad \text{and} \quad \tilde{X}_3 = X_3 \big| X_1, X_2 \quad \text{and so on}$$

with cumulative distribution functions $F_{\tilde{X}_1}, F_{\tilde{X}_2}, \ldots$. These independent random variables can now be transformed in turn to standard normal random variables $U_1, U_2, \ldots$ via the probability integral transform

$$F_{\tilde{X}_j}(x) = \Phi(u_j) \quad \text{for } j = 1, 2, \ldots$$

where Φ is the cumulative distribution function of the standard normal distribution. The probability of failure is then evaluated using

$$p_F = \int_{g_U(u) \leq 0} \phi(u) du$$

for transformed failure surface g_U, where ϕ is the probability density function of a set of independent standard normal random variables.

Thus, the contours of constant probability density of the integrand are concentric circles (bivariate case) or hyper-spheres in higher dimensions.

To facilitate solution, the integration boundary $g_U(u) = 0$ is simplified by truncating its Taylor expansion about an as yet unknown point, u^*, to first order (FORM) or second order (SORM). u^* is the point that has the highest probability density on $g_U(u) = 0$, to minimise accuracy loss (the integrand function quickly diminishes away from the expansion point), and is referred to as the Most Probable Point (MPP).

The MPP is found by minimising $\|u\|$ for $g_U(u) = 0$, the minimum distance from origin to the failure surface. The minimum distance $\beta = \|u^*\|$ is called the reliability index, and the probability of failure is now simply $p_F = 1 - \phi(\beta)$. The value of u^* can be transformed back to a corresponding x^* (in terms of the original variables) to establish the failure design set.

3. Response-independent methods

Response-independent methods for establishing combinations of environmental variables for design require joint distributions that describe the behaviour of the variables when one or more is extreme to be established directly from the environmental variables themselves. A particular combination of variables with a given low probability of occurrence can then be specified. Reference to a response variable is not required but could be used to further optimise selection of variables. In this sense, response-based methods are different only in that they involve finding the most likely combination of environmental variables to produce a target response value.

In the case of FORM or SORM, the failure surface is the target and a failure probability is calculated, but conversely if the target is a failure probability, a design point can be calculated on an associated failure surface. Winterstein et al. (1993) demonstrate this approach, which they refer to as inverse FORM, to calculate probability contours of joint occurrences of environmental variables. The design point, u^*, is found by minimising $g_U(u)$ for $\|u\| = \beta$. The FORM and SORM failure surfaces are tangential to the contour at u^*, but for design the behaviour of the system can be checked to ensure the actual failure surface is outside the contour for that probability. Nerzic et al. (2007) use inverse FORM contours for a West Africa location.

An example of the application of inverse FORM is given in Fig. 1. The plot shows contours of equal probability density for significant wave height, H_S, and spectral period, T_P, following the joint probability model proposed by Haver and Nyhus (1986).

The inverse FORM approach requires us again to express the set of environmental variables in terms of a product of independent random variables. In the bivariate case, we might model the distribution of X_1 and $X_2|X_1$, if there is good physical justification for doing so. In the model of Haver and Nyhus (1986), H_S is modelled with a Weibull distribution and $T_P|H_S$ is modelled with a log-normal distribution. These model forms are motivated by good fitting performance to the body of a sample of data, but their validity for extremes is not known. In addition, inverse FORM is difficult to model beyond two variables, requiring a model for the probability density of a variable conditional on the occurrence of the others. In the case of three variables, the objective is to estimate $f(X_1, X_2, X_3)$ with

$$f(X_3, X_2, X_1) = f(X_3|X_2, X_1) f(X_2|X_1) f(X_1)$$

but the difficulty lies in justifying the choice $X_3|X_2$, X_1 on physical grounds, and then estimating $f(X_3|X_2, X_1)$ — which is not straightforward.

