Modeling the Seasonality of Extreme Waves in the Gulf of Mexico

Philip Jonathan
Shell Technology Centre Thornton,
P.O. Box 1,
Chester CH1 3SH, UK
e-mail: philip.jonathan@shell.com

Kevin Ewans
Shell International Exploration and Production,
P.O. Box 60,
2280 AB Rijswijk, The Netherlands
e-mail: kevin.ewans@shell.com

Statistics of storm peaks over threshold depend typically on a number of covariates including location, season, and storm direction. Here, a nonhomogeneous Poisson model is adopted to characterize storm peak events with respect to season for two Gulf of Mexico locations. The behavior of storm peak significant wave height over threshold is characterized using a generalized Pareto model, the parameters of which vary smoothly with season using a Fourier form. The rate of occurrence of storm peaks is also modeled using a Poisson model with rate varying with season. A seasonally varying extreme value threshold is estimated independently. The degree of smoothness of extreme value shape and scale and the Poisson rate with season are regulated by roughness-penalized maximum likelihood; the optimal value of roughness is selected by cross validation. Despite the fact that only the peak significant wave height event for each storm is used for modeling, the influence of the whole period of a storm on design extremes for any seasonal interval is modeled using the concept of storm dissipation, providing a consistent means to estimate design criteria for arbitrary seasonal intervals. The characteristics of the 100 year storm peak significant wave height, estimated using the seasonal model, are examined and compared with those estimated ignoring seasonality.
[DOI: 10.1115/1.4002045]

1 Introduction

The duration of some offshore activities is limited to a certain period of time. It can be advantageous, therefore, to specify structural design criteria for short periods of the order of weeks. Here, we consider specification of seasonal design criteria for two locations in the Gulf of Mexico.

The availability of comprehensive metocean data allows the effect of the heterogeneity of extremes with respect to direction, season, and location to be accommodated in the estimation of design criteria. Capturing covariate effects of extreme sea states is important when developing design criteria. Design criteria derived from a model that adequately incorporates covariate effects can be materially different from a model that ignores those effects. In previous work (e.g., Refs. [1,2]), it has been shown that omnidirectional storm peak H_{S100} derived from a directional model can be heavier tailed than that derived from a direction-independent approach, indicating that large values of storm peak H_S are more likely than we might anticipate where we base our beliefs on estimates that ignore directionality. There is a large body of statistics literature that routinely models covariate effects in extreme value analysis (see, e.g., Ref. [3] or Ref. [4]). Anderson et al. [5] performed a seasonal analysis. The case for adopting an extreme value model incorporating covariate effects is clear, unless it can be demonstrated statistically that a model ignoring covariate effects is no less appropriate. Chavez-Demoulin and Davison [6] and Coles [7] provided straightforward descriptions of a nonhomogeneous Poisson model in which occurrence rates and extremal properties are modeled as functions of covariates. Scotto and Guedes-Soares [8] described modeling using nonlinear thresholds. A Bayesian approach is adopted by Coles and Powell [9] using data from multiple locations and by Scotto and Guedes-Soares [10]. Spatial models for extremes [11,12] have also been used, as well as models [13,14] for the estimation of predictive distributions, which incorporate uncertainties in model parameters. Ledford and Tawn [15] and Heffernan and Tawn [16] discussed the modeling of dependent joint extremes. Chavez-Demoulin and Davison [6] also described the application of a block bootstrap approach to estimate the parameter uncertainty and the precision of extreme quantile estimates, applicable when dependent data from neighboring locations are used. Guedes-Soares and Scotto [17] discussed the estimation of quantile uncertainty. Here, we investigate the effect of the seasonal covariate using a nonhomogeneous Poisson model.

The outline of the article is as follows. In Sec. 2, we describe the present application and illustrate the data. In Sec. 3, we outline the extreme value model used and describe the results. In Sec. 4, we discuss the estimation of design criteria for arbitrary seasonal intervals. Conclusions are drawn and recommendations are made in Sec. 5.

2 Data

The data examined are significant wave height H_S values from the proprietary Gulf of Mexico Oceanographic Study (GOMOS) [18] for the period September 1900 to September 2005, inclusive at 30 min intervals. For two typical Gulf of Mexico locations (henceforth, "A" and "B"), we selected 78 grid points arranged on a 13×6 rectangular lattice with spacing of 0.125 (14 km). For each storm period for each grid point, we isolated a single storm peak significant wave height H_S^{sp} for modeling purposes, together with the corresponding wave direction at storm peak θ, henceforth referred to as the storm peak direction, and the corresponding storm peak season ϕ. We expect storm peak significant wave heights for different storms at the same location to be statistically independent. For convenience and consistency, we define season ϕ on the interval [0,360) corresponding to 1 year and refer to a value of ϕ as a *seasonal degree*, approximately equal to a day of the year. Thus, extreme value modeling is performed on storm peak wave height data (one value for each storm at each of 78 locations, giving 315 storm peak values for each location in the period 1900–2005).

