Ocean Engineering 62 (2013) 91–109

Contents lists available at SciVerse ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

Review

Statistical modelling of extreme ocean environments for marine design: A review

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

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

ARTICLE INFO

Article history:
Received 8 August 2012
Accepted 5 January 2013

Keywords:
Review
Extreme
Covariate
Multivariate
Bayesian
Return value

ABSTRACT

We review aspects of extreme value modelling relevant to characterisation of ocean environments and the design of marine structures, summarising basic concepts, modelling with covariates and multivariate modelling (including conditional and spatial extremes). We outline Bayesian inference for extremes and reference software resources for extreme value modelling. Extreme value analysis is inherently different to other empirical modelling, in that estimating the tail (rather than the body) of a distribution from a sample of data, and extrapolation beyond the sample (rather than interpolation within) is demanded. Intuition accumulated from other areas of empirical modelling can be misleading. Careful consideration of the effects of sample size, measurement scale, threshold selection and serial dependence, associated uncertainties and implications of choices made is essential. Incorporation of covariate effects when necessary improves inference. Suitable tools (e.g. based on additive models, splines, random fields, spatial processes) have been developed, but their use is restricted in general to academia. Effective modelling of multivariate extremes will improve the specification of design conditions for systems whose response cannot be easily characterised in terms of one variable. Approaches such as the conditional extremes model are easily implemented, and provide generalisations of existing marine design approaches (e.g. for primary and associated variables). Software is available, but again generally only for academic use. Modelling spatial dependence rigourously will provide single extreme value models applicable to spatial neighbourhoods including complete ocean basins, avoiding the need for procedures such as site pooling. Indeed, once the model is established, the metocean engineer may not ever need to perform further extreme value analysis for that basin in principle. Spatial extremes is an area of active research in the statistics community. A limited number of appropriate models have been deployed (e.g. for precipitation, temperature and metocean applications). Software is available, but again for specialist use. Bayesian inference provides a consistent framework for inference and is rapidly becoming the standard approach in academia. It appears inevitable that, in time, Bayesian inference will also be regarded as the standard in ocean engineering applications. Implementation of Bayesian methods requires some expertise. Software is available, but again generally only used by statistical specialists.

* Corresponding author. Tel.: +44 151 272 5421; fax: +44 151 44 5384.
E-mail addresses: philip.jonathan@shell.com (P. Jonathan), kevin.ewans@shell.com (K. Ewans).

0029-8018/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.oceaneng.2013.01.004

Contents

1. Introduction

Maritime structures from breakwaters to oil and gas producing facilities in thousands of metres of water must be designed for extreme environmental conditions. 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. To define the environmental loading, metocean criteria therefore need to be specified to an appropriate return period, typically 100 years, but sometimes to 10,000 years for the required failure probabilities. Extreme value (henceforth EV) analysis of data from measurements and/or hindcasts are undertaken to derive these criteria, but the resulting uncertainties associated with long return period criteria are usually large, as the number of data available for analysis is usually small by comparison. In addition, there is no standard approach to analysis within the metocean community, and it remains a subject of continuous debate and active research.

For most ocean basins in the world, adequate historical meteorological records exist to allow numerical simulation of hindcast data, some for periods as long as 100 years. For example, in the Gulf of Mexico, the proprietary GOMOS hindcast data set covers the period 1900–2008. The latest version, GOMOS08, includes a 29-year continuous wind and wave hindcast for the period 1980–2008, 379 tropical events (hurricanes) for the period 1900–2008, and 68 extra-tropical events (winter storms) for the period 1950–2008 (see Oceanweather, 2005). Such hindcasts include the relevant parameters of the wind and sea state and provide a good basis upon which analysis can be undertaken, including joint extremal analysis. Many physical systems respond to environmental conditions in a manner that cannot be represented by a single variable or parameter. For example, the pitch of a vessel is as much a function of the wave period or wave length as it is of the wave height. Accordingly, the extreme conditions must be specified jointly, necessitating a multivariate analysis involving a number of variables. A hindcast study often also incorporates modelling of currents at the surface and through the water column providing the means to develop joint design criteria. Another important feature of a typical hindcast for an ocean basin is that it provides a spatial, directional and temporal description of extreme events, and in particular, it allows for incorporation of location as a covariate and modelling of spatial dependence, avoiding the need for (e.g.) site pooling. Clustering of large values in time series similarly demands careful modelling.

The requirement to consider the effect of covariates in developing extreme criteria has been well demonstrated. For example, extremes of storm peak significant wave height are dependent on time of year or the directionality of the sea state concerned. Although these effects have long been known (e.g. Graham, 1981), and rigourous techniques for dealing with covariates in estimates have also been available for some time, it is only recently that such methods have been adopted for establishing metocean design criteria. A good case in point is directionality. Graham (1981) notes inconsistencies associated with estimating directional wave extremes with the omni-directional extreme, and much debate has ensued. Forristall (2004) notes the importance of ensuring that directional criteria are consistent with omnidirectional criteria in the sense that they yield consistent failure probabilities in design. In particular, he points out that directional criteria that are simply scaled to omnidirectional criteria estimated independently give inconsistent probabilities. Moreover, the distribution of extreme significant wave heights associated with (e.g.) monsoonal surges would not be expected to be the same as those for extreme significant wave heights from (e.g.) typhoons, even though return values corresponding to some return periods might be similar, since they emerge from different physical phenomena. Covariate models can be useful in this case also.

There have been considerable recent advances in EV theory and method within the statistics community, and a developing literature on EV modelling applied to many scientific, engineering and financial disciplines. A considerable proportion of the literature on applications of extreme value analysis to weather and climate change is relevant to the metocean engineer, such as methods used to predict extreme precipitation (e.g. Cooley et al., 2007). Chandler (2005) proposes addition of generalised linear models, able to model spatio-temporal systems, to the climatologist's toolkit. The relative merits of single location versus regional estimates for extremes values are also a concern (e.g. Kysely et al., 2011), as are methods (such as regional frequency models) to provide best single site estimates. Detection of climate change effects and their incorporation within return level estimation is another area of common ground (e.g. Bernier et al., 2007; Kyselý et al., 2010). Statistical downscaling (e.g. Vrac and Naveau, 2007; Michelangeli et al., 2009) provides a framework within which the (extreme) characteristics of local short-term (e.g. hindcast-based) data can be combined reliably with large-scale long-term (e.g. climate simulation based) data to provide estimates of return levels far into the future. Some of these developments are sufficiently mature to be useful for marine design. Yet the number of instances of applications of modern EV methods to metocean design is rather small. It is hoped that this review will encourage more ocean engineers to consider using modern methods.

There are cornerstone texts in extreme value modelling for environmental applications that the interested reader might find useful, including Coles (2001), Kotz and Nadarajah (2000), Davison (2003) and Beirlant et al. (2004).

The outline of this review is as follows. Section 2 gives an introduction to modelling peaks over threshold. Section 3 suggests practical considerations relevant to modelling peaks over threshold. Motivated by Section 3, we then review models incorporating covariate effects for peaks over threshold in Section 4. Methodologies of multivariate extreme value analysis of peaks over threshold and block maxima are considered in Section 5. Section 6 motivates Bayesian inference for extremes. The review is illustrated using examples from previously-published work by the authors. In Section 7, we draw conclusions and make recommendations for future applications and developments of EV analysis within ocean engineering. Two appendices provide an informal introduction to relevant aspects of the mathematics of univariate and multivariate extremes (Appendices A and B respectively).

2. Modelling peaks over threshold

We choose to motivate extreme value modelling from the perspective of peaks over threshold. Pickands (1975) showed that the conditional distribution function $F_u(x)$ of a random variable X, with continuous cumulative distribution function F, given exceedences $X>u$ of a threshold u can be approximated by a generalised Pareto (GP) distribution (see Appendix A.5) for sufficiently high threshold. For a peak X over threshold u, the form of the generalised Pareto distribution with shape parameter ξ and scale parameter σ is

$$F_u(x) = \frac{F(x)-F(u)}{1-F(u)} = \Pr(X \le x | X > u)$$
$$= 1-\left(1+\frac{\xi}{\sigma}(x-u)\right)^{-1/\xi} \quad \text{for } \xi \neq 0 \tag{1}$$

for $u \in (-\infty,\infty)$, $\sigma \in (0,\infty)$ and $\xi \in (-\infty,\infty)$ with $x > u$, $1+\xi(x-u)/\sigma > 0$, with the right hand side taken to be $1-\exp(-(x-u)/\sigma)$ when $\xi = 0$. The form of the unconditional distribution of peaks over threshold is then

$$F(x) = \Pr(X \le x) = (1-p_u) + p_u F_u(x)$$

where p_u is the probability of threshold exceedence.

Fitting samples of peaks over threshold using the GP distribution is defensible if the threshold used is large enough. Using diagnostic tests, such as the mean residual life plot (e.g. Embrechts et al., 2003), comparing the variability of parameter estimates with expected behaviour as threshold is varied and confirming the stability of extreme quantile estimates with threshold, the plausibility of the GP model form can be ascertained (e.g. Coles, 2001). However, there is no direct evidence that the chosen threshold is sufficiently high that the GP distribution is an adequate model. In general, it would be advantageous to have sufficient understanding of the underlying physics to suggest an appropriate model selection, so that as much of the sample as possible (not just the largest values) could contribute to parameter estimation.

There is a close correspondence between modelling of peaks over threshold (or threshold exceedences) and so called block maxima such as daily or monthly extremes. The emphasis in this review is on modelling of threshold exceedences, which we favour for reasons of statistical efficiency, since more data are in principle available with which to make inferences. Appendix A provides a motivation for the block maximum approach leading to the generalised extreme value (GEV) distribution. Appendix A also demonstrates the duality between peaks over threshold and block maxima approaches.

There are other general characteristics of modelling of peaks over threshold which are worthy of note.

2.1. The value of the shape parameter

The shape parameter ξ determines tail behaviour. When ξ is negative, the support of extreme values is bounded on the right hand side. Otherwise, the distribution is unbounded on the right. For modelling of storm peaks of significant wave height H_S, numerous authors (e.g. Elsinghorst et al., 1998; Jonathan and Ewans, 2007a, 2008) have reported estimates for EV shape ξ that were found to be generally more often negative than positive (for arbitrary selections of covariate values). $\xi < 0$ implies a finite upper bound for storm peak H_S, considered reasonable because of physical constraints on storm size (such as water depth and fetch limitations, or physical limitations on pressure fields, winds and storm intensity). It should be noted that a finite upper bound on storm peak H_S does not imply a finite upper bound for individual waves within a storm, although it seems reasonable to expect an upper bound on individual waves will also exist.

2.2. Alternative distributional assumptions

Many practitioners in engineering disciplines adopt specific model forms for describing exceedances x of threshold u, e.g. fitting the conditional distribution $F_u(x)$ (see Eq. (1)) with $F(x)$ given by the classical Weibull distribution with cumulative distribution function:

$$F(x) = 1-\exp\left(-\left(\frac{x-\mu}{\beta}\right)^{\alpha}\right)$$

for $\mu \in (-\infty,\infty)$, $\alpha,\beta > 0$ with $x > \mu$, or the Gumbel distribution with cumulative distribution function:

$$F(x) = \exp\left(-\exp\left(-\left(\frac{x-\mu}{\beta}\right)\right)\right)$$

for $\mu \in (-\infty,\infty)$ and $\beta > 0$, although the GP form for $F_u(x)$ is recommended by some on the basis of goodness of fit alone (e.g. Li et al., 2012) regardless of theoretical considerations. There is often little evidence in the data to distinguish between quality of fit for different forms. Yet fitting using the classical Weibull distribution, for example, implicitly constrains the solution to be unbounded on the right hand side. Modelling using GP avoids this constraint. For this reason, fitting using classical Weibull and GP is expected to give similar estimates for return values corresponding to return periods of the order of that of the sample, but different estimates for return values corresponding to long return periods.

2.3. Tail estimation from a sample

EV analysis involves the most unusual events in the sample, rather than the typical. Uncertainties of extreme values (from measurement or hindcast) are likely to be different to those of the bulk of the data. The few largest values in the sample are typically highly influential for estimates (Davison and Smith, 1990); the model is most sensitive to the most informative observations. Compounded with limited sample size, selection of the optimal model form that characterises the observed extreme values is therefore difficult. Minor differences in model form or parameter estimates, which would be of little consequence for interpolation within the domain of the data, can produce material discrepancies in return values, especially at long return periods. The quality of interference from EV analysis of data from hindcast models depends critically on the extent to which extremes from hindcasts are physically realistic (Coles and Simiu, 2003).

2.4. Return values

One aim of EV modelling is the estimation of an extreme quantile corresponding to a particular return period, P, or the maximum value observed in P years. Typically, a single value is reported, such as the value of H_S which is exceeded with probability $1/P$ per annum. Yet the P-year maximum is itself a random variable. For most random variables of environmental interest, such as (significant) wave height, given perfect model specification and parameter estimation, the P-year maximum event will follow a skewed distribution with long right-hand tail. Model and parameter uncertainty broaden this distribution. Typically, the value reported is near the mode of this distribution, which because of skewness, is less than both the expected and median values. Further, the ratio of the 95th percentile of the distribution of the P-year maximum to its mode also varies systematically with the shape parameter of the GP used. This ratio is greater in Gulf of Mexico conditions (ξ near -0.1) than in the northern North Sea conditions ($\xi = -0.3$). That is, for the same value of most probable 100-year H_S, we expect more exceptionally large values in the Gulf than in the Northern North Sea (Jonathan and Ewans, 2007b). Expressions for return values from GP and GEV models are given in Appendices A.5 and A.3 respectively.

2.5. Period of data and return period

The period of data available for EV estimation is usually shorter than the return period of interest. Therefore the analysis entails extrapolation beyond the domain of measurements. Even after assuming homogeneity and regularity, uncertainties associated with estimates will be high compared with the usual and preferred case in statistical modelling corresponding to interpolation between data within the domain of the sample (Anderson, 1990). Extreme events with low rates of occurrence have an important impact on long return values. If 1000 years of data were available with which to estimate the distribution of 100-year maximum at a given location over the past (approximately stationary) millennium, progress would be relatively straight forward, even in the presence of directional, seasonal and possible long term variation with time. This regrettably is not usually the case in offshore design.

3. Practicalities

The statistical aspect of developing marine design conditions for a given metocean environment requires being mindful of multiple (often competing) considerations. Some of these are summarised here, as a means of introduction to practical analysis, and a motivation for methods presented in subsequent sections.

3.1. Are observations independent in time?

Observations (e.g. of time series of extreme H_S) often show serial dependence in time. Since EV models generally assume independent observations, the sample must be de-clustered prior to analysis. De-clustering can be performed based on physical considerations (e.g. selecting storm peak events to represent an interval of time series corresponding to each storm) or statistical considerations (e.g. Smith and Weissman, 1994; Ferro and Segers, 2003; Ledford and Tawn, 2003; see Appendix A.7 for an introduction to the extremal index). Both approaches select a subset of the original observations as effectively independent data for modelling.