The motivation for application of asymptotic distributions in extreme value analysis is the remarkable concept of max-stability. It can be shown that (appropriately shifted and scaled versions of) maxima from a large class of ("max-stable") probability distributions have very similar statistical characteristics and the same distributional form. As
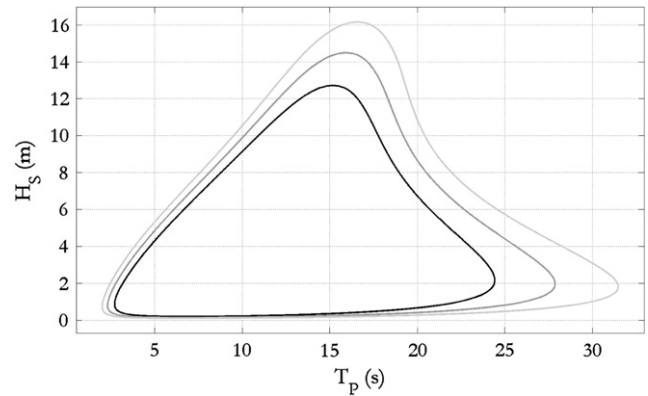

Fig. 1. Contours of equal probability density for significant wave height and spectral peak period, for annual exceedance probabilities corresponding to 10-year (black), 100-year (dark grey), and 1000-year (light grey) return periods.

a result, in the univariate case, there is justification for modelling extremes of block maximum data (e.g. monthly maxima) with a generalised extreme value distribution and peaks over threshold data with a generalised Pareto distribution. However, in the multi-dimensional case, max-stability is possible only when (often unrealistic) component-wise maxima assumptions are appropriate. Nevertheless, the max-stable concept has been used for spatial extremes, with implicit asymptotic dependence assumed (Jonathan and Ewans, 2013).

The conditional extremes model of Heffernan and Tawn (2004) provides a more general framework based on a (more realistic) limit assumption. It involves modelling the conditional distribution of one variable when the value of the conditioning variate is large, but a distinct advantage over a typical FORM analysis is that no prior knowledge of the forms of the distributions is required. Instead, asymptotic distributional forms are used.

The method is most clearly and most easily described in the case of two variables (X,Y) but can be trivially extended to multi-dimensions. The marginal distribution of each variable is expressed on a Gumbel scale, by modelling variables in turn using a generalised Pareto distribution (assuming threshold exceedances) and then transforming using the probability integral transform. A parametric form then applies for the conditional distribution of one variable given a large value of the other variable

$$(Y|X = x) = \alpha x + x^\beta Z \text{ for } x > u$$

for an appropriate threshold u, where $\alpha \in (0,1]$ is the scale parameter, $\beta \in (-\infty, 1]$ is the shape parameter, and Z is a random variable, independent of X, converging with increasing x to a non-degenerate limiting distribution, G (which is assumed Gaussian for model fitting purposes only). Threshold u is selected by inspecting a number of model fit diagnostics (see, e.g. Jonathan et al., 2012); the smallest value of threshold u for which acceptable model fit diagnostics are observed, admitting the largest possible sample for model fitting, is generally used. In application, using sample $\{x_i, y_i\}_{i=1}^n$ of n values of X and Y respectively, both on Gumbel scale, the residuals

$$\hat{z}_i = \frac{y_i - \hat{\alpha} x_i}{x_i^{\hat{\beta}}} \text{ for } i = 1, 2, \ldots$$

are assumed to provide a sample from the distribution of Z, where $\hat{\alpha}$ and $\hat{\beta}$ are the fitted values of α and β respectively. Then estimates of conditional extremes of Y given X are obtained by simulation by

- Drawing a threshold exceedance value x of X randomly from its standard Gumbel distribution,
- Drawing a value z of Z randomly from the set of estimated values of $\hat{z}$,

- Calculating $(y|x) = \hat{\alpha}x + x^{\hat{\beta}}z$, and finally
- Transforming the pair (x,y) from Gumbel to original physical scale using the probability integral transform.