Contributed by the Ocean Offshore and Arctic Engineering Division of ASME for publication in the JOURNAL OF OFFSHORE MECHANICS AND ARCTIC ENGINEERING. Manuscript received July 8, 2008; final manuscript received October 28, 2009; published online December 6, 2010. Assoc. Editor: Takeshi Kinoshita.

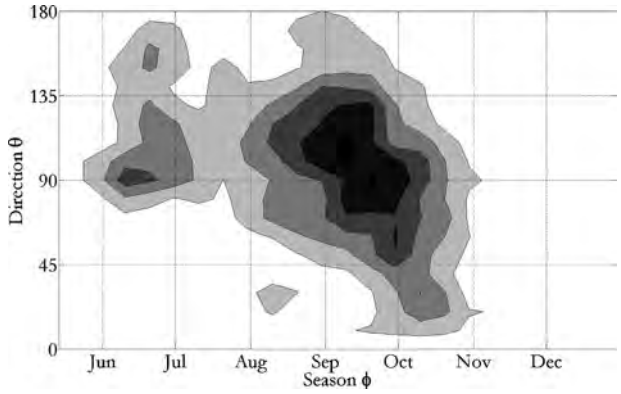

Fig. 1 Empirical density of storm peak events at location A. Darker shading represents higher density.

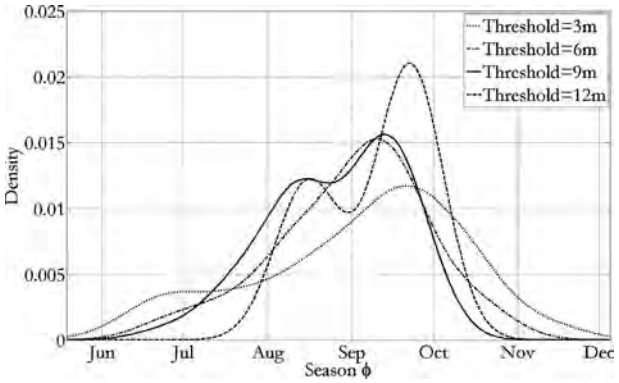

Fig. 4 Conditional density of storm peaks over threshold at location A

Figure 1 shows an empirical density of storm peak events at location A. Storm events occur in the second half of the year only (with two exceptions in 105 years; these are not shown here for convenience but are used for extreme value modeling and estimation of design criteria below). June storms are associated with a storm peak direction of approximately 90 deg. This continues approximately until October and November when storm peak direction shifts to around 45 deg. The figure suggests that the effects of storm peak direction and season on extreme storm events are related. Figures 2 and 3 give quantile estimates of H_S^{sp} with respect to season and direction, respectively, for location A. The seasonal resolution possible in Fig. 2 is less than the directional resolution possible in Fig. 3 because any storm is more localized across locations with respect to season than it is with respect to direction. In a previous application [1] the authors illustrated that storm peak direction varies between locations for a given storm by more than 30 deg. This is not the case for storm season; the storm peak event occurs at almost the same seasonal degree across locations for a given storm; there are effectively only 315 values of seasonal degree present in these data across locations, whereas the number of storm peak directions available is much larger. In this sense, site averaging of data for seasonal modeling is less advantageous than that for directional modeling.