Asymptotic theory (Leadbetter, 1991) states that the distribution of the (original) exceedances and storm peak or cluster maxima are the same, but recent work (Fawcett and Walshaw, 2007) suggests a bias in estimating the distribution of all threshold exceedances using cluster maxima only. Assuming independent extremes, when in fact they are clustered, has implications for inference. In particular, standard errors of parameter estimates are under-estimated. It can be seen that the implications of summarising storms using storm peak events need to be considered carefully, especially when covariate effects are present. For example, a storm event typically spans a range of storm directions, yet the storm peak event is associated with just one direction. In retaining only the storm peak event for modelling, it is essential also to retain the directional influence of that storm for estimation of directional design conditions (using ideas such as directional dissipation, see Jonathan and Ewans, 2007a).

3.2. Are observations from multiple locations?

When attempting to estimate an EV model for a single location using observations of the same set of storm events at a neighbourhood (or pool) of locations (typically centred on the location of interest), observations per storm across locations will be dependent since they derive from the same physical phenomena. The effects of using dependent observations in a model which assumes independence are difficult to predict in general. In the limiting case of perfectly dependent observations, no bias is present in the estimated model, but parameter uncertainty is considerably larger than would be naively expected (Jonathan and Ewans, 2007b). One way to avoid modelling dependence and yet to take account of it in eventual inferences is to preserve it by using an appropriate bootstrap (see, e.g. Davison and Hinkley, 1997, and the discussion of estimating model parameters and their uncertainties below). A pragmatic approach to modelling dependent data would then be: (1) assume that observations are independent for model fitting (potentially introducing bias and increased model uncertainty) and (2) adjust for bias and uncertainty using the so called block bootstrap scheme (see, e.g. Chandler and Bate, 2007; Chavez-Demoulin and Davison, 2005).

If attempting to characterise joint extremes over two or more locations, we might estimate a model for the joint extremal structure over those locations. In the metocean context, this might be achieved using the conditional extremes model of Heffernan and Tawn (2004) (which has been used to characterise joint extremes of H_S for whole ocean basins, or joint extremes of current with depth, Jonathan et al., 2012). This is discussed further in Section 5. Pooling observations (e.g. Heideman and Mitchell, 2009) over a neighbourhood assumes that their marginal characteristics are identical. If this is the case, pooling increases sample size and therefore reduces uncertainty in a naive model; the degree of improvement increases as the dependence between observations from different locations reduces. If however marginal characteristics are not identical, pooling can cause bias (unless we model marginal characteristics appropriately using covariates). Again, the overall effect of pooling on the bias and uncertainty of EV models is application specific and difficult to predict in general. When interested in joint extremes across locations, application of a spatial extremes model to componentwise maxima is possible but requires specialist statistical knowledge (see Section 5). Methods of composite likelihood provide a practical approach to estimate spatial models. These models are generally limited to situations where extremes from different locations can be assumed to be asymptotically dependent (see Appendix B.1). The conditional model of Heffernan and Tawn (2004) admits both asymptotic dependence and asymptotic independence.

3.3. Are covariate effects likely?

If the characteristics of observed extreme values vary systematically with one or more covariates (e.g. H_S might vary with storm direction and location), it is essential to accommodate these effects in the model. Otherwise, estimated models will be unreliable (see Jonathan et al., 2008). The most straight forward approach treats model parameters as smooth functions of covariates using appropriate bases (e.g. additive models, splines, Fourier series, random fields). These forms can be extended to multidimensional (e.g. spatio-temporal, spatio-directional) covariates using (e.g.) tensor products of marginal bases. Modelling of covariates is discussed further in Section 4.

Some practitioners reduce the effects of covariates by processing observations (sometimes referred to as “whitening” of “coloured” observations, e.g. Eastoe and Tawn, 2009; Jonathan and Ewans, 2011b) prior to modelling, arguing certain theoretical and pragmatic advantages. However, the covariate “whitening” procedure is generally performed independently of subsequent analysis, producing a disjointed inference scheme, within which it is difficult to understand the effects of one step on others.

3.4. Are there multiple variables?

When observations consist of more that one variable (e.g. H_S and spectral peak period, T_P), interest may lay in estimating EV models for each variable independently (i.e. marginal modelling), or in joint (or conditional) modelling. For example, we might be interested in extremes values of H_S and associated values for T_P at the extreme value of H_S. Sometimes physical arguments provide a motivation for joint EV model forms (e.g. a classical Weibull distribution for H_S and a log-normal for the conditional distribution of T_P given the value of H_S, see Haver, 1985). When no such physical arguments apply, the conditional model of Heffernan and Tawn (2004) can be used (as discussed in Section 5).

3.5. Threshold selection

EV models are motivated by asymptotic arguments. They are useful when observations correspond to values from the tail of a distribution. It is almost always necessary to select a threshold for modelling (which effectively defines where the "tail" of the distribution starts). In practice it is rarely clear where the threshold should be set. The objective of threshold selection is to select a threshold such that the model fits the data well above the threshold. The threshold is also generally set as low as possible (so that sample size for modelling is maximised) subject to the estimated model behaving in the anticipated manner for all thresholds higher than the threshold chosen. Graphical approaches are generally used to assess this. Threshold selection is a critical step in many other branches of statistical modelling, and is generally difficult to do well. It is essential to re-estimate models (and resulting design conditions) using a number of different plausible threshold values to check their consistency. Inconsistency indicates unsatisfactory modelling. Threshold selection is more difficult when covariate effects are present, since the threshold may itself be a function of covariates. The threshold is usually set as a local (covariate-dependent) quantile (and the problem of threshold selection transformed into one of specifying the relevant covariates and the appropriate threshold quantile level). When modelling multivariate extremes, threshold selection is further complicated by potential interactions between variables and (typically) insufficient data to characterise the joint tail.

There have been many attempts to overcome the need to select a threshold: (1) by extending the EV model into the body of the distribution (so that more observations are available for modelling, and the location of the threshold can be estimated or integrated away, sometimes within a Bayesian framework (e.g. Wadsworth et al., 2010; MacDonald et al., 2011; Tancredi et al., 2006) or (2) by averaging the model over many possible threshold choices. None of these approaches has been adopted as a standard approach in the statistics literature, suggesting that none provides material improvements over the ad-hoc graphical approach. When covariate effects are present, non-linear threshold methods (e.g. Scotto and Guedes-Soares, 2000) using perhaps quantile regression methods (e.g. Koenker, 2005; Thompson et al., 2009) are useful to model covariate-dependent thresholds, potentially with respect to multiple or multivariate covariates.

3.6. Is measurement scale appropriate?

Typical results of an EV analysis on a sample of observations are generally not invariant to measurement scale. In an ocean engineering context, we might model H_S to estimate a return value corresponding to some return period. We might also choose to model the squared values H_S^2, motivated by the fact that wave forcing on a drag-dominated structure is proportional to the square of wave height (see Tromans and Vanderschuren, 1995), then estimate the return value and square-root it. Estimates for return values from inferences on the linear and square scales will be different in general. The modeller should estimate EV models and hence return values for observations on a range of plausible scales to check that estimates of return values are consistent. Inconsistency indicates unsatisfactory estimation. The reason for the apparent inconsistency in estimation of return values on different measurement scales is poor convergence of the distribution of threshold exceedences or block maxima to the corresponding asymptotic distributional forms used to model the sample, on one or both of the measurement scales. The rate of convergence of maxima or threshold exceedances from the distribution from which the data are drawn to the asymptotic form can be improved by judicious choice of measurement scale, thereby reducing bias in estimation of return values. Reeve et al. (2012) assess the implications for H_S in the North Sea, based on the model of Wadsworth et al. (2010), motivated by the work of Cook (1982), Cook et al. (2003) and Cook and Harris (2004).

3.7. Estimating model parameters and their uncertainties

EV models are typically estimated using maximum likelihood estimation (e.g. Kalbfleisch, 1985; Pawitan, 2001), providing a consistent framework for statistical inference. Maximum likelihood estimation of a generalised Pareto model for univariate extremes in the absence of covariates is straight forward, on paper at least. Parameter uncertainty can be assessed using asymptotic results, for example using the so called delta method (e.g. Davison, 2003), or bootstrapping. When covariates are present, maximum penalised likelihood estimation is often appropriate, requiring the estimation of additional tuning parameters.

There is a large literature on estimation of extreme value characteristics based on considerations of order statistics or quantiles. Beirlant et al. (2004) provide an introduction.

3.8. Are there less than 50 independent observations from the tail?

Inferences are critically dependent on sample size. When sample size is small, parameter and return value estimates will typically be highly uncertain, casting considerably doubt on model utility. In this situation, it might be possible to supplement the sample with other complimentary data expected to have similar extremal characteristics, or constrain parameters using expert prior knowledge, perhaps within a Bayesian framework. Consideration of peaks over threshold typically allows more observations to be retained for model estimation than in the case of block maxima, and should be preferred.

It is possible to constrain one or more model parameters in a number of ways: (1) by fixing parameters to known values based on previous analysis, (2) by performing a Bayesian analysis in which model parameters are given prior distribution which can constrain the values they take, (3) by assuming a restrictive parametric form for the model. Robust methods (e.g. Vandewalle, 2004) are also available, but are generally computationally more challenging to estimate and less intuitive to interpret. L-moments (Hosking, 1990, e.g. Pandey et al., 2001), being linear combinations of the order statistics of a sample, are more robust to outliers than their conventional counterparts. Non-parametric approaches (e.g. maximum entropy estimation with moment constraints) have also been proposed (e.g. Petrov et al., 2013).

4. Incorporating directionality and other covariate effects

Conventionally, design wave criteria are based on omnidirectional estimates – that is, sea state criteria that are relevant for any direction of approach. However, to take advantage of directional effects in optimising cost of offshore facilities, and with increasing availability of good wave direction data, it is common to provide directional estimates of extreme significant wave heights, in addition to the omni-directional estimate. The practice has been to associate an appropriate wave direction, such as the mean wave direction or the direction of spectral peak, with the significant wave height of a given sea state. The significant wave height values selected for analysis are then binned into directional sectors, with sector limits chosen arbitrarily either on a fixed number (typically eight) centred on the cardinal directions, or on the basis of a perceived directionality in the data. Both sectoring approaches make the assumption of homogeneity within each sector. Extremal analyses are then performed on

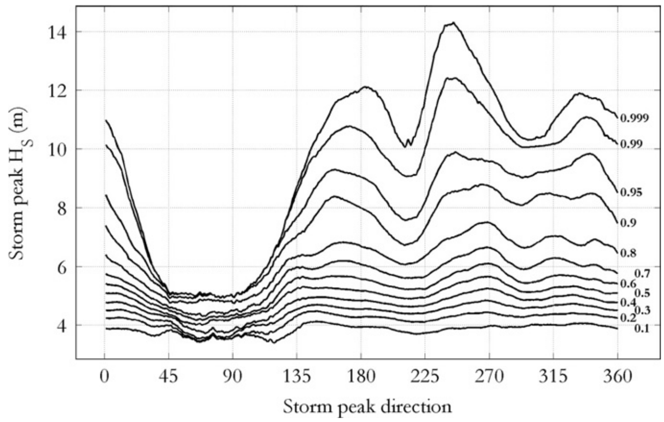

Fig. 1. Quantiles of storm peak H_S with direction for the northern North Sea hindcast data discussed in Jonathan and Ewans (2007a). The characteristics of extreme values of storm peak H_S are dependent on storm peak direction.

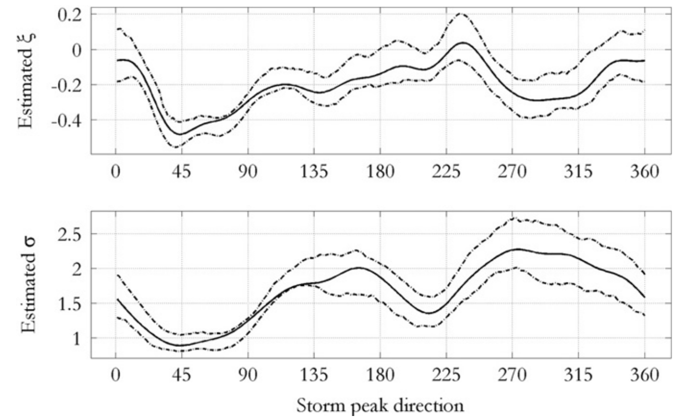

Fig. 2. Estimated values for generalised Pareto shape ξ and scale σ for the northern North Sea hindcast data discussed in Jonathan and Ewans (2007a), with 95% bootstrap confidence bands. The characteristics of the estimated parameters reflect those of the quantiles shown in Fig. 1.

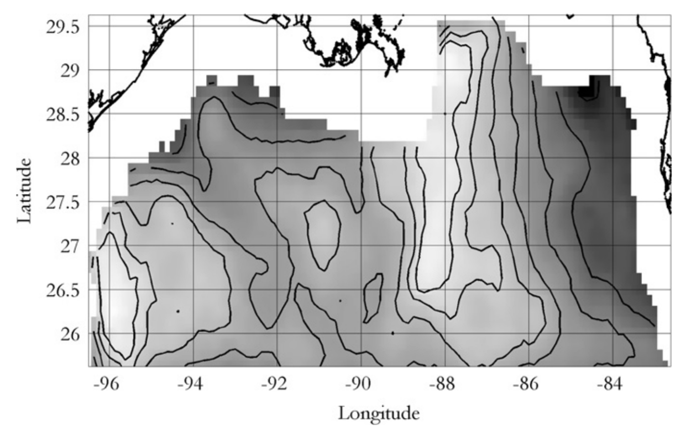

Fig. 3. Illustrative contours of estimated median 100-year storm peak H_S for a grid of locations throughout the U.S. Gulf of Mexico. Values of contour levels have been withheld for reasons of confidentiality. From Jonathan and Ewans (2011b).

the significant wave height values in each sector, and also on the complete set of values ignoring direction. Design values with a given return period are then specified for each directional sector and for all directions, the latter being the omnidirectional estimate. Unfortunately, this approach has often led to inconsistencies in design criteria, such as that the probability of exceedence of a given significant wave height when calculated from the directional criteria is different from that obtained from omni-directional criteria. A discussion of this problem is given by Forristall (2004). Guidelines such as API (2005) and ISO (2005) provide recommendations on treating directional criteria, but even when these are followed, either inconsistency remains (in the case of API), or insufficient detail is given on how to make the criteria consistent (in the case of ISO). Recently, Mendez et al. (2006), Jonathan and Ewans (2007a, 2011a), Mendez et al. (2008), and Ewans and Jonathan (2008) have borrowed from the statistics literature to model directional and seasonal effects smoothly in a way that better reflects nature and observed data in application to numerous ocean basins, including the Gulf of Mexico, the North Sea and the North East Pacific. Consistency between directional and omni-directional criteria follow directly in a rigourous way in these analyses (e.g. Ewans and Jonathan, 2008), but the analyses also demonstrate a perhaps more significant aspect to establishing omni-directional criteria perhaps not widely recognised by practitioners. Specifically, (i) a directional model generally explains the observed variation of extremes significantly better than a model which ignores directionality, (ii) omni-directional criteria developed from a directional model are different from those generated when directionality is not accounted for, and thus (iii) omni-directional criteria derived from a directional model are more realistic and should be preferred in general. These facts challenge traditional practice and warrant special attention. Accordingly, Jonathan et al. (2008) undertake specific simulation studies to demonstrate unambiguously that omni-directional criteria should be derived using a directional model in cases where good directional data exist. Covariate effects are not restricted to direction, but are equally relevant for (e.g.) time and space.