By way of example, an application of the Heffernan and Tawn method to wave data from several locations is given in Jonathan et al. (2010). Fig. 2 is a plot of measured storm peak significant wave height and associated spectral peak period from measurements in the northern North Sea, together with estimates from simulations for $H_S > 15$ m from conditional extremes modelling. The plot shows a generally increasing trend in T_P with H_S. The most probable value of T_P, which appears to be between 16 s and 17 s for $H_S > 15$ m, is significantly less than the longest in the measured data. Jonathan et al. (2010) demonstrates the improved performance of the conditional extremes model with respect to the model of Haver and Nyhus (1986) for simulated samples from known multivariate distributions. The article also suggests how the model of Haver and Nyhus (1986) might be improved to provide estimates with improved statistical characteristics.

An example of an application of the Heffernan and Tawn method to multi-dimensional problems is given by Jonathan et al. (2012), in which current profiles measured on the northwest shelf of Australia were analysed to derive extreme profiles with depth. Two and half years of current measurements, including both speed and direction, were made at eight depths through the water and were available for the analysis. The steps in the analysis involve

- Resolving currents into major and minor axes of total current at each depth,
- For each axis, separating tidal and residual components by a local harmonic analysis,
- Calculating hourly maxima for each of the tidal and residual components, with the residual maxima to be used for the extreme value analysis, and the tidal maxima to be used for recombining with the residual simulations from conditional extremes modelling,
- Applying the conditional extremes model to the residual hourly extremes
  – fitting marginals with a generalised Pareto distribution,
  – transforming to Gumbel marginal scale,
  – fitting a multi-dimensional conditional extremes model (for all residual components) of the form

$$(\mathbf{Y}_{[-k]}|Y_k = y_k) = \alpha_k y_k + y_k^{\beta_k}\mathbf{Z}_k \text{ for } k = 1, 2, \ldots$$

where $\mathbf{Y}_{[-k]}(=\{Y_1, Y_2, \ldots, Y_j, \ldots\}_{j \neq k})$ is the set random variables, excluding the conditioning variate, on Gumbel scale. y_k is the
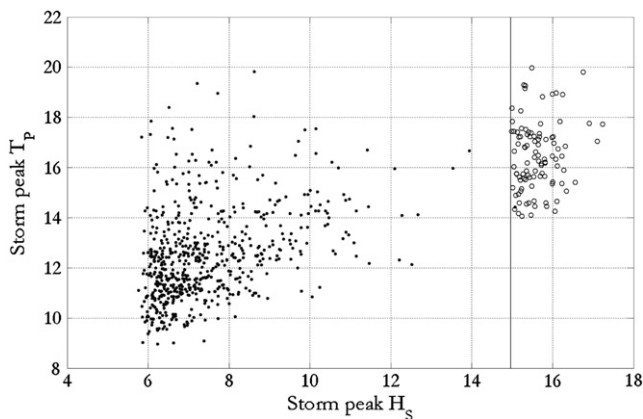
value of the conditioning variable Y_k on Gumbel scale, and α_k and β_k are vectors of parameters to be estimated as before. $\mathbf{Z}_k(=\{Z_{k1}, Z_{k2}, \ldots, Z_{kj}, \ldots\}_{j \neq k})$ is a vector random variable whose jth component describes the "residual variation" in the model for $Y_j|Y_k = y_k$, $j \neq k$. Componentwise multiplication is assumed (see Jonathan et al., 2012 for details).

- Simulating samples of joint extremes, where
  – tidal components are re-sampled with replacement, and
  – sampled tidal components and residuals are added to provide hourly estimates of hourly maxima and minima along the major and minor axes.