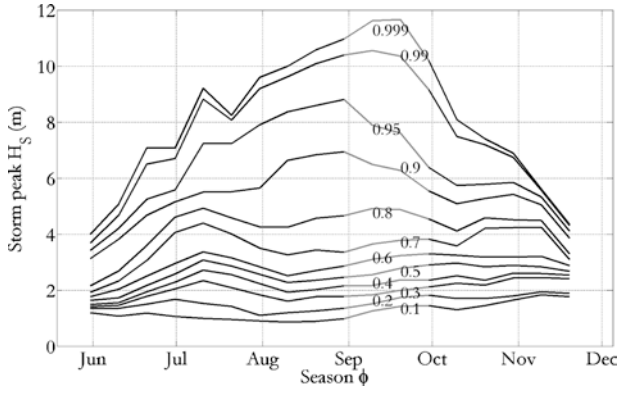

Fig. 2 Quantiles of H_S^{sp} by season at location A

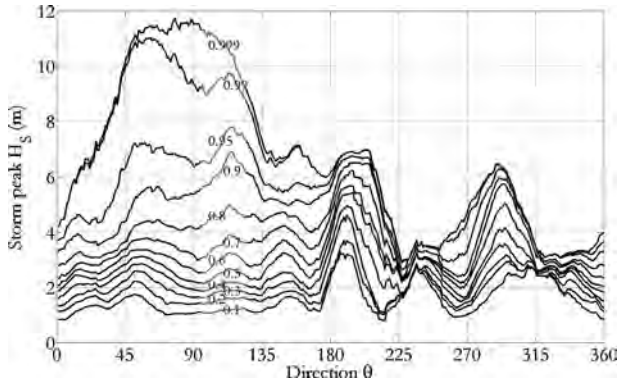

Fig. 3 Quantiles of H_S^{sp} by direction at location A

Figure 4 gives the density of H_S^{sp} for values over thresholds of 3 m, 6 m, 9 m, and 12 m, estimated using a kernel density approach. The conditional density with a 12 m threshold peaks more in September, generally consistent with Fig. 2.

For extreme value modeling with seasonal covariate, we adopt a variable threshold u to reflect variation in extreme values with season. A fixed threshold, above which all storm peaks are taken to be "extreme," appropriate in June would not be appropriate in September. We might partition the data by direction and perform independent extreme value analyses per month, e.g., assuming that months are effectively homogeneous. Here, alternatively, we adopt a threshold, which varies with storm peak season, thereby avoiding partitioning the data while accommodating seasonal heterogeneity. The variable threshold is estimated locally by identifying for each storm peak present, the nearest 300 storm peaks in terms of storm peak season. The variable threshold for that seasonal degree is then selected as a certain quantile q (e.g., the median, $q=0.5$) of H_S^{sp} for that sample of 300. The effect of varying the size of the local sample is to vary the smoothness of the estimated variable threshold profile with season. In previous work, the median was taken to provide a reasonable location for the onset of the extremal tail for all seasons. Here, we decided to use two variable thresholds corresponding to the median and the $q=0.8$ quantile. For the latter, we set the threshold such that only the highest 20% of peaks are retained locally for extreme value modeling. This enables us to assess the effect of choice of variable threshold on estimates for design conditions. For location A, variable threshold estimates are shown in Fig. 5. Interestingly, Eastoe [19] showed that, with careful model construction, forms of covariate structure invariant to choice of threshold can be achieved.

For location B, Figs. 6 and 7 correspond to Figs. 1 and 2 for location A. The empirical density in Fig. 6 is qualitatively similar to that in Fig. 1. However, Fig. 7 shows that values of the most extreme quantiles (99% and 99.9%) are larger at location B than at A.

The use of hindcast data in marine structural design is well established. Note, however, that the veracity of values for extreme events in hindcast data is dependent on the validity of the physical approximation used in the hindcast and on the quality of the observational data used to calibrate it. A useful summary of GOMOS development is given by Stiff et al. [20].

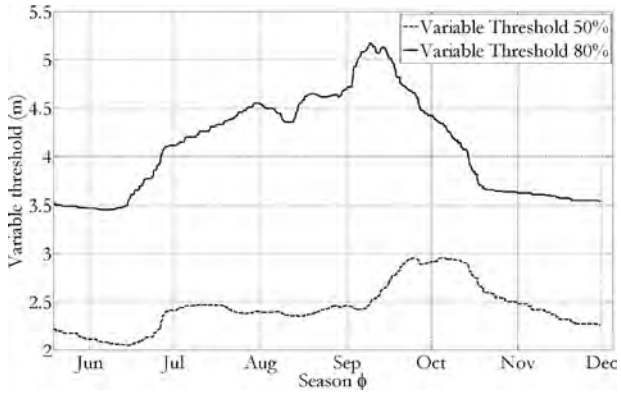

Fig. 5 Variable extreme value threshold at location A. Thresholds set to omit 50% and 80% of values for a given seasonal degree.

3 Extreme Value Modeling

Given storm peak significant wave heights $\{X_i\}_{i=1}^n$ and storm seasons $\{\phi_i\}_{i=1}^n$ (corresponding to a total of 78 different grid points for each of locations A and B) occurring in some period P_0, the distribution of storm peaks above variable threshold $u(\phi)$ can be described using the generalized Pareto (GP) distribution with cumulative distribution function $F_{X_i|\phi_i,u}$ given by

$$F_{X_i|\phi_i,u}(x) = P(X_i \le x|\phi_i, u(\phi_i)) = 1 - \left(1 + \frac{\gamma(\phi_i)}{\sigma(\phi_i)}(x - u(\phi_i))\right)_+^{-1/\gamma(\phi_i)} \tag{1}$$

for $x > u$, $\sigma > 0$, where γ is the shape parameter and σ is the scale. The subscript + notation, defined as $a_+ = \max(a, 0)$, is used.