The requirement for modelling of covariate effects must be considered in light of all other pertinent issues. When physical effects such as fetch variability with direction, seasonal cycles of storm severity, variation of water depth or climate change are anticipated, inclusion of the corresponding covariates is likely to be important for good model fit. Conversely, it is important also not to overfit by including covariate effects which are not supported by data, especially when sample size is small. Statistical tests can be used to select formally between models with or without covariate terms based on their goodness of fit to the sample. Techniques such as cross-validation offer one approach to ensuring parsimony in terms of predictive performance. Appropriate model complexity can also be regulated by formal statistical testing, or by adopting a cost-complexity criterion. The behaviour of parameter and quantile estimates with threshold should be examined and compared with expectation from theory and previous application. The stability of quantile estimates for directional sectors with respect to threshold and small changes to sector boundary specifications should also be examined. Neglecting covariates can lead to biased estimates of extreme quantiles. The extent of any bias cannot be anticipated a priori in general. Moreover, the relative size and direction of biases due to neglect of covariate(s) compared to other sources (e.g. mis-specification of threshold, of model form, outliers, sample size) is also difficult to estimate before the analysis is performed. Nevertheless, failure to model covariate effects adequately when there is evidence for their presence will likely compound errors.

Numerous authors have reported the essential features of EV modelling with covariates (e.g. Davison and Smith, 1990) and the importance of considering different aspects of covariate effects. Coles and Walshaw (1994) describe directional modelling of extreme wind speeds. Robinson and Tawn (1997) describe directional modelling of extreme sea currents. Anderson et al. (2001) report that estimates for 100-year maximum H_S from a model ignoring seasonality are considerably smaller than those obtained using a number of different seasonal models. Chavez-Demoulin and Davison (2005) and Coles (2001) provide straight-forward descriptions of a non-homogeneous Poisson model in which occurrence rates and extremal properties are modelled as functions of covariates.

4.1. Point process model

The inhomogeneous Poisson process model (outlined in Appendix A.6 and discussed, e.g. by Coles, 2001; Embrechts et al., 2003; Davison, 2003; Beirlant et al., 2004) is a three parameter model describing the rate of occurrence and intensity of extreme events. If realisations from this model are observed over some interval, the number of exceedances of threshold u is Poisson-distributed. Conditional on a given number of exceedances, values of the exceedances are then a random sample from a GP distribution (see Davison and Smith, 1990). This model is estimated using maximum likelihood estimation. The log-likelihood can be written

$$l(\lambda,\xi,\sigma)=l_N(\lambda)+l_W(\xi,\sigma)$$

where l_N is the log-density (or log-likelihood) of the total number of exceedances (with rate argument $\lambda>0$), and l_W is the log-conditional-density of exceedances (with the usual GP form, given in Eq. (1)) given a known total number $N=n$ for $n=1,2,3,\ldots$. Since the log-likelihood can be partitioned in this way, inferences on λ can be made separately from those on GP shape $\xi\in(-\infty,\infty)$ and scale $\sigma>0$. Each of the parameters is modelled as a smooth function of one or more covariates, using penalised maximum likelihood estimation, adopting a suitable smooth basis. For directional modelling, a Fourier series (e.g. Ewans and Jonathan, 2008; Jonathan and Ewans, 2011a) or spline (Green and Silverman, 1994) basis may be appropriate. For spatial covariates, Legendre polynomials (e.g. Northrop and Jonathan, 2011), random fields (e.g. Rue and Held, 2005) or spatial splines (e.g. Marx and Eilers, 1998; Ramsay, 2002; Ruppert et al., 2003; Jonathan and Ewans, 2011b) are suitable. Some approaches (e.g. Marx and Eilers, 1998) are relatively easily extended to include multiple or multivariate covariates (e.g. 2D space, direction and time). Vector generalised additive models (VGAM) for extremes (Yee and Wild, 1996; Yee and Stephenson, 2007 and the VGAM software) provide useful tools, as do Bayesian Hierarchical Models (BHM, e.g. Gilleland et al., 2006; Cooley et al., 2007; Sang and Gelfand, 2009). However, spatial applications usually require sometimes unrealistic (conditional) dependence assumptions (see, e.g. Turkman et al., 2010 and the discussion of spatial extremes in Section 5 and Appendix B).

The Poisson log-likelihood, for arrivals at times $\{t_i\}_{i=1}^{n}$ in interval $(0,T]$ is

$$l_N(\lambda)=\sum_{i=1}^{n}\log\lambda(t_i)-\int_0^T\lambda(t)\,dt$$

Chavez-Demoulin and Davison (2005) describe an approximate log-likelihood achieved by partitioning the interval $(0,T]$ into a large number m of sub-intervals of length δ, choosing δ small enough that λ is effectively constant over each sub-interval. The GP model from above (with shape parameter ξ, scale parameter σ and threshold u) is used to model the sizes of exceedances, where each of ξ, σ and u are smooth functions of covariates.

4.2. Penalised likelihood

When covariates Θ are present, since we seek parameters which are smoothly varying with respect to covariates, one approach is to maximise a penalised (negative log) likelihood $\ell^*(\Theta)$

$$\ell^*(\Theta)=l(\Theta)+\kappa R(\Theta)$$

where $l(\Theta)$ is the original sample (negative log) likelihood (e.g. for the Poisson or GP models above), and $R(\Theta)$ is the roughness of parameter estimates with respect to covariates. $\kappa>0$ is a roughness coefficient required to balance the relative importance of maximising likelihood and minimising parameter roughness. The roughness coefficient κ needs itself to be estimated, possibly using cross-validation or a measure of model cost-complexity such as the Akaike Information Criterion (AIC) or Bayes Information Criterion (BIC). In cross-validation, the original sample is typically partitioned appropriately into 5–10 groups. Each group is withheld in turn, the remaining groups used for estimation (for each of a set of plausible values of κ), and the withheld group used to assess predictive performance (e.g. Stone, 1974). Cross-validation can also be used (possibly nested) to provide an unbiased assessment of performance. Parameter uncertainty can be estimated by appropriate bootstrap resampling. The original sample is resampled with replacement a number of times, and the complete estimation procedure executed for each resample. The variability of estimated model parameters and return values across resamples provide estimates for the uncertainties of parameters and return values for the original sample (see, e.g. Davison and Hinkley, 1997; Wang and Wahba, 1995).

In covariate models, the roughness coefficient κ can sometimes be treated as a model parameter, in which case it too can be estimated (possibly using Laplace Approximations if appropriate, e.g. Rue et al., 2009). Within a Bayesian framework, model and return value uncertainty are estimated naturally as part of inference. Typically, estimation proceeds using Monte Carlo Markov chain methods to sample the posterior joint distribution of parameters and functions thereof (e.g. Gamerman and Lopes, 2006). When covariate effects or multivariate extremes are present, Bayesian procedures can be computationally demanding; approximate methods are often therefore necessary.

4.3. Modelling steps

For a sample of independent observations, we might adopt the following modelling steps: (1) Define an appropriate probability of threshold exceedence p_u, (2) estimate the threshold u as a function of covariates, e.g. using quantile regression with exceedence probability p_u, (3) estimate the rate of threshold exceedence λ using the Poisson process model, (4) estimate the shape ξ and scale σ of the generalised Pareto distribution, and (5) simulate under the model to estimate return values corresponding to desired return periods.

4.4. Illustrations

The directional dependence of large values of storm peak H_S for northern North Sea hindcast data is discussed by Jonathan and Ewans (2007a) and illustrated in Fig. 1. The land shadow of Norway is evident for storm peak directions in the interval (45,100). Storms travelling up the North Sea from the South have directions in the interval (140,200), and those travelling down from the Norwegian Sea approximately (290,360). The most extreme storm peak events, and those with the longest extreme value tail, can be seen to correspond to Atlantic storms with directions in the interval (230,270). The corresponding estimates for GP shape ξ and scale σ, with 95% bootstrap uncertainty bands, are shown in Fig. 2. The largest value of shape corresponds to Atlantic storms as expected. Minima of both shape and scale correspond to short-fetched storms from the Norwegian coast.

Fig. 3 shows contours of marginal median 100-year storm peak H_S for locations in the Gulf of Mexico, estimated from a thin-plate spline model in Jonathan and Ewans (2011b).

5. Multivariate modelling

Careful statistical description is important to understand extreme ocean environments, and design and assessment of ocean structures. For example, wave climate can be described in terms of sea-state variables such as the significant wave height, H_S, and spectral peak wave period, T_P, or mean zero up-crossing wave period, T_Z. Spectra for an extreme sea state, such as the 100-year return-period sea state are often required for assessing dynamic loads. Spectral properties of extreme sea states are typically estimated by jointly estimating values of H_S and T_P corresponding to a given return period. Alternatively, we might wish to model storm peak H_S over a spatial grid of locations; values of H_S at neighbouring locations will be dependent in general. Or we might be interested in modelling extreme current profiles with depth. In any case, were we able to justify assuming (effective) independence of random variables, modelling would be relatively straight forward – but this is usually not possible. In general, estimation of extreme wave environments requires that extremal dependence (between extreme values of different variables) be characterised adequately.

In univariate EV theory, the asymptotic form of the distribution of threshold exceedances (i.e. GP) and block maxima (i.e. GEV) can be derived theoretically. The same is not true for multivariate extremes. No finite-dimensional parameterisation of the class of multivariate EV distributions exists. Therefore, modelling typically proceeds by selecting one of a variety of sub-families of distributions or a suitable asymptotic argument. Unfortunately there is as yet no unifying approach, and the literature is rather confusing. The work of Beirlant et al. (2004) provides an accessible introduction. For practical application, we find that the conditional extremes model of Heffernan and Tawn (2004) is a flexible framework for general multivariate extreme value modelling which is easily implemented and extended. If concerned specifically with spatial extremes, methods related to max-stable processes are promising, despite unrealistic modelling assumptions regarding componentwise maxima and restrictive modelling assumptions (see Section 5.5).

Appendix B provides a motivation for extremal dependence thinking and describes simple diagnostic statistics for the different forms of extremal dependence. The appendix also motivates ideas of spatial extremes, including a description of the extremal coefficient diagnostic useful for site pooling.

5.1. Types of model for multivariate extremes

Suppose we wish to model extremes of a sample of values drawn from some multivariate distribution. We need to estimate the tail behaviour of the multivariate distribution from which the sample is drawn, based on the sample alone. There are essentially four approaches we could take, described below under the headings extremal dependence models, parametric models, conditional extremes models and max-stable models. The extremal dependence approach is useful primarily for introducing the different forms of extremal dependence (namely asymptotic dependence, asymptotic independence and perfect independence), and estimating the coefficient of tail dependence (see Section 5.2); it can also be used directly for modelling when all variables become extreme together. A large number of parametric models for multivariate extremes has been proposed, which themselves can be classified according to the form of their extremal dependence (see Section 5.3). The conditional extremes model is motivated by different asymptotic arguments, and provides the basis for modelling more general situations for all forms of extremal dependence (see Section 5.4). Max-stable models on componentwise maxima (see Section 5.5) allow extension of max-stability (see Appendix A.2) to the multivariate case, in a rather restrictive sense as explained below, and are the basis for models of spatial extremes using max-stable processes.

5.1.1. Transforming marginals

Many models for multivariate extremes are most easily described and applied when all variables follow a common marginal distribution. In practice we achieve this using the probability integral transform (see, e.g. Jonathan et al., 2009). For a sample $\{x_{ij}\}_{i=1,j=1}^{n,p}$ of n observations on p variables $X_1,X_2,\ldots,X_p$, we model variable X_j marginally independently using an appropriate distribution F_j (e.g. GP for threshold exceedances). Then for each observation x_{ij} of that variable, we evaluate the cumulative distribution function $\hat{F}_j(x_{ij})$. We then find the value of the argument x^*_{ij} of the desired common marginal cumulative distribution function $F^*(x^*_{ij})$ (which might be Gumbel or Fréchet) such that

$$\hat{F}_j(x_{ij}) = F^*(x^*_{ij}) \quad \text{for all } i,j$$

thereby establishing a transformed sample $\{x^*_{ij}\}_{i=1,j=1}^{n,p}$ exhibiting common marginal distribution F^*. The remaining dependence structure between variables can be described as a copula. Typical forms for the F^* are the standard Gumbel cumulative distribution function $F^*(x)=\exp(-\exp(-x))$ for $x\in(-\infty,\infty)$, and the standard Fréchet cumulative distribution function $F^*(x)=\exp(-1/x)$ for $x>0$.

5.1.2. Copula models

The joint distribution F of random variables $X_1,X_2,\ldots,X_p$ may be written as

$$F(x_1,x_2,\ldots,x_p) = C\{F_1(x_1),F_2(x_2),\ldots,F_p(x_p)\}$$

where $F_1,F_2,\ldots,F_p$ are univariate marginal distributions and C is a p-dimensional distribution on $[0,1]^p$ known as a copula. C is uniquely determined for well-behaved distributions F. If $F_1,F_2,\ldots,F_p$ are well-behaved, we can also write

$$C(u_1,u_2,\ldots,u_p) = F\{F_1^{-1}(u_1),F_2^{-1}(u_2),\ldots,F_p^{-1}(u_p)\}$$

for $\{u_1,u_2,\ldots,u_p\}\in[0,1]^p$. For example, we might assume that F is a multivariate Gaussian and $F_1,F_2,\ldots,F_p$ all GP with different parameters. For a copula to describe max-stable multivariate extremes, it must satisfy max-stability constraints (see Appendix B.3.2).

5.2. Extremal dependence

Consider modelling extremes of the random vector $X_1,X_2,\ldots,X_p$ (corresponding to, e.g. (H_S,T_P) for $p=2$), assuming that marginal distributions of individual variables have been transformed to Gumbel scale by means of the probability integral transform above. Extremal dependence modelling assumes particular behaviour of the joint survivor function as described in more detail in Appendix B.1, applicable for a wide range of multivariate EV distributions (but not all, see, e.g. Schlather, 2001). We describe the joint tail of a distribution using its coefficient of tail dependence, η, a constant $\in(0,1]$, which quantifies the extent of extremal dependence for the distribution. For Gumbel marginals, and large values $(x_1,x_2,\ldots,x_p)$ we have (from Appendix B.1 and Ledford and Tawn, 1997)

$$\begin{aligned}\Pr(X_1>x_1+t,X_2>x_2+t,\ldots,X_p>x_p+t)\\ \approx \exp(-t/\eta)\Pr(X_1>x_1,X_2>x_2,\ldots,X_p>x_p)\end{aligned} \quad (2)$$

for fixed $t>0$. When $\eta=1$ we say that the distribution is asymptotically dependent. Otherwise, it is asymptotically independent. As described in Appendix B.1, the value of η can be estimated directly from the sample. Eq. (2) can therefore also be used directly for estimation of probabilities associated with extreme sets of the form $(X_1>x_1+t,X_2>x_2+t,\ldots,X_p>x_p+t)$ provided that an appropriate set $(X_1>x_1,X_2>x_2,\ldots,X_p>x_p)$ exists in the sample. The problem is essentially therefore one of threshold selection.