An example of the results is given in Fig. 3, which shows median maxima hourly extremes conditioned on exceedances of the 10-year return period current values at depth D1. The figure suggests that the minor axis conditional extremes are approximately symmetric about zero at depths D1 to D3. At depth D4 however, there is systematic rotation of current components in a clockwise direction, with respect to axis directions defined using unconditioned sample at this depth. At depths D5 to D8, this trend is reversed; rotation is in an anti-clockwise direction.

The importance of accounting for covariates in univariate extreme value analyses has been demonstrated by Jonathan et al. (2008). The conditional extremes model can also be extended to include covariates in a relatively straightforward manner. The objective becomes to model the distribution of T_P (say) when H_S is extreme, as a function of storm direction θ as covariate, for which the conditional extremes model form becomes

$$(T_P|H_S = h, \theta) = \alpha_\theta h + h^{\beta_\theta}(\mu_\theta + \sigma_\theta Z)$$

where H_S and T_P are assumed on Gumbel scale and h is the conditioning value of H_S. $\alpha_\theta, \beta_\theta, \mu_\theta$ and σ_θ are now all smooth functions of θ to be estimated, and tZ is the "residual" random variable as before.

As an example of its application we give the results for hindcast storm peak H_S and associated T_P in the northern North Sea reported by Jonathan et al. (submitted for publication, draft). The objective is to model the distribution of T_P for large storm peak H_S as a function of storm direction. The location is particularly useful for application
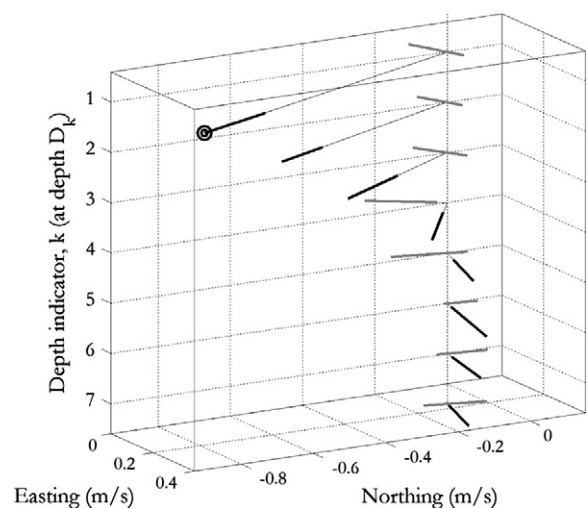
Fig. 3. Illustration of median conditional total current components with depth ($k = 1,2, \ldots,8$) for exceedances of 10-year return level of the major axis component at depth *D1* from conditional simulation of hourly extremes. Uncertainty intervals for major (thick black) and minor components (thick grey) at each depth are projected onto the positive major- and minor- axis at each depth k. The figure suggests systematic rotation of currents clockwise with respect to unconditioned currents at depth D4, but anti-clockwise rotation at larger depths. See Jonathan et al. (2012) for details.

Fig. 2. Measured storm peak significant wave height and associated spectral peak period from the northern North Sea (filled dots), and simulations of H_S and conditional T_P values for storm peak $H_S > 15$ m following conditional extremes modelling.

of the model as the wave field has identifiable characteristics for various directional sectors, as can be seen in Figs. 4 and 5. Storms with the largest sea states are those occurring in the north, south, and southwest–west sectors; less severe sea states are associated with storms from the northwest sector; and virtually no storms occur that cause waves from the easterly sector. Further, it can be seen in Fig. 5 that storm peak sea states from the northwest and southwest–west sectors are associated with the longest T_P values. These characteristics should be evident in the conditional extremes modelling and can serve as an indicator of the success of the modelling. The joint distribution of H_S and T_P below the threshold is modelled using quantile regression.

Conditional T_P values corresponding to storm peak H_S values with exceedance probability of 0.01 are illustrated in Fig. 6. The inner (black and white) dotted curves, drawn on the same scale, illustrate estimates of the storm peak H_S return value. The inner white dotted curve is an estimate for the directional variation of the storm peak H_S return value. For comparison the inner black dotted curve is an estimate for the same return value ignoring directional effects. The influence of longer fetches from south (in particular), the Atlantic and Norwegian Sea are visible.