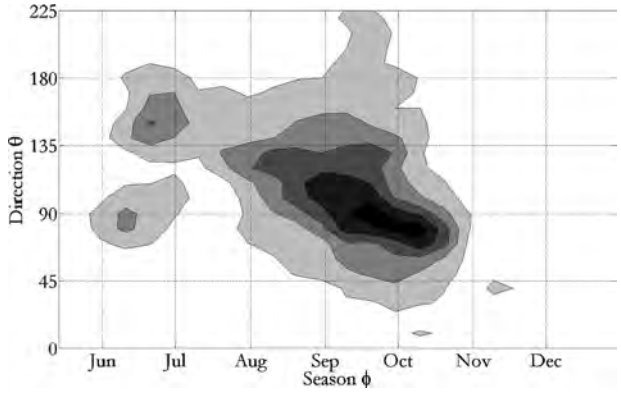

Fig. 6 Empirical density of storm peak data at location B

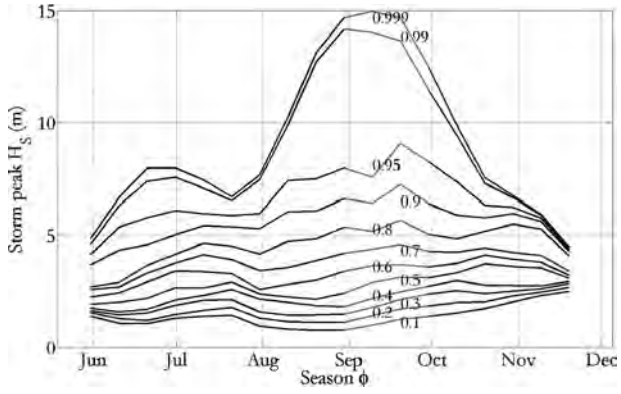

Fig. 7 Quantiles of H_S^{sp} by season at location B

We expect the extreme value parameters γ and σ to vary smoothly with season and to characterize their seasonal dependence using a Fourier series expansion $\Sigma_{k=0}^{p}\Sigma_{b=1}^{2}A_{abk}t_b(k\phi)$, where $t_1 = \cos$ and $t_2 = \sin$, with $a=1$ for γ and $a=2$ for σ. We set $A_{a20}=0$ and $a=1,2$ to avoid parameter redundancy. p is the order of the Fourier model and $p=0$ corresponds to a constant model. A Fourier form for γ and σ guarantees that parameter estimates are periodic with respect to season and permits a straightforward calculation of asymptotic covariances. The asymptotic covariance matrix of parameter estimates is given by the inverse I^{-1} of the information matrix, $I = E_X(\{-\partial^2 l/\partial A_{abk}\partial A_{\alpha\beta\kappa}\})$, where l is the log likelihood defined below. Asymptotic variances for functions of parameters, e.g., H_{S100}, can also be derived from I^{-1}. Asymptotic variances are useful as part of a bootstrapping resampling analysis to obtain reliable estimates for parameter uncertainties. The estimates $\hat{\gamma}$ and $\hat{\sigma}$ are intercorrelated, so that smoothing one functional form smoothes the other. Pooling dependent data samples (from the 78 grid points for locations A and B) with the same extremal characteristics is advantageous since the sample size for modeling is increased. However, resulting asymptotic estimates for uncertainties of model parameters and design criteria are too small due to data dependency. Techniques, such as bootstrapping, are required to obtain realistic estimates of parameter uncertainties.

It is important to note that the objective of the analysis is to estimate extreme quantiles for a *single* typical location rather than for the set of 78 dependent locations. To achieve this, we assume that the (marginal) extremal properties of each of the 78 locations are identical.

We estimate the parameters A_{abk}, $a=1,2$, $b=1,2$, and $k=1,2,\ldots,p$ using roughness-penalized maximum likelihood estimation. An order 5 model is sufficiently flexible to capture the directional dependence of γ and σ. A roughness term is incorporated for model fitting to penalize the functional forms of γ and σ, which are not smooth. The penalized negative log likelihood to be minimized takes the form

$$l^* = \sum_{i=1}^{n} l_i + \lambda\left(R_\gamma + \frac{1}{w}R_\sigma\right) \tag{2}$$

Here, the unpenalized negative log likelihood, for $i=1,2,\ldots,n$, is

$$l_i = \log\,\sigma(\phi_i) + \left(\frac{1}{\gamma(\phi_i)} + 1\right)\log\left(1 + \frac{\gamma(\phi_i)}{\sigma(\phi_i)}(X_i - u(\phi_i))\right)_+ \tag{3}$$