Asymptotic independence requires careful thought. For example, note that for a bivariate Normal distribution with correlation ρ, the coefficient of tail dependence is given by $\eta=(1+\rho)/2$. In the case $\rho=1$ we have perfect dependence between the variables and also asymptotic dependence. When $\rho<1$ we have $\eta<1$ and hence variables will be asymptotically independent (even for ρ close to unity). Wrongly assuming a model imposing asymptotic independence will reduce the rate of joint occurrence of rare events.

5.3. Parametric models

Numerous authors in the statistic literature (e.g. Tawn, 1988a, 1988b; Joe, 1990; Coles and Tawn, 1991) have proposed parametric forms of bivariate and multivariate extreme value distributions exhibiting different dependence structures. Kotz and Nadarajah (2000) and Beirlant et al. (2004) provide relatively recent reviews. These models are typically limited to low dimensions. Modelling the sample requires fitting the distribution of choice, estimating its parameters, then confirming goodness of model fit and modelling assumptions made using diagnostic tests. The approach is valid when the distributional form is chosen appropriately. Almost all parametric distributions appropriate for modelling extremes of multivariate samples can be classified in terms of the value of their coefficient of tail dependence η, a function of the parameters of the distribution. However, it is often difficult to decide between different parametric forms (with possibly different extremal dependence structures) based on goodness of fit statistics. Direct estimation of η may be useful in these circumstances to advise model form.

In the engineering literature, Ferreira and Guedes-Soares (2002) present a method to estimate bivariate distributions of significant wave height, H_S and mean wave period, T_Z. The method consists of (1) transforming individual variables using the Box–Cox transformation so that, marginally, transformed values are consistent with samples from a Normal distribution (as described above), (2) fitting a bivariate Normal model to transformed data. The approach is illustrated in application to modelling of Waverider buoy data at Figuera da Foz, Portugal. The approach is appropriate for modelling the body of a bivariate distribution, provided that the Box–Cox transformations provide fit adequately. However, the method imposes a particular structure to extreme tail behaviour which may not be supported by the sample. In a similar vain, methods of (multivariate) kernel density estimation, appropriate for modelling the body of a distribution, impose the kernel structure on extreme tails also with little or no justification. Dong et al. (2013) develop bivariate models with specified parametric forms for the joint distribution of significant wave height and peak period, using maximum entropy estimation; this work also suffers from a lack of asymptotic justification for the model form assumed.

5.4. Conditional extremes

There are short-comings to direct fitting of parametric multivariate EV distributions, since typically the appropriate distributional form is unknown, and different distributional forms are often impossible to distinguish based on empirical evidence. Direct estimation using Eq. (2) is also problematic, e.g. for samples which are not simultaneously extreme in all their components, since the set $(X_1>x_1+t, X_2>x_2+t, \ldots, X_p>x_p+t)$ is likely to be empty if $(X_1>x_1, X_2>x_2, \ldots, X_p>x_p)$ is not. Moreover the approach becomes unmanageable as p increases; for example, we might wish jointly to model extremes of wave spectral parameters (H_S,T_P) over 1000 locations, requiring a 2000-dimensional model, which would not easily be estimated without making simplifying assumptions – parameter estimation is intractable. Authors have therefore taken a different conditional approach, the basis of which is to take each variable in turn as a conditioning variate, and model remaining variables conditional on a large value of the conditioning variate. In a bivariate case we might model $X_2|X_1=x_1$ for large x_1, and $X_1|X_2=x_2$ for large x_2. The advantage of this approach is that it might extend to higher dimensions, and to samples for which not all variables are simultaneously extreme. A disadvantage until recently, was that the form of the conditional model had to be assumed (or itself estimated) from the sample. For example, Haver (1985) proposed a model for extremes of (H_S,T_P) data, based on (1) modelling the extreme tail of H_S using a Weibull form, (2) modelling the distribution of T_P, for a given (range of) H_S using a log-normal form, (3) estimating values for the parameters in the log-normal distribution for $T_P|H_S=h$ corresponding to extreme values h (beyond the sample) by regressing the log-normal parameter estimates on H_S for the sample and extrapolating, and hence estimating extreme quantiles of (H_S,T_P). Justification for the choice of log-normal conditional model, and for the forms of models for log-normal parameters as a function of H_S was goodness of fit to the sample; these may not be adequate when extrapolating beyond the sample.

The conditional approach of Heffernan and Tawn (2004) goes some way to overcoming these difficulties. Based on asymptotic arguments, they derive a parametric equation for the form for one variable conditional on a large value of another, valid for extremes from a wide class of multivariate distributions with Gumbel marginals. The conditional approach consists of (1) transforming each variable so that marginally its values are consistent with a sample from the standard Gumbel distribution, (2) estimating the parameters of the parametric model for values of one variable conditional on large values of the other, (3) simulating (and back-transforming) to characterise extremal behaviour of the joint distribution (on the original scale). The major advantage of the Heffernan and Tawn conditional approach compared with previous conditional models, is that the functional forms for marginal fitting and conditional modelling are motivated by asymptotic arguments. There is also no need to estimate the coefficient of tail dependence nor restrict attention to componentwise maxima. Model form, motivated by the assumption of a particular limit representation for the conditional distribution, is appropriate to characterise the conditional behaviour of a wide range of theoretical examples of bivariate (and higher-dimensional) distributions for extremes. For positively associated variables $X_1,X_2,\ldots,X_p$ the model for large x_j is

$$(X_{-j}|X_j=x_j)=a_{-j}x_j+x_j^{b_{-j}}Z_{-j} \tag{3}$$

where X_{-j} represents all variables except X_j, and a_{-j} and b_{-j} are vectors of location and scale parameters respectively to be estimated, with $a_{-j}\in[0,1]^{p-1}$ and $b_{-j}\in(-\infty,1)^{p-1}$, and componentwise multiplication is understood. Z_{-j} are random $(p-1)$-vectors, independent of X_j, converging with increasing x_j to a non-degenerate limiting distribution G_{-j}. Joint tail behaviour is then characterised by $\{a_{-j}\}_{j=1}^p$, $\{b_{-j}\}_{j=1}^p$ and $\{G_{-j}\}_{j=1}^p$. The form of distributions G is not specified by theory. Note that modelling using Eq. (3) reduces to estimating a (large) number of straight forward regression models for pairs of variables.

From a practical perspective, the main advantages of conditional extremes (and related) models are that they admit all forms of extremal dependence (as opposed to approaches based on max-stable models), and that they are relatively simple to implement and apply in high dimensions.

5.4.1. Illustrations

The conditional approach has been used as the basis for studies by us, including joint extremes of wave spectral parameters (Jonathan et al., 2010) and extreme current profiles with depth (Jonathan et al., 2012). In Jonathan et al. (2010), the joint characteristics of storm peak period (T_P) and storm peak H_S are quantified using the conditional model for measured and hindcast data taken from different locations. Fig. 4 illustrates estimates for conditional model parameters a and b from Eq. (3) for the four applications considered. The figure suggests that the northern North Sea and Gulf of Mexico measured samples in particular are different to one another with respect to parameter a. We would expect this difference to be reflected in design criteria for the two locations, for example with respect to specification of design contours (Jonathan et al., 2011c).

In Jonathan et al. (2012), observed vector ocean currents at each of a number of water depths are expressed as components with respect to major- and minor-axis of current variation at that depth using Principal Components Analysis. Each current component is then independently decomposed into the sum of (deterministic periodic) tidal and (random) non-tidal currents using a local harmonic model. The marginal and dependence structure of extremes of hourly maxima and minima of non-tidal components is characterised using the conditional extremes model above. We simulate under this model to estimate characteristics of extreme current profiles corresponding to arbitrary return periods, and quantify the uncertainty of those estimates. For a sample collected over a 2.5-year period in 250 m water on the outer shelf of North Western Australia, the model predicts monthly instantaneous extreme conditional profiles well. Fig. 5 compares median conditional total current profiles (along the major axis of current variation at that depth) with measured profiles for a 10 year return period. The conditional profile provides an upper bound for the instantaneous observed profile.

In Ewans and Jonathan (accepted for publication) the conditional extremes model is extended to incorporate covariate effects.

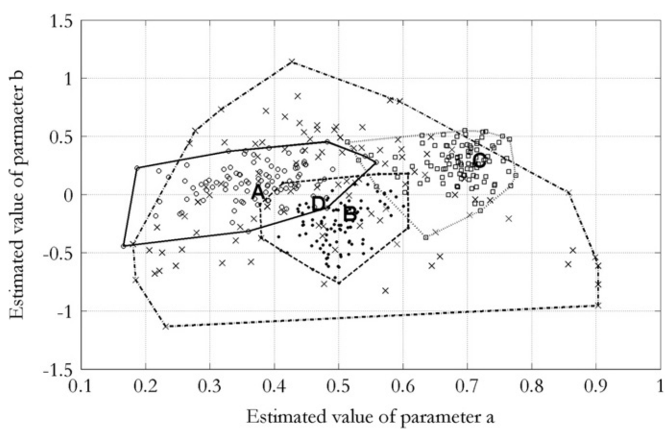

Fig. 4. Point and bootstrap convex hull estimates for conditional model parameters a and b for four applications of modelling storm peak T_P given storm peak H_S. Point estimates (A–D) and bootstrap convex hulls shown corresponding to a measured northern North Sea sample (A, circles), a northern North Sea hindcast sample (B, dots), a measured Gulf of Mexico sample (C, squares), and a Australian North-West shelf hindcast (D, crosses). From Jonathan et al. (2009).

5.5. Multivariate max-stable models

Here, ideas of max-stability (see Appendix A.2) are extended from a univariate to a multivariate context. In order to achieve this, we are forced to consider modelling of so called componentwise maxima. For a set of n observations $\{x_{ij}\}_{i=1,j=1}^{n,p}$ on p variables $X_1, X_2, \ldots, X_p$, the value of the componentwise maximum for variable j is $m_j = \max_i\{x_{ij}\}$, the maximum value across all observations for that variable. In the componentwise maximum approach we perform EV analysis on multivariate observations of the form $m = \{m_j\}_{j=1}^{p}$. Note that m may not actually exist in the sample! Componentwise maxima are useful to model certain types of dependence structure (namely asymptotically dependence and perfectly independence) but not others (namely asymptotically independence). Despite these restrictions,

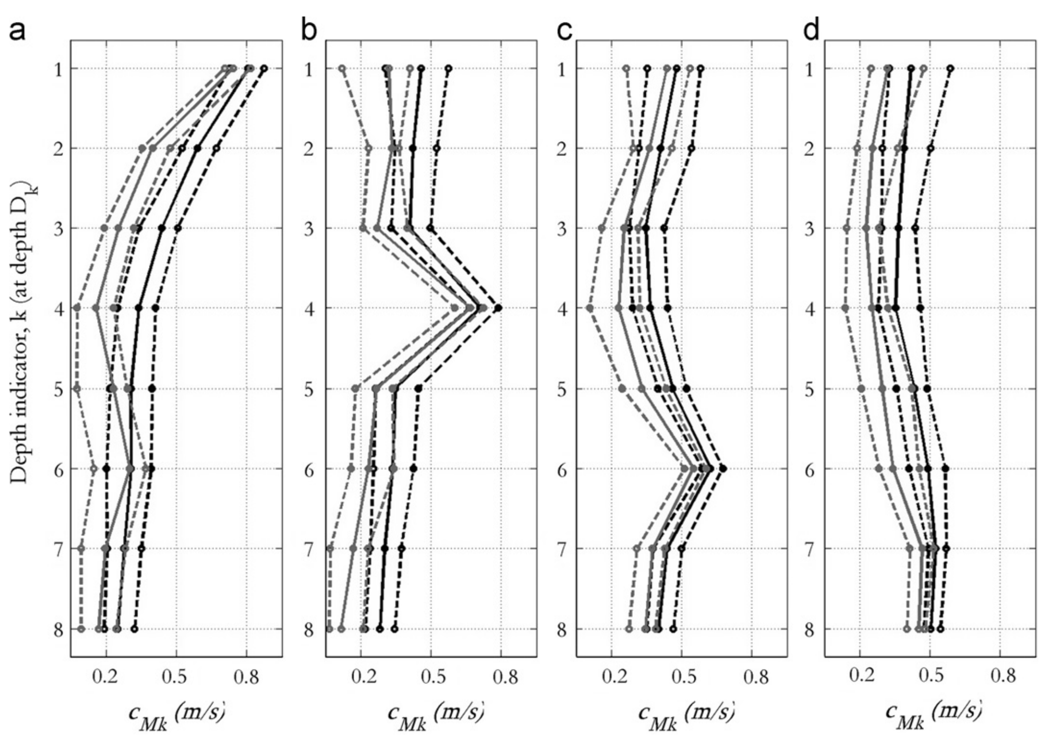

Fig. 5. Median (solid) conditional total current profiles along major axis, c_{Mk}, with depth ($k = 1,2,\ldots,8$) for monthly maxima of (a) c_{M1}, (b) c_{M4}, (c) c_{M6} and (d) c_{M8} estimated from conditional simulation of hourly extremes (black) and measurements of measured profiles (grey). Also shown are corresponding 25 percentiles and 75 percentiles (dashed). Conditional profiles from simulation provide an upper bound for the instantaneous profile. From Jonathan et al. (2012).

max-stable models provide a useful framework for multivariate extreme value analysis, particularly spatial extremes (see below). We stress that max-stable models apply to componentwise maxima only, and may therefore exaggerate dependence effects, although recent work (see Appendix B.2) suggests how these restrictions may be overcome.

5.5.1. *Spatial extremes*

The last decade has seen the emergence of useable statistical models based on max-stable processes, at least within in academia. The application of max-stable processes is complicated due to unavailability of the full multivariate density function. Padoan et al. (2010) develop inferentially practical, likelihood-based methods for fitting max-stable processes derived from a composite likelihood approach. The procedure is sufficiently reliable and versatile to permit the simultaneous modelling of marginal and dependence parameters in the spatial context at a moderate computational cost. Davison and Gholamrezaee (2012) describe an approach to flexible modelling for maxima observed at sites in a spatial domain, based on fitting of max-stable processes derived from underlying Gaussian random process models. Generalised extreme-value margins as assumed throughout the spatial domain, and models incorporate standard geo-statistical correlation functions. Estimation and fitting are performed through composite likelihood inference applied to observations from pairs of sites. Davison et al. (2012) also provide a good introduction and review. Erhardt and Smith (2011) use approximate Bayesian computation to circumvent the need for a joint likelihood function by instead relying on simulations from the (unavailable) likelihood avoiding the need to construct composite likelihoods at higher computational cost.