The outer (black and white) curves, drawn on the same scale, illustrate estimates for return values of T_P conditional on exceedances of the corresponding storm peak H_S value. Solid lines represent median values, and dashed lines 95% uncertainty bands, incorporating (white) or ignoring (black) directional effects. The results clearly show increased associated periods from the Atlantic and Norwegian sectors, as expected. When directionality is ignored, associated T_P values are underestimated for some sectors and overestimated for others. The importance of this difference for design can be seen in the response of a simple system with a transfer function characteristic of the roll or heave of a floating system with a natural period of around 17 s. Response is over-estimated by more than 30% in directional sectors with short fetches, but under-estimated by as much as 20% in sectors with long fetches, particularly the Atlantic sector (see Jonathan et al., submitted for publication, draft).

4. Discussion

Metocean engineers are often tasked with estimating joint extremes. This task is often framed as the estimation of associated values of a set of environmental variables given the occurrence of large values of a dominant environmental or structural response variable. Typical examples might be the specification of associated peak period corresponding to a significant wave height with given return period, or the specification of significant wave height and spectral peak period corresponding to a large value of a particular structural response. The task is therefore to model joint extremes of a sample of values, assumed drawn from some multivariate distribution. The practitioner needs to estimate the tail behaviour of the multivariate distribution from which the sample is drawn, based on the sample alone, and might adopt one of at least three generic approaches to complete the task well.

Firstly, the practitioner might choose to fit one or more of a large number of parametric models for multivariate extremes. These models, for the joint tail of a multivariate distribution only (rather than the whole distribution) are well understood (see, e.g. Tawn, 1988 or Kotz and Nadarajah, 2000). Different model forms have different tail characteristics, different extremal dependence structures (see, e.g. Ledford and Tawn, 1997), leading to different estimates for (joint) design values from the same sample. Given sufficient sample, the most appropriate parametric model can be selected using statistical tests. A number of simple diagnostic tools are available to characterise extremal structure, as summarised e.g. by Eastoe et al. (2013). However, in practice due to limited sample size, it is often difficult if not impossible to select appropriate model forms based on quality of fit alone, leading to considerable ambiguity in design values.

Secondly, the practitioner might choose to model the dependence structure of the extreme values using an extremal dependence model motivated by asymptotic statistical arguments. This approach involves first transforming extreme values to standard marginal scale using an appropriate univariate extreme value distribution (e.g. generalised
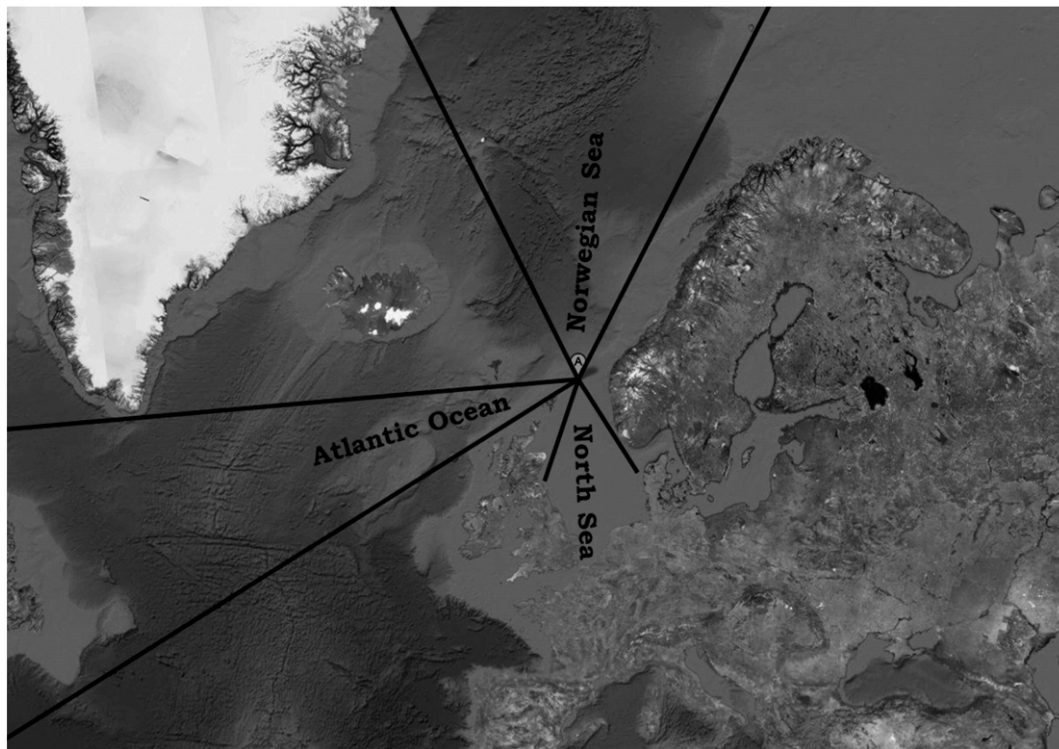