The roughness of γ is given by

$$R_\gamma = \int_0^{2\pi}\left(\frac{\partial^2\gamma}{\partial\phi^2}\right)^2 d\phi = \sum_{k=1}^{p}\pi k^4\left(\sum_{b=1}^{2}A_{1bk}^2\right) \tag{4}$$

The roughness of σ is given by

$$R_\sigma = \int_0^{2\pi}\left(\frac{\partial^2\sigma}{\partial\phi^2}\right)^2 d\phi = \sum_{k=1}^{p}\pi k^4\left(\sum_{b=1}^{2}A_{2bk}^2\right) \tag{5}$$

The constant w is set prior to modeling. In this work, we choose to penalize the roughness of γ exclusively. The value of roughness parameter λ is selected using cross validation to maximize model predictive performance at locations not used for fitting as follows. Using a set of ten locations (on a 3×3 grid covering the region with an additional near-central location added), models are fitted using data from nine locations only. Data from the remaining location are used to test how well the model works for prediction. This procedure is repeated until each location has been used exactly once for prediction for a range of possible values of λ. Then, we select that value of λ, which gives best predictive performance across all locations. Figure 8 illustrates overall model fitting error and predictive error as a function of λ for location A. Model *fitting* error improves with decreasing λ, whereas the *predictive* error is minimized for $\lambda = 3\times10^{-5}$, at which predictive performance is

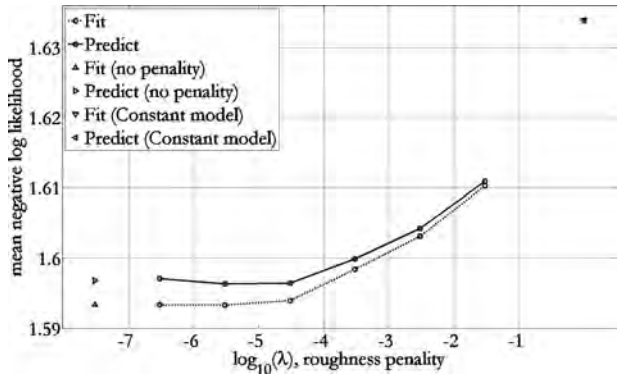

Fig. 8 Overall GP model fitting and prediction error as a function of λ with a 50% variable threshold at location A

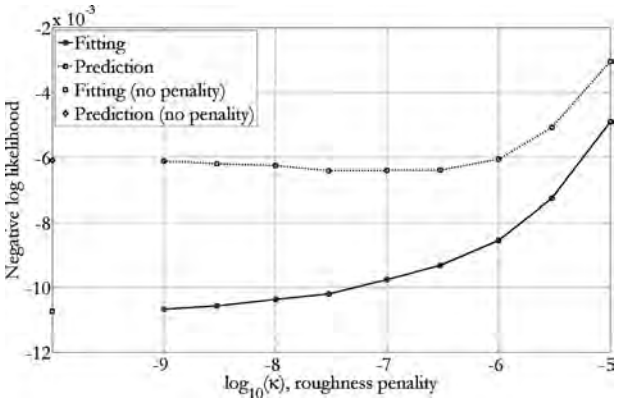

Fig. 10 Extreme value shape γ and scale σ by season with 50% and 80% variable thresholds at location A

optimal. Alternative cross-validation schemes are possible, for example, because of the dependency between locations for a given storm; we also recommend performing cross-validation studies based on omitting and reinstating storms rather than locations to confirm consistency of estimates for roughness parameters.

Corresponding functional forms of $\hat{\gamma}$ and $\hat{\sigma}$ with season, evaluated using data for all storms at all locations with $\lambda = 3 \times 10^{-5}$, are shown in Fig. 9 for location A (only June–December period is shown, but the model is estimated for the full year). The value of $\hat{\gamma}$ is seen to vary relatively smoothly from approximately -0.3 to 0, indicating considerable heterogeneity in extremal behavior. Figures 16 and 17 in the Appendix give corresponding forms with 50% and 80% thresholds at locations A and B, respectively. For location A, estimates for 80% threshold are less smoother than their 50% threshold counterparts. For location B, note that $\hat{\gamma}$ exceeds zero in September for both threshold choices. Various diagnostic plots confirm that the models explain the data well.