These spatial extremes methods are potentially of value to metocean design, since they provide a framework within which extremal behaviour of complete ocean basins can be modelled, incorporating appropriate marginal and dependence structure and avoiding the need for site pooling in particular. It might be possible that in future, only one EV modelling task would be necessary per hindcast. That model, for the whole ocean basin, could then be interrogated routinely to estimate design values for a particular location. The ocean engineer would no longer in principle need to perform further site-specific analysis. For this reason, there follows a brief motivation for max-stable processes and the composite likelihood methods to estimate them. The outline draws heavily on the work of Smith (1990) and the article of Padoan et al. (2010).

Typically, marginal modelling is thought of separately to modelling of dependence structure; we assume GEV or GP marginals, then worry about a model for the multivariate dependence structure or copula. These models only admit certain types of dependence structure, namely perfect independence or asymptotic dependence; they do not admit asymptotic independence in particular (see, e.g. Gabda et al., 2012), whereas the model of Heffernan and Tawn (2004) does. This is possibly not of great concern if diagnostics suggest that asymptotic dependence is appropriate, given other advantages the models offer and other mitigating recent developments.

5.5.2. *Max-stable processes*

From the practical perspective, max-stable processes are attractive since they provide a standard way of thinking about spatial extremes, including a framework for estimating EV models and making inferences (e.g. simulating) using the models. They describe the distribution of maxima (in time) within an arbitrary spatial domain. Appendix B.2 gives a mathematical outline. Max-stable processes can be thought of as infinite-dimensional generalisations of EV distributions. They require the definition of a function $f(s,r)$ called the event profile function, which describes (in a typical metocean context) how a storm centred at location s influences extremes (perhaps of H_S) at location r. In 2-D space, we might assume a Gaussian form for f

$$f(s,r)=f_0(s-r)=(2\pi|\Sigma|^{1/2})^{-1}\exp\left\{-\frac{1}{2}(s-r)'\Sigma^{-1}(s-r)\right\}$$

parameterised by the diagonal elements (σ_{11} and σ_{12} and the off-diagonal σ_{12} of the covariance matrix Σ). In fitting this model using a sample, we estimate the elements of Σ. For mainly bivariate choices of f (including the above) we can write down the joint distribution and, in particular, the joint density and hence the pairwise likelihood of extremes at any two locations in closed form.

5.5.3. *Composite likelihood*

The analysis of spatial extremes is concerned with the joint modelling of a spatial process at large numbers of locations. The full likelihood function describes the dependence across all locations. However, for max-stable processes, it is not possible to write down this likelihood for more than pairs of locations in general. Standard maximum likelihood estimation is therefore impossible. Fortunately, methods of composite likelihood approximation have been developed (Lindsay, 1988; Cox and Reid, 2004; Varin, 2008) which permit the simultaneous and consistent modelling of joint and marginal parameters in the spatial context at a moderate computational cost.

For a sample $\{\{x_{ij}\}_{i=1}^{n_j}\}_{j=1}^{p}$ of $n=\sum_{j=1}^{p} n_j$ observations from p locations, the full likelihood function would take the form

$$L(\text{parameters related to all locations};\ \{\{x_{ij}\}_{i=1}^{n_j}\}_{j=1}^{p})$$

were it accessible. Were we able to assume (unrealistically) that observations corresponding to different locations were independent, the full likelihood function would be

$$\prod_{j=1}^{p} L_j\ (\text{parameters related to location } j;(\{x_{ij}\}_{i=1}^{n_j}))$$

where L_j, $j=1,2,\ldots,p$ is the likelihood corresponding to the jth location. In a spatial context, we might expect that dependence between pairs of locations decreases with increasing distance, motivating an approximate likelihood of the form

$$\prod_{j=1}^{p}\prod_{k=1,k\neq j}^{p} w_{jk}L_{jk}\ (\text{parameters related to locations } j,k;\{x_{ij}\}_{i=1}^{n_j},\{x_{ik}\}_{i=1}^{n_k})$$

where L_{jk}, $j,k=1,2,\ldots,p, j\neq k$ is the pairwise likelihood for jth and kth locations and $w_{jk}\geq 0$, j, $k=1,2,\ldots,p$ are distance-related weightings. Since the pairwise likelihood between locations is available in closed form, we have a procedure to model marginal and spatial extreme characteristics across space. Although not fully understood theoretically, composite likelihood methods have been found to perform well in many applications, including spatial extremes (e.g. Ribatet et al., 2012). The efficiency of pairwise likelihood methods for simple time series problems has received careful attention. Efficiency tends to be poor in (e.g.) long-memory auto-regressive and in moving-average processes (e.g. Davis and Yau, 2011).

In principle, the parameters of the marginal and spatial EV models might be considered smooth functions of covariates (e.g. location, storm direction). The ideas introduced in Section 4 are therefore relevant here too, as are random fields and spatial (Gaussian) process models for covariates (e.g. Davison et al., 2012). The max-stable Gaussian EV process with event profile $f(s,r)$ described above has an equivalent max-stable copula representation.

6. Bayesian inference

Informally, Bayesian inference seeks to update prior information about (the parameters or variables Θ of) a model with information gained from observation to generate improved posterior information. The prior information is specified by a probability density $f(\Theta)$. The information gained from observation is specified by the likelihood $f(D|\Theta)$ of a sample D ($=\{x_i\}_{i=1}^n$ say) of observations related to the precise model form being estimated. These are then combined using Bayes theorem

$$f(\Theta|D) = \frac{1}{\int_\Theta f(\Theta)f(D|\Theta)\,d\Theta} \times f(\Theta) \times f(D|\Theta) \quad (4)$$

to yield posterior information specified in the posterior density $f(\Theta|D)$. Bayesian inference provides an intuitively appealing and self-consistent basis for modelling. It enables the modeller to combine data from observation with past experience and expertise in a consistent manner. It enables prediction and modelling uncertainty to be estimated naturally. For example, we can estimate the predictive density $f(q|D)$ of extreme quantile q given the posterior density

$$f(q|D) = \int_\Theta f(q|\Theta)f(\Theta|D)\,d\Theta$$

thereby capturing and systematically incorporating sources of uncertainty in the modelling procedure. Specialised methods for estimation using Bayesian inference have evolved. Perhaps the most important is Monte Carlo Markov chain (MCMC, see Gamerman and Lopes, 2006). This is a simulation approach which generates (after suitable burn-in) a serially correlated sample from the full joint posterior $f(\Theta|D)$. Notably, MCMC avoids the need to integrate the first (integral) term on the right hand side of Eq. (4) numerically. Bayesian inference is fast becoming the default tool in academic statistics, and spreading into business and industry. Gibbs sampling from full conditionals provides slick computation, and the Metropolis–Hastings algorithm a general-purpose sampler (see Gamerman and Lopes, 2006). Reversible jump MCMC (Green, 1995) allows model complexity itself to be inferred.

There are difficulties however with Bayesian inference. Fundamentally, it is subjective. Different modellers will specify different priors and therefore obtain different inferences from the same data. Specifying prior information as a (multivariate) density is problematic. Estimation using MCMC is a specialist task in general. Estimation in high dimensions can become prohibitively complex computationally. All these areas are actively researched in the statistics community. Despite these difficulties, the appeal of Bayesian inference is in the authors' opinion overwhelming.

There is already a large literature on Bayesian inference relating to extremes and metocean applications. Coles and Powell (1996), Coles (2001) and Beirlant et al. (2004) review Bayesian methods in EV analysis. Scotto and Guedes-Soares (2007) outline a Bayesian inference for parameter estimation in EV analysis using the Metropolis-Hastings algorithm implemented in an appropriate Markov Chain Monte Carlo scheme. MacDonald et al. (2011) develop a flexible mixture model for extreme values. Wadsworth et al. (2010) and Reeve et al. (2012) adopt a Bayesian approach to estimating appropriate measurement scales for EV analysis.

7. Discussion

Extreme value analysis is inherently different to other empirical modelling, in that estimation of tails (rather than bodies) of distributions, and extrapolation beyond the sample (rather than interpolation within) is demanded. Intuition accumulated from other areas of empirical modelling can be misleading. This review outlines areas of opportunity for improved extreme value modelling in ocean engineering. Four main areas are identified: (1) improved understanding and use of basic EV techniques and tools, (2) modelling with covariates, (3) modelling dependence including spatial modelling, and (4) Bayesian inference. These areas have received considerable attention for statisticians, and are available for application in marine design.

Incorporation of covariate effects will improve model fits in general, and reduce modelling uncertainty. Suitable tools (e.g. based on spline methods, random fields, spatial processes) have been developed. Regarding modelling of multivariate extremes, effective modelling will improve the specification of design conditions (e.g. for H_S and associated values for other variables). Approaches such as the conditional model of Heffernan and Tawn (2004) are easily implemented, and provide generalisations of methods such as I-FORM (e.g. Winterstein et al., 1993). Modelling spatial dependence well will provide single EV models applicable to complete ocean basins, avoiding the need for site pooling. Once the model is established, the metocean engineer may not ever need to perform EV analysis in principle; at the very least the basin model would provide a reliable first estimate for establishing site-specific criteria. Modelling of spatial extremes is an active area of research in the statistics community. Bayesian inference provides a consistent framework for statistical modelling which is rapidly becoming the standard approach in academia.

There is considerable academic software available for extreme value modelling, mostly written in the R language. These packages are available (from CRAN) for general academic use, but not for commercial purposes. Modules are generally written to a reasonable programming standard, have received considerable attention for statisticians. Stephenson and Gilleland (2006) provide an overview of existing software for EV analysis, summarising that software currently available for the analysis of extreme values goes some way towards allowing practitioners to perform complicated tasks relatively easily. They note a confusing mixture of different methods and techniques and call for more coherent software development. Not one of the packages has been developed or tailored to metocean requirements (with the possible exception of WAFO). Packages such as TEXMEX (see CRAN) and VGAM have elements which are particularly useful for modern metocean applications.

The potential for applying modern methods of extreme value analysis to ocean engineering tasks is clear. The mathematical foundations of modern extreme value theory are relatively straight forward and intuitive. Indeed, many of the ideas developed by statisticians (e.g. max-stable process, extremal coefficient) have a natural and useful interpretation in the metocean context. The hurdle to wider adoption of such methods lies in a combination of unfamiliar but not impenetrable mathematical terminology, unfamiliar academic statistical software and decades of engineering practice and codification based on an empirical mix of well-established but more elementary and limited approaches. Awareness in the ocean engineering community, leading to careful consideration and incorporation of the effects of sample size, measurement scale, threshold, covariates, extremal dependence – all sources of modelling uncertainty and implications of modelling choices made – will lead eventually to better marine design.

Acknowledgments

The authors gratefully acknowledge discussions with John Heidemann, David Randell and Michael Vogel at Shell, with Eric Gilleland at NCAR, and with Ross Towe at Lancaster University.

We also thank two reviewers for thorough and thoughtful comments which improved initial drafts considerably.

Appendix A. Univariate extremes

This section attempts to provide informal motivation for extreme value thinking useful to the ocean engineer, starting with a motivation based on the distribution of the maximum of a sample (A.1), then introducing the concept of max-stability (A.2) and the generalised extreme value (GEV) distribution (A.3) suitable for modelling block maxima. A.4 illustrates the duality between the GEV distribution for block maxima and the generalised Pareto (GP) distribution for peaks over threshold (A.5). A.5 also provides a fuller motivation for the result due to Pickands (1975). A.6 outlines the use of the Poisson process model in extreme value analysis, and A.7 introduces the extremal index used to characterise the extent of serial correlation or clustering in a time series of extremes.

A.1. Motivation

Suppose $\{X_i\}_{i=1}^n$ are independent, identically distributed random variables with cumulative distribution function F with upper end point x_+, defined as the smallest value of x for which $F(x)=1$. Then

$$\Pr(\max(X_1,X_2,\ldots,X_n) \le x) = F^n(x)$$

For $x < x_+$, as $n \to \infty$, $F^n(x) \to 0$. That is, the limiting distribution of the maximum is degenerate, a point mass at x_+. A degenerate limiting form is not useful for empirical modelling of maxima data.

Therefore to model the tail behaviour of F, we choose to rescale the random variables. Consider $Y_n = a_n^{-1}(\max(X_1,X_2,\ldots,X_n) - b_n)$ for (sequences of) constants $a_n(>0)$ and $b_n \in (-\infty,\infty)$. Can we find values of a_n,b_n such that $Y_n \to Y$ as $n \to \infty$, such that Y no longer has a degenerate distribution? By definition

$$\begin{aligned}\Pr(Y_n \le y) &= F^n(b_n + a_n y)\\ &= \left(1 - \frac{n(1-F(b_n+a_n y))}{n}\right)^n\\ &\approx \exp(-n(1-F(b_n+a_n y))) \text{ for large } n\end{aligned}$$

using $(1+x/n)^n \approx \exp(x)$ for large n in the last line. The corresponding limit as $n \to \infty$ exists if and only if $\lim_{n\to\infty} n\ (1-F(b_n+a_n y))$ exists.

A.2. Max-stable distributions

Constants a_n and b_n can be found for most continuous distributions F. Remarkably, in all these cases, the limiting distribution is the same. We say that these limiting distributions are max-stable. The form of the limiting distribution is

$$n(1-F(b_n+a_n y)) \to \left(1+\xi\frac{y-\mu}{\sigma}\right)^{-1/\xi} \quad \text{as } n\to\infty \quad \text{for } \xi \ne 0$$
$$\left(\to \exp\left(-\frac{y-\mu}{\sigma}\right) \quad \text{when } \xi = 0\right)$$

for location parameter $\mu \in (-\infty,\infty)$, scale parameter $\sigma > 0$ and shape parameter $\xi \in (-\infty,\infty)$, with $y > \mu$ and $1+\xi(y-\mu)/\sigma > 0$. If ξ is negative, the limiting distribution has a finite upper end-point, $y_+ = \mu - \sigma/\xi$.

A.3. Generalised extreme value distribution

Combining the results in A.1 and A.2, the unique limiting distribution for maxima from max-stable distributions is the generalised extreme value distribution (GEV) with cumulative distribution function G:

$$G(x) = \exp\left(-\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-1/\xi}\right) \quad \text{for } \xi \ne 0$$

with the obvious result when $\xi = 0$ (see A.2), for location $\mu \in (-\infty,\infty)$, scale $\sigma > 0$ and shape $\xi \in (-\infty,\infty)$, with $1+\xi(x-\mu)/\sigma > 0$. Note also that maxima of max-stable distributions are also max-stable, so that $G^k(x) = G(ax+b)$ for $a>0, b \in (-\infty,\infty)$ and $k=1,2,\ldots,$ and that distributions F_1 and F_2 are said to be of the same type (or to be equivalent) if $F_2(x) = F_1(ax+b)$.

A.3.1. Domain of attraction

Any max-stable distribution lies within the domain of attraction (DOA) of the GEV for some ξ. For example, the classical Weibull distribution converges to the GEV with $\xi = 0$. That is, for cumulative distribution function $F(x) = 1-\exp(-cx^\tau)$, some algebra show that appropriate normalising constants are $a_n = (c\tau)^{-1}(c^{-1}\log n)^{(1/\tau)-1}$ and $b_n = (c^{-1}\log n)^{1/\tau}$ to leading order for large n (see, e.g. Embrechts et al., 2003).