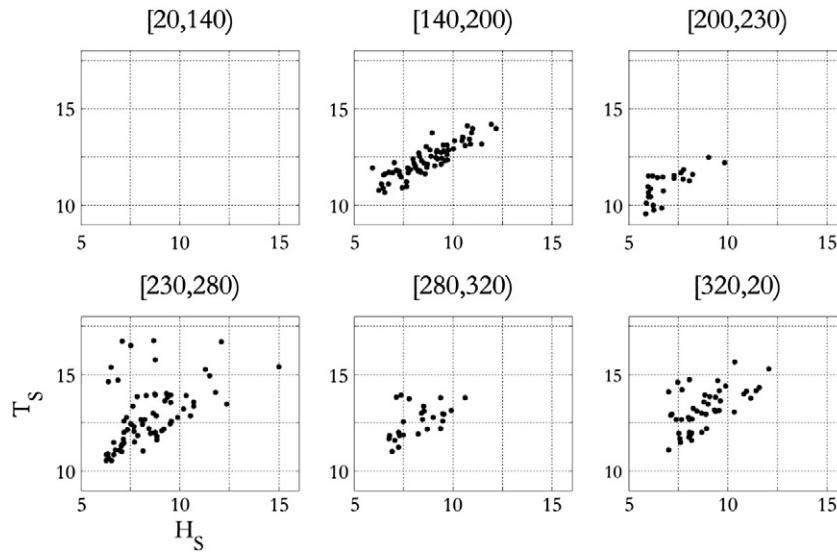**Fig. 4.** Northern North Sea location and directional sectors with distinctive wave characteristics.

Fig. 5. Scatter plots of associated T_P against storm peak H_S for different directions of arrival.

Pareto for threshold exceedances, or generalised extreme value for sample maxima) and the probability integral transform (see Section 3). With the sample now on common marginal scale, the modeller can fit one or more appropriate dependence models, including the conditional extremes model and extreme value copulas. The critical point here is that the dependence models adopted must be appropriate to model multivariate extremes. Different copula models impose different extremal dependence structures on the sample, and may not be appropriate.