The current approach in the statistics literature to modeling peaks over the threshold is to adopt the inhomogeneous Poisson process model (see, e.g., Refs. [7,21]). This three parameter model describes the rate of occurrence and intensity of extreme events. If data from this model are observed over some interval, the number of threshold exceedances is Poisson distributed. Conditional on a given number of exceedances, values of the exceedances are then a random sample from a GP distribution (see Ref. [3]). The log-likelihood can be written as

$$l(\mu, \gamma, \sigma) = l_N(\mu) + l_W(\gamma, \sigma) \tag{6}$$

where l_N is the log-density of the total number of exceedances (with rate argument μ) and l_W is the log-conditional-density of exceedances given a known total number N (with the usual GP form). Since the log-likelihood can be partitioned in this way, inferences on μ can be made separately from those on γ and σ. The Poisson process log-likelihood for arrivals at times $\{t_i\}_{i=1}^n$ in period P_0 is

$$l_N(\mu) = \sum_{i=1}^{n} \log \mu(t_i) - \int_{P_0} \mu(t)dt \tag{7}$$

Chavez-Demoulin and Davison [6] described an approximate log-likelihood achieved by partitioning P_0 into a large number m of subintervals of length δ and choosing δ small enough that μ is effectively constant over each subinterval. Then,

$$\hat{l}_N(\mu) = \sum_{j=1}^{m} c_j \log \mu(j\delta) - \delta \sum_{j=1}^{m} \mu(j\delta) \tag{8}$$

where $\{c_j\}_{j=1}^m$ is the number of occurrences in each of the m subintervals. Using this approach, we estimate the storm occurrence rate based on counts of the number of storms per location per annum per seasonal degree from the hindcast data. As for GP parameters γ and σ, we adopt a Fourier form for Poisson intensity μ, penalizing its roughness R_μ by maximizing the penalized log-likelihood,

$$\hat{l}_N^*(\mu) = \hat{l}_N(\mu) - \kappa R_\mu \tag{9}$$

where R_μ has a form identical to that of R_γ or R_σ above. The optimal value of the penalty κ is chosen using the cross-validation strategy described above. Figure 10 shows that a roughness penalty of approximately $\kappa = 3 \times 10^{-7}$ is appropriate for location A with a 50% variable threshold. The corresponding estimate for Poisson rate with season is given in Fig. 11. This figure also

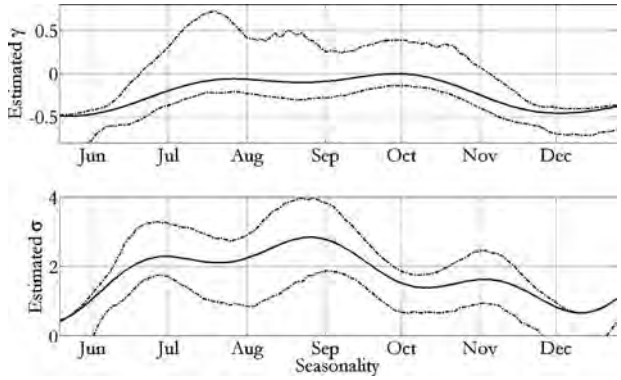

Fig. 9 Extreme value shape γ and scale σ by season with a 50% variable threshold at location A

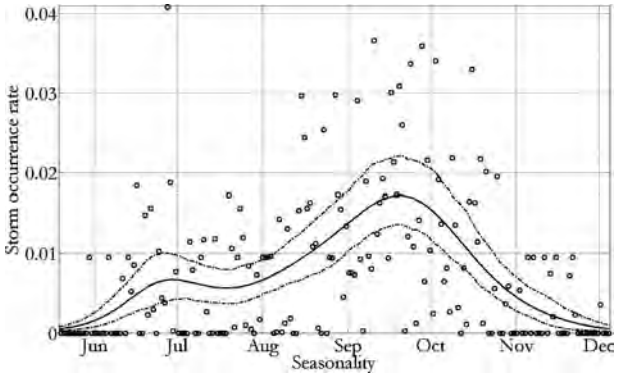

Fig. 11 Extreme value shape γ and scale σ by season with 50% and 80% variable thresholds at location B

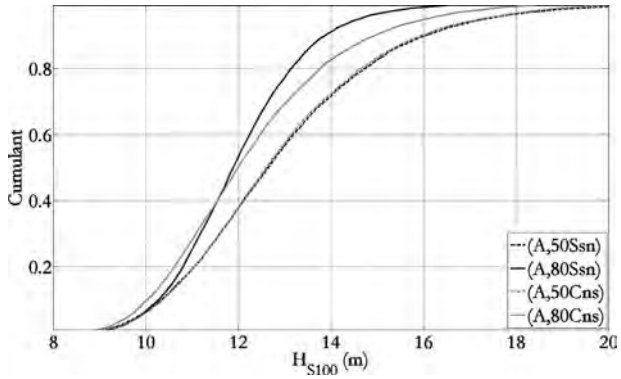

Fig. 12 Overall Poisson model fitting and prediction error as a function of κ with a 50% variable threshold at location A