A.3.2. Return values from GEV

We are interested in estimating the return value corresponding to a period $P = 1/p$ (typically longer than 1 year). Formally we define this return value, x_p, as the value that will be exceeded with probability p annually. Informally, this is also close to the most probable maximum for a return period of $1/p$ years. If annual maxima are GEV distributed with parameters ξ, σ and μ, the return period is defined by

$$1-p = \exp\left(-\left(1+\xi\frac{x_p-\mu}{\sigma}\right)^{-1/\xi}\right) \quad \text{for } \xi \ne 0$$

so that

$$x_p = \mu - \frac{\sigma}{\xi}(1-(-\log(1-p))^{-\xi})$$

In practice, for a sample $\{x_i\}_{i=1}^n$ corresponding to time period of exactly one year, we estimate the GEV model obtaining parameter estimates $\hat{\xi}$, $\hat{\sigma}$ and $\hat{\mu}$ corresponding to annual occurrence. We can use these as plug-in estimates in the equation above to estimate the return value.

In reality, we rarely fit the GEV distribution directly to maxima, since there are alternative approaches which make use all occurrences exceeding a threshold (which are therefore more efficient).

A.4. Correspondence between maxima and threshold exceedances

If X is a max-stable random variable with distribution function F, the following approximate argument provides the form of the distribution of threshold exceedances. For a sample $\{x_i\}_{i=1}^n$ of independent draws from F, since draws are independent we have

$$\Pr(\text{maximum of sample} \le x) = F^n(x)$$

Provided n is large enough, the distribution of the maximum is well approximated by the GEV

$$F^n(x) \approx \exp(-(1+\xi(x-\mu)/\sigma)^{-1/\xi}) \quad \text{for } \xi \ne 0$$

Taking logarithms

$$n \log F(x) \approx -(1+\xi(x-\mu)/\sigma)^{-1/\xi}$$

Then a Taylor expansion of $\log F(x)$ about 1 gives $\log F(x) \approx -(1-F(x))$ so that

$$\Pr(X>x)=1-F(x)\approx \frac{1}{n}(1+\xi(x-\mu)/\sigma)^{-1/\xi}$$

Simple rearrangement then gives

$$\Pr(X>x|X>u)=(1-F(x))/(1-F(u))\approx(1+\xi(x-u)/\tilde{\sigma})^{-1/\xi}$$

where $\tilde{\sigma}=\sigma+\xi(u-\mu)$. The last line is the generalised Pareto (GP) distribution. Note that the shape parameter is common to GP and GEV.

Equivalently, for independent Poisson-distributed occurrences of threshold exceedances with GP-distributed sizes, we can also show informally that the distribution of the maximum (corresponding to some period) is GEV-distributed. Recalling that $\Pr(N=k)=(\lambda^k/k!)\exp(-\lambda),\ k=0,1,2,\ldots$ for Poisson-distributed random variable N with rate $\lambda>0$ in the second line in the following equation:

$$\begin{aligned}\Pr(\text{maximum in period}\le x) &= \sum_{k=0}^{\infty}\Pr(k \text{ storms in period})F^k(x)\\ &= \sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\exp(-\lambda)F^k(x)\\ &= \exp(-\lambda(1-F(x)))\\ &= \exp\left(-\lambda\left(1+\xi\frac{x-u}{\tilde{\sigma}}\right)^{-1/\xi}\right)\end{aligned}$$

where λ is the expected number of exceedances in the period, and $\tilde{\sigma}=\sigma+\xi(u-\mu)$. If we choose λ to be $(1+\xi(u-\mu)/\sigma)^{-1/\xi}$ (without loss of generality), we recover the original GEV form.

A.5. Generalised Pareto distribution

The generalised Pareto distribution has cumulative distribution function H given by

$$H(x)=1-\left(1+\xi\frac{x-u}{\tilde{\sigma}}\right)^{-1/\xi}\quad \text{for } \xi\neq 0$$

for threshold $u\in(-\infty,\infty)$, scale $\tilde{\sigma}>0$ and shape $\xi\in(-\infty,\infty)$ with $x>u$ and $1+\xi(x-u)/\tilde{\sigma}>0$, with the right hand side taken to be $1-\exp(-(x-u)/\tilde{\sigma})$ when $\xi=0$.

A.5.1. GP tail approximation

Consider a random variable X with continuous cumulative density function F with upper end point x_+. Pickands (1975) showed that the conditional distribution of X given $X>u$ for some threshold u can be approximated by the generalised Pareto distribution defined above (and see also, e.g. Falk et al., 2004).

Restricting discussion to max-stable limiting distributions, continuing the development in A.2, we see after some algebra that

$$\frac{1-F(b_n+a_n y)}{1-F(b_n)}\to\left(1+\xi\frac{y}{\sigma^*}\right)^{-1/\xi}\ \text{as } n\to\infty\quad \text{for } \xi\neq 0$$

where $\sigma^*=\sigma-\xi\mu$ and $y>0$. Setting $b_n=u$ and $a_n=a(u)>0$, then as $n\to\infty$ (or equivalently as $u\to x_+$)

$$F_u(u+a(u)y)=\frac{F(u+a(u)y)-F(u)}{1-F(u)}=1-\frac{1-F(u+a(u)y)}{1-F(u)}\to 1-\left(1+\xi\frac{y}{\sigma^*}\right)^{-1/\xi}$$

for any $y>0$. Setting $x=u+a(u)y$ we see that

$$F_u(x)\approx 1-\left(1+\xi\frac{x-u}{\tilde{\sigma}}\right)^{-1/\xi}$$

where $\tilde{\sigma}=a(u)\sigma^*$, for exceedances x of high threshold u. In theory, the function $a(u)$ is given by the reciprocal hazard function for F evaluated at u (see, e.g. Reeve et al., 2012). In practice, ξ and $\tilde{\sigma}$ are estimated empirically from data.

A.5.2. Return values from GP

Assuming a GP model with parameters u, $\tilde{\sigma}$ and ξ is appropriate for threshold exceedances of X, the $P=1/p$-year return value x_p is defined by

$$1-p=\exp\left(-\lambda\left(1+\xi\frac{x_p-u}{\tilde{\sigma}}\right)^{-1/\xi}\right)\quad \text{for } \xi\neq 0$$

with the obvious result for $\xi=0$, where λ is the expected number of exceedances per annum.

A.6. Poisson process

Suppose we count the number of events $N_i=0,1,2,\ldots$ occurring on each of $i=1,2,\ldots,n$ intervals $\{(a_i,b_i]\}_{i=1}^n$ on an interval $(0,T]$ of the real line, with $b_i\le a_{i+1}$. The set $\{N_i\}_{i=1}^n$ follows a Poisson process if

$$\Pr(N_i=n_i,\ i=1,2,\ldots,n)=\prod_{i=1}^{n}\frac{\Lambda_i^{n_i}}{n_i!}e^{-\Lambda_i}$$

where $\Lambda_i=\int_{a_i}^{b_i}\lambda(t)dt$ and $\lambda(t)>0$ is the rate of the Poisson process at $t\in(0,T]$. The number of points in each interval is Poisson-distributed and the numbers of points in disjoint intervals are independently distributed. The probability of obtaining exactly one event in each interval $\{(a_i,b_i]\}_{i=1}^n$ is therefore $\prod_{i=1}^n\Lambda_i e^{-\Lambda_i}$, and the probability of obtaining no events elsewhere is $e^{-(\Lambda-\sum_{i=1}^n\Lambda_i)}$, where $\Lambda=\int_0^T\lambda(t)\,dt$. Since the corresponding events are independent, the probability of obtaining exactly n events on $(0,T]$ at points $\{t_i\}_{i=1}^n$ (each with the corresponding interval) is

$$\Pr(t_1,t_2,\ldots,t_n)=\prod_{i=1}^{n}\Lambda_i e^{-\Lambda_i}\times e^{-\Lambda-\sum_{i=1}^n\Lambda_i}=e^{-\Lambda}\prod_{i=1}^{n}\Lambda_i$$

Now assume that the intervals $(a_i,b_i]$ are of constant width Δ and divide the expression above by Δ^n. For sufficiently small Δ we see that $\Lambda_i\approx\lambda(t_i)\Delta$. In the limit as $\Delta\to 0$, we obtain the density (or likelihood)

$$f(t_1,t_2,\ldots,t_n)=e^{-\Lambda}\prod_{i=1}^{n}\lambda_i^n$$

But the probability of obtaining exactly n events in $(0,T]$ is also given by

$$\Pr(N=N_{(0,T]}=n)=\frac{\Lambda^n}{n!}e^{-\Lambda}$$

and hence the conditional density $f(t_1,t_2,\ldots,t_n|N=n)$ becomes

$$f(t_1,t_2,\ldots,t_n|N=n)=n!\prod_{i=1}^{n}\frac{\lambda(t_i)}{\Lambda}$$

which can be interpreted as the density of n independently distributed distinguishable points, each with density

$$f(t|N=n)=\frac{\lambda(t)}{\Lambda}$$

This is the non-homogeneous Poisson process, a useful general basis for many spatio-temporal models for different choices of rate and domain. It can be generalised to higher dimensions. In particular, if we assume an intensity $\lambda(x,t)$ defined on $[u,\infty)\times(0,1]$

$$\lambda(x,t)=\frac{1}{\sigma}\left(1+\xi\frac{u-\mu}{\sigma}\right)^{-1/\xi-1}\quad \text{for } \xi\neq 0$$

with the obvious result for $\xi=0$, for location $\mu\in(-\infty,\infty)$, scale $\sigma>0$ and shape $\xi\in(-\infty,\infty)$, with $x>\mu$ and $1+\xi(x-\mu)/\sigma>0$, the conditional distribution $\Pr(X\le x|T=t)$ has the generalised Pareto form given in Section A.5. Hence, with this choice of intensity, we can use the Poisson process for maximum likelihood inference in extreme value analysis.

A.7. Estimating time series dependence using the extremal index

Suppose $X_t^F, t=1,2,3,\ldots$ is a time series of random variables with unit Fréchet marginal distribution, $\Pr(X_t^F \le f) = e^{-1/f}, f>0$. If the random variables are independent, the scaled maximum of $n=1,2,3,\ldots$ consecutive points, $M_n^F = \max\{X_i^F, X_{i+1}^F, \ldots, X_{i+n-1}^F\}/n$ for any $i=1,2,3,\ldots$ will also satisfy $\Pr(M_n^F \le f) = e^{-1/f}$. With appropriate (and plausible) assumptions (see, e.g. Ledford and Tawn, 2003) an analogous asymptotic result emerges for dependent random variables:

$$\Pr(M_n^F \le f) = e^{-\theta/f}$$

for $f>0$, where $\theta \in [0,1]$ is known as the extremal index. The extremal index θ (see Leadbetter et al., 1983 and Smith and Weissman, 1994) is a measure of within-series extremal dependence. Independent data have $\theta=1$, although $\theta=1$ does not necessarily imply independence in the full data set, merely independence between extremes. The closer θ is to zero, the stronger the dependence between extremes. For series in which the extremes are not independent, the extreme points will tend to occur in independent clusters; the mean cluster size is given by the reciprocal θ^{-1}. The above equation provides a straight forward estimation procedure for θ from a sample of time series: first transform to unit Fréchet marginal, then calculate scaled cluster maxima for different cluster sizes n, then estimate the value of θ for exceedances of threshold u. For sufficiently large n and u, the value of θ should stabilise. See Ferro and Segers (2003) and Ledford and Tawn (2003) for refinements and more recent developments.

Appendix B. Multivariate extremes

The following is an informal outline of some concepts in multivariate EV analysis possibly useful to the practising ocean engineer. We start by describing different forms of multivariate extremal dependence (B.1), discussing the different forms of multivariate extremal dependence and how to estimate them. We then look at spatial extremes (B.3), motivating the field of max-stable processes using the so called Smith model, and outlining the extremal coefficient useful to quantify the extent of spatial dependence.

B.1. Extremal dependence

The bivariate random variable (X,Y) with common marginal distribution is said to be asymptotically independent if $\lim_{x\to\infty} \Pr(X>x|Y>x)=0$ and asymptotically dependent if $\lim_{x\to\infty}\Pr(X>x|Y>x)>0$. In multivariate EV analysis it is usual to consider the properties of quantities such as the joint survivor function $\Pr(X>x, Y>x)$, and the conditional probability $\Pr(X>x|Y>x)$, for large x.

Let (X_F, Y_F) be a bivariate random variable with unit Fréchet marginal distributions (i.e. $\Pr(X_F \le f) = e^{-1/f}, f>0$). In the special case that X_F and Y_F are independent, then $\Pr(X_F>f|Y_F>f) = \Pr(X_F>f) \to 0$ as $f\to\infty$. Hence X_F and Y_F and also asymptotically independent. In the case $X_F = Y_F$, then $\Pr(X_F>f|\ Y_F>f)=1$ so that X_F and Y_F are asymptotically dependent. In general, we can assume that $\Pr(X_F>f, Y_F>f)$ is regularly varying at infinity with index $-1/\eta$, $\eta \in (0,1]$ (see Bingham et al., 1987). That is

$$\lim_{f\to\infty} \frac{\Pr(X_F>sf, Y_F>sf)}{\Pr(X_F>f, Y_F>f)} = s^{-1/\eta} \text{ for some fixed } s>0$$

For sufficiently large values of f, this equation suggests

$$\Pr(X_F>sf, Y_F>sf) \approx s^{-1/\eta}\Pr(X_F>f, Y_F>f)$$

If we transform variables from Fréchet (X_F, Y_F) to Gumbel (X_G, Y_G) marginal distributions, such that $\Pr(X_G<g) = \exp(-e^{-g}) = \Pr(X_F<e^g), g\in(-\infty,\infty)$, we obtain for $t>0$

$$\begin{aligned}\Pr(X_G>g+t, Y_G>g+t) &= \Pr(X_F>e^{g+t}, Y_F>e^{g+t})\\ &\approx e^{-t/\eta}\Pr(X_F>e^g, Y_F>e^g) \quad \text{for large } g\\ &= e^{-t/\eta}\Pr(X_G>g, Y_G>g)\end{aligned}$$

The index η provides a measure of the dependence between the marginal tails, effectively the rate of decay of the joint survivor function for large arguments. It is clear, if we approach multivariate EV modelling from the perspective of multivariate regular variation, that estimation η is important.