Only some (so-called max-stable) copula models are appropriate for extreme value modelling of spatial extremes, for example. Similarly, FORM and in particular inverse FORM offers the possibility for more realistic estimation of joint extremes and has been used frequently by the offshore industry since the early 90s, but FORM generally relies on the estimation of conditional distributions developed from the body of the distribution rather than the tail, is difficult to extend to multi-dimensions, and imposes a particular (possible unjustified) extremal dependence

Fig. 6. Return values of storm peak H_S (inner circle) and associated conditional values of T_P (outer circle). Inner dashed lines (on common scale): storm peak H_S with non-exceedance probability 0.99 (in 34 years), with (white) and without (black) directional effects. Outer solid lines (on common scale): median associated T_P with (white) and without (black) directional effects; outer dashed lines give corresponding 2.5% and 97.5% percentile values for associated T_P. Design values are withheld for reasons of confidentiality.

structure. The major advantage of the conditional extremes approach is that it allows the user to estimate the nature of the extremal dependence present in the sample directly from the sample, and reflect any modelling uncertainty in estimated design values in a natural way using simulation.

Thirdly, if the practitioner is interested in estimating spatial extremes, such as the values of significant wave height at each point in a neighbourhood, methods of spatial extreme value analysis may be used. Methods based on max-stable processes, motivated by the work of Smith (1990), are increasingly popular in the statistics literature. Despite the fact that the full multivariate probability density function cannot be written in closed form, composite likelihood methods provide one approach to inference as illustrated e.g. by Padoan et al. (2010), Davison and Gholamrezaee (2012) and Davison et al. (2012). Censored likelihood methods provide an approach to making these models, which assume componentwise maxima, available for analysis of threshold exceedances (e.g. Huser and Davison, 2012, draft). Furthermore, Wadsworth and Tawn (2012) propose the adoption of inverted multivariate extreme value distributions with which to admit hybrid extremal dependence structures within the framework of max-stable processes. However, these methods are technically demanding, and are generally only used in the statistics community. Computationally, they do not scale well in general with increasing numbers of locations. Some applications to extreme precipitation have been reported (see, e.g., Cooley et al., 2007). Much work is needed before spatial extremes methods can be used reliably in real-world engineering applications. In contrast, the conditional extremes model provides a pragmatic alternative for spatial analysis also.

Covariate effects are important in extreme value analysis, for individual variables or for joint modelling. Incorporating covariate effects in multivariate extreme value models is challenging in general. However, as demonstrated in this paper, incorporation of covariate effects in the conditional extremes model is possible. Moreover, for the northern North Sea application illustrated, estimated models including covariate effects are different from those excluding covariates, reflect physical reality more adequately, and lead to different estimates for return values.

5. Conclusion

Estimation of joint occurrences of extremes of environmental variables is crucial for design of offshore facilities and achieving consistent levels of reliability. Specification of joint extremes in design criteria has often been somewhat ad hoc, being based on fairly arbitrary combination of extremes of variables estimated independently. Such approaches are even outlined in design guidelines. More rigorous methods for modelling joint occurrences of extremes of environmental variables are now available. In particular, the conditional extremes model provides a straightforward approach to joint modelling of extreme values, based on solid theory. It admits different forms of extremal dependence, ensuring that the data (rather than unwittingly made modelling assumptions) drive the estimation of design values. The model admits uni- and multi-variate covariate effects, is scalable to high dimensions and allows uncertainty analysis via simulation. For this reason, we recommend the conditional extremes model for joint extremes modelling in both response-based and response-independent metocean design.

References

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., 2004. Statistics of Extremes: Theory and Applications. Wiley.

Cooley, D., Nychka, D., Naveau, P., 2007. Bayesian spatial modelling of extreme precipitation return levels. J. Am. Stat. Assoc. 102:0, 824–840.

Davison, A.C., Gholamrezaee, M.M., 2012. Geostatistics of extremes. Proc. R. Soc. A 468:0, 581–608.

Davison, A.C., Padoan, S.A., Ribatet, M., 2012. Statistical modelling of spatial extremes. Stat. Sci. 27:0, 161–186.

Eastoe, E., Koukoulas, S., Jonathan, P., 2013. Statistical measures of extremal dependence illustrated using measured sea surface elevations from a neighbourhood of coastal locations. Ocean Eng. 62, 68–77.