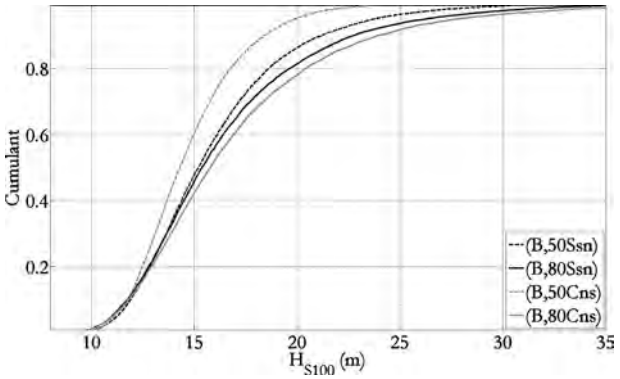

Fig. 13 Storm counts (o) and estimated Poisson rate μ by season with a 50% variable threshold at location A. Rate is defined as the number of storms per location per seasonal degree per annum.

shows the count data $\{c_j\}_{j=1}^{m}$ to which the model was fit. Figures 18 and 19 in the Appendix give corresponding estimates with either variable threshold choice for locations A and B.

4 Design Criteria

To estimate design criteria such as the 100 year storm peak event H_{S100} for an arbitrary seasonal interval (e.g., of length 1 week or 1 month), we can use our extreme value model to estimate the number of occurrences of storms in the interval and the values of H_S^{sp} for those occurrences. However, we also need to incorporate the influence of storm events whose peaks occur near but not within the seasonal interval at hand since these storm events also contribute to the extremal characteristics of the interval. To achieve this, we estimate the seasonal dissipation of storm events as follows. First, we partition the seasonal domain $[0,360)$ into 36 periods of ten seasonal degrees. Then, using hindcast data alone, we estimate the maximum (fractional) H_S^{sp} seen in each interval for a storm with any peak season. The resulting seasonal dissipation $\rho_S(\phi)$ quantifies the influence of a storm with peak season ϕ on seasonal interval S. If the value of ϕ falls within S, then $\rho_S(\phi)=1$; otherwise, $\rho_S(\phi)\in[0,1]$.

We then proceed to estimate the cumulative distribution of H_{S100} for seasonal interval S using simulation as follows. Using the Poisson model, we generate a random number of occurrences n_i of storms for each seasonal degree ϕ_i corresponding to a period of 100 years, obtaining the set $\{n_i\}_{i=1}^{360}$. Then, for occurrence j at seasonal degree ϕ_i, using the generalized Pareto model, we generate random values $\{X_{ij}\}_{i=1}^{360}{}_{j=1}^{n_i}$ for H_S^{sp}. For seasonal interval S, we retain the maximum value of H_S^{sp} observed, $\max_{i,j}\{\rho_S(\phi_i)X_{ij}\}$ from storm events occurring within or near S using dissipation $\rho_S(\phi_i)$ to factor down for peak seasonal degrees ϕ_i outside S. We repeat the procedure 10,000 times to obtain an empirical estimate for the cumulative distribution function of H_{S100}. By setting $S=[0,360)$, we can also estimate omniseasonal H_{S100}.

Figure 12 shows omniseasonal H_{S100} using 50% and 80% variable thresholds at location A. For comparative purposes, we include corresponding estimates based on a *constant* model, which ignores seasonal variability in extreme value parameters but does encapsulate variable threshold and nonhomogeneous Poisson occurrence rate. In this sense, it is important to note that the constant model here also captures a number of influential seasonal effects not typically incorporated in engineering practice. It is apparent from the figure that there is little difference between the seasonal and constant models for the 50% variable threshold case and that the 80% threshold case is less heavily tailed. The effect of choice of variable threshold appears more influential than the incorporation of seasonally varying extreme value parameters. Figure 13 gives corresponding curves for location B. Here, the 80% threshold case is heavier tailed, but again there is general consistency between the four curves shown.

Figure 14 gives cumulative distribution functions for monthly H_{S100} at location A using the seasonal model. As might be expected, the September cumulant is nearest to the omniseasonal. For comparison, Fig. 15 gives corresponding monthly cumulants assuming the constant model. There is now less variability between monthly cumulants. (Figures 20 and 21 in the Appendix give corresponding cumulants for location B. Trends are similar; September dominates for estimates based on the seasonal model, and months are more similar under the constant model.)