Ledford and Tawn (1996) introduce the model $\Pr(X_F>f, Y_F>f) = \ell(f)f^{-1/\eta}$ where $\ell(f)$ is a slowly varying function (i.e. $\lim_{f\to\infty}\ell(sf)/\ell(f)=1$), applicable to a broad range of distributions with unit Fréchet marginal variables, with η referred to as the coefficient of tail dependence. Under this model

$$\begin{aligned}\Pr(X_F>f|Y_F>f) &= \frac{\Pr(X_F>f, Y_F>f)}{\Pr(Y_F>f)}\\ &= \ell(f)f^{-1/\eta}(1-e^{-1/f})^{-1}\\ &\approx \ell(f)f^{1-1/\eta} \quad \text{for large } f\\ &\approx \ell(f)\Pr(Y_F>f)^{(1/\eta)-1} \quad \text{for large } f\end{aligned}$$

Thus at $\eta=1$ (and $\lim_{f\to\infty}\ell(f)\neq 0$), X_F and Y_F are asymptotically dependent. Otherwise, for $\eta<1$ (or $\eta=1$ and $\lim_{f\to\infty}\ell(f)=0$), X_F and Y_F are asymptotically independent. In practice, η can be estimated by noting, for the minimum L_F of X_F and Y_F, that $\Pr(L_F>f) = \Pr(X_F>f, Y_F>f) = \ell(f)f^{-1/\eta}$. Hence, conditionally, for $s>0$

$$\Pr(L_F>f+s|L_F>f) = \frac{\ell(f+s)}{\ell(f)}\left(1+\frac{s}{f}\right)^{-1/\eta} \approx \left(1+\frac{s}{f}\right)^{-1/\eta} \text{ for large } f$$

The final expression takes generalised Pareto (GP) form; the shape parameter from a GP fit to L_F (with threshold f) therefore yields an estimate for η.

Ledford and Tawn (1997) introduce a more flexible asymptotic form for the joint survivor function. For modelling purposes, this reduces to assuming $\Pr(X_F>x, Y_F>y) \approx \ell(x,y)x^{-c_X}y^{-c_Y}$, where $c_X+c_Y=1/\eta$ with $x,y>0$ again simultaneously large. Now $\ell(x,y)$ is a bivariate slowly varying function with limit $\phi(x,y)$ (i.e. $\lim_{f\to\infty}\ell(fx,fy)/\ell(f,f)=\phi(x,y)$, and $\phi(fx,fy)=\phi(x,y)$ for all $f>0$). Under this model again, for $t>0$

$$\begin{aligned}\Pr(X_G>x+t, Y_G>y+t) &= \Pr(X_F>e^{x+t}, Y_F>e^{y+t})\\ &\approx \ell(e^xe^t, e^ye^t)e^{-t(c_X+c_Y)}e^{-xc_X}e^{-yc_Y} \quad \text{for large } x \text{ and } y\\ &= \frac{\ell(e^xe^t, e^ye^t)}{\ell(e^x, e^y)}e^{-t/\eta}(\ell(e^x, e^y)e^{-xc_X}e^{-yc_Y})\\ &\approx e^{-t/\eta}\Pr(X_F>e^x, Y_F>e^y) \quad \text{for large } x \text{ and } y\\ &= e^{-t/\eta}\Pr(X_G>x, Y_G>y)\end{aligned}$$

Summary statistics such as the coefficient of tail dependence η characterise the strength of extremal association concisely. In application, if an incorrect assumption is made on the form of the extremal dependence, then any inferences drawn from a model will be invalid. For example, we might wrongly assume two variables to be asymptotically dependent, when in fact they are asymptotically independent; a model would then overestimate the probability of joint extreme events.

B.2. Estimating extremal dependence using χ and $\bar{\chi}$

Coles et al. (1999) introduce the diagnostics pair χ , $\bar{\chi}$ to summarise the extremal dependence of an arbitrary pair of random variables. For random variables X and Y with common marginal distribution we have

$$\chi = \lim_{x\to\infty} \Pr(X > x | Y > x)$$

and

$$\overline{\chi} = 2\eta - 1$$

In practice, χ is easily estimated as the limit as x increases of $\chi(x) = \Pr(X_G > x | Y_G > x)$ for threshold x following transformation to Gumbel (or other suitable) marginal scale; $\overline{\chi}$ is best estimated from η. χ quantifies asymptotic dependence: $\chi = 0$ for asymptotic independence and $\chi = c$ for some $c \in [0,1]$ for asymptotic dependence. As c increases, so does the extent of asymptotic dependence. The value of χ is also related to the exponent measure (see B.3.2). Conversely, $\overline{\chi}$ quantifies asymptotic independence: $\overline{\chi} = 1$ for asymptotically dependence, otherwise $\overline{\chi} = c$ for some $c \in (-1,1)$ for asymptotic independence, with $c = 0$ indicating perfect independence. If $c > 0$ we say that the association between variables is positive, and if $c < 0$ the association is negative. Eastoe et al. (accepted for publication) estimate $\chi, \overline{\chi}$ and η for the 8 m array at the USACE Field Research Facility. The conditional extremes model (see Section 5) is sufficiently flexible to accommodate different forms of asymptotic dependence and independence whereas methods requiring a max-stability assumption are not.

For random variables X and Y with joint cumulative distribution function F and marginal cumulative distribution functions F_1 and F_2 respectively, Coles et al. (1999) introduce useful estimates $\chi(u)$ and $\overline{\chi}(u)$ with $u \in (0,1)$ for χ and $\overline{\chi}$ respectively based on the bivariate copula function C and its corresponding copula survivor function $\overline{C}$:

$$\chi(u) = 2 - \frac{\log_e(C(u,u))}{\log_e(u)} \quad \text{for } 0 < u < 1$$

and

$$\overline{\chi}(u) = \frac{2\log_e(1-u)}{\log_e(\overline{C}(u,u))} \quad \text{for } 0 < u < 1$$

where

$$C(u_1,u_2) = \Pr(X \le F_1^{-1}(u_1), Y \le F_2^{-1}(u_2)) = F(F_1^{-1}(u_1), F_2^{-1}(u_2))$$

and

$$\overline{C}(u_1,u_2) = \Pr(X > F_1^{-1}(u_1), Y > F_2^{-1}(u_2)) = 1 - u_1 - u_2 + C(u_1,u_2)$$

such that $\chi(u) \to \chi$ and $\overline{\chi}(u) \to \overline{\chi}$ as $u \to 1$. Poon et al. (2003) provide alternative estimates for χ and $\overline{\chi}$ under an appropriate tail model within a censored likelihood framework.

B.3. Spatial extremes

This section provides an informal introduction to the rapidly growing field of spatial extreme value modelling of componentwise maxima, using the so called Smith model as motivation.

B.3.1. Max-stable process

Following Smith (1990), suppose we have a region S of 2-D space within which storm centres can occur, and a (frequency) measure v of the distribution of storms over S. Suppose that $h_i > 0$ represents the magnitude of a storm indexed by $i = 1,2,3,\ldots$, and that $h_i f(s_i, r)$ represents the value of significant wave height (H_S) at location r (from a set of locations $r \in$ index set R) for a storm of size h_i centred at s_i; the function f represents the distribution of H_S corresponding to each storm. We are interested in the maximum value Z_r at location r of H_S over independent storms:

$$Z_r = \max_i\{h_i f(s_i, r)\}, \quad r \in R$$

We know that Poisson processes have proved invaluable in describing extremes in 1-D. We can construct a Poisson process to describe the current situation also, reducing to the existing univariate EV results marginally, using the following approach. We suppose that events $\{h_i, s_i\}$ form a Poisson process with intensity measure $h^{-2}\, dh\, dv$ (where the choice of $h^{-2}\, dh$ guarantees that all marginal distributions will be unit Fréchet). We also assume (a convenient normalisation only) that $\int_S f(s,r)\, dv = 1$.

Now consider the event $(Z_1 \le z_1, Z_2 \le z_2, \ldots)$, i.e. that the maximum values of H_S jointly across all locations due to all storms, do not exceed the specified values:

$$\begin{aligned}
\Pr(Z_1 \le z_1, Z_2 \le z_2, \ldots) &= \Pr(Z_r \le z_r, \text{ for all } r) \\
&= \Pr\left(\max_i\{h_i f(s_i,r)\} \le z_r\right), \text{ for all } r) \\
&= \Pr(h_i f(s_i,r) \le z_r, i = 1,2,\ldots, \text{for all } r) \\
&= \exp\left(-\int_S \int_0^\infty \mathbf{I}\left(h \max_r\{z_r^{-1} f(s,r)\} > 1\right) h^{-2}\, dh\, dv\right) \\
&= \exp\left(-\int_S \max_r\{z_r^{-1} f(s,r)\}\, dv\right)
\end{aligned}$$

where $\mathbf{I}$ is the indicator function ($=1$ when its argument is true and $=0$ otherwise), and the fourth line derives from the fact that for any Poisson process (e.g. with respect to variable x with intensity $\lambda(x)$)

$$\Pr(\text{time of last occurrence} \le x^*) = \exp\left(-\int_{x^*}^{\infty} \lambda(x)\, dx\right)$$

Now we also see, because of the normalising assumption $\int_S f(s,r)\, dv = 1$, that marginally

$$\Pr(Z_r \le z_r) = \exp(-1/z_r) \text{ for any particular } r \in R$$

That is, all marginal distributions are unit Fréchet. The remaining (dependence) structure of the process is governed by our specification of f, as outlined in Section 5.5.

de Haan (1984) provides a constructive representation of max-stable processes which includes the Smith model and others, including that of Schlather (2002). de Haan and Pickands (1986) study the minima of min-stable processes, and Wadsworth and Tawn (2012) show how the constructive representation can also be used for min-stable and inverted max-stable processes which are asymptotically independent. Wadsworth and Tawn (2012) further suggest a hybrid spatial dependence model as the weighted maximum of max-stable and inverted max-stable processes incorporating both asymptotic dependence and asymptotic independence. Further, Huser and Davison (2012) show that a max-stable model for spatial componentwise maxima also provides a model for spatial exceedences of high thresholds. Using a censored likelihood approach, they illustrate application of spatial extremes models to peaks over threshold. Wadsworth and Tawn (2012) include a similar illustration. These developments suggest a route to large scale modelling of spatial peaks over threshold with a range of extremal dependence structures.

B.3.2. Exponent measure

It is convenient to express $\Pr(Z_r \le z_r, r = 1,2,\ldots)$ in the following form:

$$\Pr(Z_1 \le z_1, Z_2 \le z_2, \ldots) = \exp(-V(z_1, z_2, \ldots))$$

where $V(z_1, z_2, \ldots)$ is known as the exponent measure. It can be shown (e.g. Beirlant et al., 2004) that max-stability constrains V to be homogeneous of order -1, that is

$$V(rz_1, rz_2, \ldots) = r^{-1} V(z_1, z_2, \ldots)$$

Only distributions which correspond to a V which is homogeneous order -1 form valid max-stable processes. In particular, not all copula models are max-stable. Because of their homogeneity property, max-stable processes can also be motivated from the perspective of a spectral representation (e.g. Schlather, 2002).

B.3.3. Extremal coefficient and site pooling

For exponent measure V, the extremal coefficient $\theta_D = V(1,1,\ldots)$, i.e. the value of the exponent measure at $(1,1,\ldots)$, is a useful summary of the extent of spatial dependence, and can be thought of as the effective number of sites being pooled in a pooling analysis. For q locations, the extremal coefficient is defined by the following equation:

$$\Pr(Z_1 \le z, Z_2 \le z, \ldots, Z_q \le z) = \exp(-V(1,1,\ldots)/z) = \exp(-\theta_D/z)$$

for $z > 0$, $\theta_D \in [1,q]$ since V is homogeneous order -1. Further, since the marginal distribution of any particular Z is Fréchet, we see that

$$\exp(-\theta_D/z) = \{\exp(-1/z)\}^{\theta_D} = \Pr(Z_r \le z)^{\theta_D}$$

providing the pooling interpretation of θ_D. The extremal coefficient can be estimated empirically from data (once transformed to unit Fréchet marginals), summarising the degree of dependence and hence the effectiveness of any site pooling strategy. For two locations, the extremal coefficient also has a natural interpretation in terms of conditional bivariate exceedances:

$$\lim_{z\to\infty} \Pr(Z_1 > z | Z_2 > z) = 2 - \theta_D$$

showing its relationship to the dependence diagnostic χ, namely $\theta_D = 2 - \chi$ (see B.1).

References

Anderson, C.W., 1990. Discussion of Davison, A.C., Smith, R.L. "Models for exceedances over high thresholds". Journal of the Royal Statistical Society: Series B 52, 393.

Anderson, C.W., Carter, D.J.T., Cotton, P.D., 2001. Wave Climate Variability and Impact on Offshore Design Extremes. Report commissioned from the University of Sheffield and Satellite Observing Systems for Shell International.

API, 2005. API Recommended Practice 2A-WSD (RP 2A-WSD), Recommended Practice for Planning, Designing and Constructing Fixed Offshore Platforms – Working Stress Design. API.

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

Bernier, N.B., Thompson, K.R., Ou, J., Ritchie, H., 2007. Mapping the return periods of extreme sea levels: allowing for short sea level records, seasonality, and climate change. Global and Planetary Change 57, 139–150.

Bingham, N.H., Goldie, C.M., Teugels, J.L., 1987. Regular Variation. Cambridge University Press.

Chandler, R.E., 2005. On the use of generalized linear models for interpreting climate variability. Environmetrics 16, 699–715.

Chandler, R.E., Bate, S., 2007. Inference for clustered data using the independence loglikelihood. Biometrika 94, 167–183.

Chavez-Demoulin, V., Davison, A.C., 2005. Generalized additive modelling of sample extremes. Journal of the Royal Statistical Society: Series C: Applied Statistics 54, 207.

Coles, S., 2001. An Introduction to Statistical Modelling of Extreme Values. Springer.

Coles, S., Heffernan, J., Tawn, J., 1999. Dependence measures for extreme value analyses. Extremes 2, 339–365.

Coles, S.G., Powell, E.A., 1996. Bayesian methods in extreme value modelling: a review and new developments. International Statistics Review 64, 119–136.

Coles, S.G., Simiu, E., 2003. Estimating uncertainty in the extreme value analysis of data generated by a hurricane simulation model. Journal of Engineering Mechanics 129, 1288.

Coles, S.G., Tawn, J.A., 1991. Modelling multivariate extreme events. Journal of the Royal Statistical Society: Series B 53, 377–392.

Coles, S.G., Walshaw, D., 1994. Directional modelling of extreme wind speeds. Applied Statistics 43, 139–157.

Cook, N.J., 1982. Towards better estimation of extreme winds. Journal of Wind Engineering and Industrial Aerodynamics 9, 295–323.

Cook, N.J., Harris, R.I., 2004. Exact and general ft1 penultimate distributions of extreme wind speeds drawn from tail-equivalent Weibull parents. Structural Safety 26, 391–420.

Cook, N.J., Harris, R.I., Whiting, R., 2003. Extreme wind speeds in mixed climated revisited. Journal of Wind Engineering and Industrial Aerodynamics 91, 403–422.

Cooley, D., Nychka, D., Naveau, P., 2007. Bayesian spatial modelling of extreme precipitation return levels. Journal of American Statistical Society 102, 824–840.

Cox, D.R., Reid, N., 2004. A note on pseudolikelihood constructed from marginal densities. Biometrika 91, 729–737.

CRAN. The Comprehensive R Archive Network. URL ⟨cran.r-project.org⟩.

Davis, R.A., Yau, C.Y., 2011. Comments on pairwise likelihood in time series models. Statistica Sinica, 255–277.

Davison, A.C., 2003. Statistical Models. Cambridge University Press.