Haver, S., Nyhus, K.A., 1986. A wave climate description for long term response calculations. Proc. 5th OMAE Symp., IV:0, pp. 27–34.

Hawkes, P.J., 2008. Joint modelling of environmental parameters for extreme sea states incorporating covariate effects. J. Hydraul. Res. 46:0, 246–256.

Hawkes, P.J., Gouldby, B.P., Tawn, J.A., Owen, M.W., 2002. The joint probability of waves and water levels in coastal defence design. J. Hydraul. Res. 40:0, 241–251.

Heffernan, J.E., Tawn, J.A., 2004. A conditional approach for multivariate extreme values. J. R. Stat. Soc. B 66:0, 497.

Huser, R., Davison, A.C., 2012. Space-time modelling of extreme events. Draft at arxiv.org/abs/1201.3245 (draft).

Jonathan, P., Ewans, K., 2013. Statistical modelling of extreme ocean environments with implications for marine design: a review. Ocean Eng. 62:0, 91–109.

Jonathan, P., Ewans, K.C., Forristall, G.Z., 2008. Statistical estimation of extreme ocean environments: The requirement for modelling directionality and other covariate effects. Ocean Eng. 35:0, 1211–1225.

Jonathan, P., Flynn, J., Ewans, K.C., 2010. Joint modelling of wave spectral parameters for extreme sea states. Ocean Eng. 37:0, 1070–1080.

Jonathan, P., Ewans, K.C., Flynn, J., 2012. Joint modelling of vertical profiles of large ocean currents. Ocean Eng. 42:0, 195–204.

Jonathan, P., Ewans, K.C., Randell, D., 2013. Joint modelling of environmental parameters for extreme sea states incorporating covariate effects. Coastal Engineering, draft at www.lancs.ac.uk/~jonathan (submitted publication).

Keef, C., Papastathopoulos, I., Tawn, J.A., 2013. Estimation of the conditional distribution of a vector variable given that one of its components is large: additional constraints for the Heffernan and Tawn model. J. Multivar. Anal. 17:0, 396–404.

Kotz, S., Nadarajah, S., 2000. Extreme Value Distributions: Theory and Applications. Imperial College Press, London, UK.

Ledford, A.W., Tawn, J.A., 1997. Modelling dependence within joint tail regions. J. R. Stat. Soc. B 59:0, 475–499.

Nerzic, R., Frelin, C., Prevosto, M., Quiniou-Ramus, V., 2007. Joint distributions of wind, waves and current in West Africa and derivation of multivariate extreme i-form contours. Proc. 17th International Offshore and Polar Engineering Conference, Lisbon, Portugal, July 1–6, 2007.

Padoan, S.A., Ribatet, M., Sisson, S.A., 2010. Likelihood-based inference for max-stable processes. J. Am. Stat. Assoc. 105:0, 263–277.

Smith, R.L., 1990. Max-stable processes and spatial extremes. Unpublished, available from http://www.stat.unc.edu/postscript/rs/spatex.pdf.

Tawn, J.A., 1988. Modelling multivariate extreme value distributions. Biometrika 77:0, 245–253.

Tromans, P.S., Vanderschuren, L, 1995. Risk based design conditions in the North Sea: application of a new method. Offshore Technology Conference, Houston (OTC–7683).

Wadsworth, J.L., Tawn, J.A., 2012. Dependence modelling for spatial extremes. Biometrika 99:0, 253–272.

Winterstein, S., Engebretsen, K., 1998. Reliability-based prediction of design loads and responses for floating ocean structures. Proc. 27th International Conf. on Offshore Mechanics and Arctic Engineering, Lisbon, Portugal.

Winterstein, S.R., Ude, T.C., Cornell, C.A., Bjerager, P., Haver, S., 1993. Environmental parameters for extreme response: inverse FORM with omission factors. Proc. 6th Int. Conf. on Structural Safety and Reliability, Innsbruck, Austria.