Davison, A.C., Gholamrezaee, M.M., 2012. Geostatistics of extremes. Proceedings of the Royal Society of London A 468, 581–608.

Davison, A.C., Hinkley, D.A., 1997. Bootstrap Methods and their Application (Cambridge Series in Statistical and Probabilistic Mathematics). Cambridge University Press, Cambridge, UK.

Davison, A.C., Padoan, S.A., Ribatet, M., 2012. Statistical modelling of spatial extremes. Statistical Science 27, 161–186.

Davison, A.C., Smith, R.L., 1990. Models for exceedances over high thresholds. Journal of the Royal Statistical Society: Series B 52, 393.

de Haan, L., 1984. A spectral representation for max-stable processes. Annals of Probability 12, 1194–1204.

de Haan, L., Pickands, J., 1986. Stationary min-stable stochastic processes. Probability Theory and Related Fields 72, 477–492.

Dong, Sheng, Wang, Nannan, Liu, Wei, Guedes Soares, C., 2013. Bivariate maximum entropy distribution of significant wave height and peak period. Ocean Engineering 59, 86–99.

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 Engineering, accepted for publication, draft at ⟨www.lancs.ac.uk/~jonathan⟩.

Eastoe, E.F., Tawn, J.A., 2009. Modelling non-stationary extremes with application to surface level ozone. Applied Statistics 58, 22–45.

Elsinghorst, C., Groeneboom, P., Jonathan, P., Smulders, L., Taylor, P.H., 1998. Extreme value analysis of North Sea storm severity. Journal of Offshore Mechanics Arctic Engineering 120, 177–183.

Embrechts, P., Klueppelberg, C., Mikosch, T., 2003. Modelling Extremal Events for Insurance and Finance. Springer-Verlag.

Erhardt, R.J., Smith, R.L., 2011. Approximate Bayesian Computing for Spatial Extremes. Available from ⟨http://arxiv.org/pdf/1109.4166.pdf⟩.

Ewans, K.C., Jonathan, P., 2008. The effect of directionality on Northern North Sea extreme wave design criteria. Journal of Offshore Mechanics Arctic Engineering 130, 10.

Ewans, K.C., Jonathan, P. Evaluating environmental joint extremes for the offshore industry. Journal of Marine Systems, accepted for publication. URL ⟨http://arxiv.org/abs/1211.1365⟩.

Falk, M., Husler, J., Reiss, R.D., 2004. Laws of Small Numbers: Extreme and Rare Events. Birkhauser, Basel, Switzerland.

Fawcett, L., Walshaw, D., 2007. Improved estimation for temporally clustered extremes. Environmetrics 18, 173–188.

Ferreira, J.A., Guedes-Soares, C., 2002. Modelling bivariate distributions of significant wave height and mean wave period. Applied Ocean Research 21, 31–45.

Ferro, C.A.T., Segers, J., 2003. Inference for clusters of extreme values. Journal of Royal Statistics Society B 65, 545–556.

Forristall, G.Z., 2004. On the use of directional wave criteria. Journal of Waterway Port Coastal and Ocean Engineering 130, 272–275.

Gabda, D., Towe, R., Wadsworth, J., Tawn, J., 2012. Discussion of statistical modeling of spatial extremes by A.C. Davison, S.A. Padoan, M. Ribatet. Statistical Science 27, 189–192.

Gamerman, D., Lopes, H.F., 2006. Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference. Taylor & Francis.

Gilleland, E., Nychka, D.W., Schneider, U., 2006. Hierarchical modelling for the environmental sciences: statistical methods and applications. In: Clark, J.S., Gelfand, A. (Eds.), Spatial Models for the Distribution of Extremes. Oxford University Press, New York, pp. 170–183. (Chapter 9).

Graham, C.G., 1981. Problems with the design statistical approach. In: Extreme Value Analysis of Directional Wind and Wave Data, International Conference on Wave and Wind Directionality, Paris, pp. 379–401.

Green, P.J., Silverman, B.W., 1994. Nonparametric Regression and Generalised Linear Models: a Roughness Penalty Approach. Chapman and Hall, London, UK.

Green, P.J., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika 82, 711–732.

Haver, S., 1985. Wave climate off northern Norway. Applied Ocean Research 7, 85–92.

Heffernan, J.E., Tawn, J.A., 2004. A conditional approach for multivariate extreme values. Journal of Royal Statistics Society B 66, 497.

Heideman, J.C., Mitchell, D.A., 2009. Grid point pooling in extreme value analysis of hurricane hindcast data. Journal of Waterway Port Coastal and Ocean Engineering 135, 31–38.

Hosking, J.R.M., 1990. L-moments: analysis and estimation of distributions using linear combinations of order statistics. Journal of Royal Statistics Society Series B 52, 105–124.

Huser, R., Davison, A.C., 2012. Space-time Modelling of Extreme Events. Draft at arxiv.org/abs/1201.3245, 2012, draft.

ISO, 2005. BS EN ISO 19901-1:2005, Petroleum and Natural Gas Industries – Specific Requirements for Offshore Structures – Part 1: Metocean Design and Operating Conditions British Standards Institute.

Joe, H., 1990. Families of min-stable multivariate exponential and multivariate extreme value distributions. Statistics & Probability Letters 9, 75–81.

Jonathan, P., Ewans, K., 2011a. Modelling the seasonality of extreme waves in the Gulf of Mexico. ASME Journal of Offshore Mechanics and Arctic Engineering 133, 021104.

Jonathan, P., Ewans, K., 2011b. A spatiodirectional model for extreme waves in the Gulf of Mexico. ASME Journal of Offshore Mechanics and Arctic Engineering 133, 011601.

Jonathan, P., Ewans, K.C., 2007a. The effect of directionality on extreme wave design criteria. Ocean Engineering 34, 1977–1994.

Jonathan, P., Ewans, K.C., 2007b. Uncertainties in extreme wave height estimates for hurricane dominated regions. Journal of Offshore Mechanics and Arctic Engineering 129, 300–305.

Jonathan, P., Ewans, K.C., 2008. On modelling seasonality of extreme waves. In: Proceedings of the 27th International Conference on Offshore Mechanics and Arctic Engineering, 4–8 June, Estoril, Portugal.

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 Engineering 35, 1211–1225.

Jonathan, P., Flynn, J., Ewans, K.C., 2009. Joint modelling of wave spectral parameters for extreme sea states. In: Proceedings ofn the 11th International Workshop on Wave Hindcasting and Forecasting, Halifax, Nova Scotia, Canada.

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

Jonathan, P., Ewans, K., Flynn, J. 2011c. On the estimation of ocean engineering design contours, In: OMAE2011-49886, Proceedings of the 30th International Conference on Offshore Mechanics and Arctic Engineering, June 19 – 24, 2011, Rotterdam, The Netherlands.

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

Kalbfleisch, J.G., 1985. Probability and Statistical Inference. Statistical Inference, vol. 2. Springer.

Koenker, R., 2005. Quantile Regression. Cambridge University Press.

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

Kyselý, J., Picek, J., Beranová, R., 2010. Estimating extremes in climate change simulations using the peaks-over-threshold method with a non-stationary threshold. Global and Planetary Change 72, 55–68.

Kysely, J., Gaal, L., Picek, J., 2011. Comparison of regional and at-site approaches to modelling probabilities of heavy precipitation. International Journal of Climatology 31, 1457–1472.

Leadbetter, M.R., 1991. On a basis for 'peaks over threshold' modeling. Statistics and Probability Letters 12, 357–362.

Leadbetter, M.R., Lindgren, G., Rootzen, H., 1983. Extremes and Related Properties of Random Sequences and Series.

Ledford, A.W., Tawn, J.A., 1996. Statistics for near independence in multivariate extreme values. Biometrika 83, 169–187.

Ledford, A.W., Tawn, J.A., 1997. Modelling dependence within joint tail regions. Journal of Royal Statistics Society Series B 59, 475–499.

Ledford, A.W., Tawn, J.A., 2003. Diagnostics for dependence within time series extremes. Journal of Royal Statistics Society Series B 65, 521–543.

Li, F., Bicknell, C., Lowry, R., Li, Y., 2012. A comparison of extreme wave analysis methods with 1994–2010 offshore Perth dataset. Coastal Engineering 69, 1–11.

Lindsay, B.G., 1988. Composite likelihood methods. Contemporary Mathematics 80, 221–239.

MacDonald, A., Scarrott, C.J., Lee, D., Darlow, B., Reale, M., Russell, G., 2011. A flexible extreme value mixture model. Computational Statistics and Data Analysis 55, 2137–2157.

Marx, B.D., Eilers, P.H.C., 1998. Direct generalised additive modelling with penalized likelihood. Computational Statistics and Data Analysis 28, 193–209.

Mendez, F.J., Menendez, M., Luceno, A., Medina, R., Graham, N.E., 2008. Seasonality and duration in extreme value distributions of significant wave height. Ocean Engineering 35, 131–138.

Mendez, F.J., Menendez, M., Luceno, A., Losada, I.J., 2006. Estimation of the long-term variability of extreme significant wave height using a time-dependent pot model. Journal of Geophysical Research 11, C07024.

Michelangeli, P.A., Vrac, M., Loukos, H., 2009. Probabilistic downscaling approaches: application to wind cumulative distribution functions. Geophysical Research Letters 36, 1–6, L11708.

Northrop, P., Jonathan, P., 2011. Threshold modelling of spatially-dependent non-stationary extremes with application to hurricane-induced wave heights. Environmetrics 22, 799–809.

Oceanweather, 2005. GOMOS Gulf of Mexico Hindcast. Oceanweather Inc.

Padoan, S.A., Ribatet, M., Sisson, S.A., 2010. Likelihood-based inference for max-stable processes. Journal of American Statistics Society 105, 263–277.

Pandey, M.D., Van Gelder, P.H.A.J.M., Vrijling, J.K., 2001. The estimation of extreme quantiles of wind velocity using L-moments in the peaks-over-threshold approach. Structural Safety 23, 179–192.

Pawitan, Y., 2001. In all Likelihood: Statistical Modelling and Inference using Likelihood. Clarendon Press, Oxford, UK.

Petrov, V., Guedes Soares, C., Gotovac, H., 2013. Prediction of extreme significant wave heights using maximum entropy. Coastal Engineering 74, 1–10.

Pickands, J., 1975. Statistical inference using extreme order statistics. Annals of Statistics 3, 119–131.

Poon, S.-H., Rockinger, M., Tawn, J., 2003. Modelling extreme value dependence in international stock markets. Statistica Sinica 13, 929–953.

Ramsay, T., 2002. Spline smoothing over difficult regions. Journal of Royal Statistics Society Series B 64, 307–319.

Reeve, D.T., Randell, D., Ewans, K.C., Jonathan, P., 2012. Accommodating measurement scale uncertainty in extreme value analysis of North Sea storm severity. Ocean Engineering 53, 164–176.

Ribatet, M., Cooley, D., Davison, A.C., 2012. Bayesian inference from composite likelihoods, with an application to spatial extremes. Statistica Sinica 22, 813–845.

Robinson, M.E., Tawn, J.A., 1997. Statistics for extreme sea currents. Applied Statistics 46, 183–205.

Rue, H., Held, L., 2005. Gaussian Markov Random Fields. Chapman and Hall.

Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent gaussian models by using integrated nested laplace approximations. Journal of Royal Statistics Society Series B 71, 319–392.

Ruppert, D., Wand, M.P., Carroll, R.J., 2003. Semiparametric Regression. Cambridge University Press.

Sang, H., Gelfand, A.E., 2009. Hierarchical modeling for extreme values observed over space and time. Environmental and Ecological Statistics 16 (3), 407–426.

Schlather, M., 2001. Examples for the coefficient of tail dependence and the domain of attraction of a bivariate extreme value distribution. Statistics and Probability Letters 53, 325–329.

Schlather, M., 2002. Models for stationary max-stable random fields. Extremes 5, 33–44.

Scotto, M.G., Guedes-Soares, C., 2000. Modelling the long-term time series of significant wave height with non-linear threshold models. Coastal Engineering 40, 313.

Scotto, M.G., Guedes-Soares, C., 2007. Bayesian inference for long-term prediction of significant wave height. Coastal Engineering 54, 393.

Smith, R.L., 1990. Max-stable Processes and Spatial Extremes, unpublished, available from ⟨http://www.stat.unc.edu/postscript/rs/spatex.pdf⟩.

Smith, R.L., Weissman, I., 1994. Estimating the extremal index. Journal of Royal Statistics Society Series B 56, 515–528.

Stephenson, A., Gilleland, E., 2006. Software for the analysis of extreme events: the current state and future directions. Extremes 8, 87–109.

Stone, M., 1974. Cross-validatory choice and assessment of statistical predictions. Journal of Royal Statistics Society 36, 111–147.

Tancredi, A., Anderson, C.W., O'Hagan, A., 2006. Accounting for threshold uncertainty in extreme value estimation. Extremes 9, 87–106.

Tawn, J.A., 1988a. Bivariate extreme value theory: models and estimation. Biometrika 75, 397–415.

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

VGAM. Vector generalized linear and additive models. URL ⟨www.stat.auckland.ac.nz/~yee⟩.

WAFO. Wave Analysis for Fatigue and Oceanography. URL ⟨www.maths.lth.se/matstat/wafo⟩.

Thompson, P., Cai, Y., Reeve, D., Stander, J., 2009. Automated threshold selection methods for extreme wave analysis. Coastal Engineering 56, 1013–1021.

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

Turkman, K.F., Turkman, M.A.A., Pereira, J.M., 2010. Asymptotic models and inferences for extremes of spatio-temporal data. Extremes 13, 375–397.

USACE Field Research Facility. US Army Corps of Engineers, Duck, North Carolina, USA. URL ⟨http://www.frf.usace.army.mil/⟩.

Vandewalle, B., 2004. Some Robust and Semi-Parametric Methods in Extreme Value Theory. Ph.D. thesis, Katholieke Universiteit Leuven.

Varin, C., 2008. On composite marginal likelihoods. Advances in Statistical Analysis 92, 1–28.

Vrac, M., Naveau, P., 2007. Stochastic downscaling of precipitation: from dry events to heavy rainfalls. Water Resources Research 43, 1–13. (W07402).

Wadsworth, J.L., Tawn, J.A., Jonathan, P., 2010. Accounting for choice of measurement scale in extreme value modelling. Annals of Applied Statistics 4, 1558–1578.

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

Wang, Y., Wahba, G., 1995. Bootstrap confidence intervals for smoothing splines and their comparison to Bayesian confidence intervals. Journal of Statistical Computation and Simulation 51, 263–279.

Winterstein, S.R., Ude, T.C., Cornell, C.A., Bjerager, P., Haver, S., 1993. Environmental parameters for extreme response: inverse form with omission factors. In: Proceedings of the 6th International Conference on Structural Safety and Reliability, Innsbruck, Austria.

Yee, T.W., Stephenson, A., 2007. Vector generalized linear and additive extreme value models. Extremes 10, 1–19.

Yee, T.W., Wild, C.J., 1996. Vector generalized additive models. Journal of Royal Statistics Society Series B 58, 481–493.