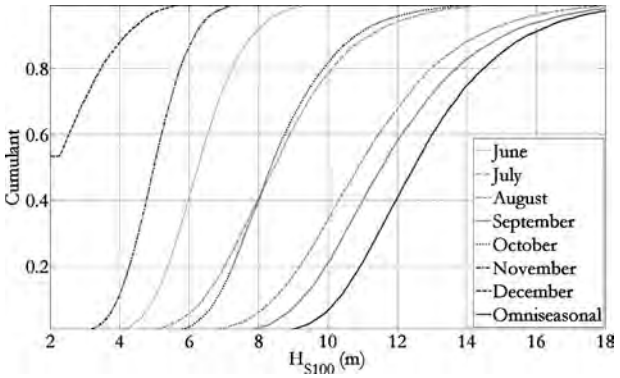

Fig. 14 Estimated Poisson rate μ by season with 50% and 80% variable thresholds at location A

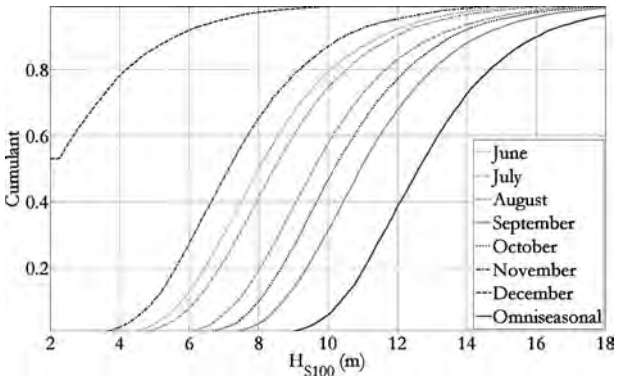

Fig. 15 Estimated Poisson rate μ by season with 50% and 80% variable thresholds at location B

Table 1 Design values corresponding to median omniseasonal H_{S100} for location A, comparing variable threshold values and model forms. NEP is nonexceedance probability.

	Variable threshold set at 50%				Variable threshold set at 80%			
	Omniseasonal		Equal monthly NEP		Omniseasonal		Equal monthly NEP	
Month	Hs (m)	NEP	Hs (m)	NEP	Hs (m)	NEP	Hs (m)	NEP
				Seasonal model				
January	12.5	1.00	3.0	0.93	11.9	1.00	5.5	0.94
February	12.5	1.00	3.8	0.93	11.9	1.00	3.9	0.94
March	12.5	1.00	3.9	0.93	11.9	1.00	4.2	0.94
April	12.5	1.00	3.3	0.93	11.9	1.00	5.8	0.94
May	12.5	1.00	4.5	0.93	11.9	1.00	5.0	0.94
June	12.5	1.00	8.2	0.93	11.9	1.00	7.8	0.94
July	12.5	0.96	11.9	0.93	11.9	0.92	12.2	0.94
August	12.5	0.74	15.2	0.93	11.9	0.75	13.1	0.94
September	12.5	0.65	15.9	0.93	11.9	0.69	14.1	0.94
October	12.5	0.97	11.4	0.93	11.9	0.96	11.1	0.94
November	12.5	1.00	6.4	0.93	11.9	1.00	6.5	0.94
December	12.5	1.00	4.6	0.93	11.9	1.00	4.3	0.94
				Constant model				
January	12.6	1.00	3.2	0.94	12.0	1.00	5.0	0.94
February	12.6	1.00	3.9	0.94	12.0	1.00	4.0	0.94
March	12.6	1.00	4.1	0.94	12.0	1.00	4.3	0.94
April	12.6	1.00	3.7	0.94	12.0	1.00	5.7	0.94
May	12.6	1.00	6.5	0.94	12.0	0.99	8.1	0.94
June	12.6	0.94	12.5	0.94	12.0	0.97	10.9	0.94
July	12.6	0.93	12.9	0.94	12.0	0.91	12.5	0.94
August	12.6	0.87	14.1	0.94	12.0	0.84	13.5	0.94
September	12.6	0.75	15.3	0.94	12.0	0.73	14.6	0.94
October	12.6	0.83	14.5	0.94	12.0	0.88	13.2	0.94
November	12.6	0.96	11.4	0.94	12.0	0.98	10.2	0.94
December	12.6	1.00	6.6	0.94	12.0	1.00	7.3	0.94

Seasonal cumulative distribution functions are the basis for design values for locations A and B. We consider two different approaches, both consistent with the same given omniseasonal design criterion, in this case the median omniseasonal H_{S100}, and quantify their relative characteristics. For definiteness, we also focus on the estimation of seasonal design values for calendar months. If the design criteria are specified in terms of an omniseasonal nonexceedance probability $q_{100\text{Omni}}$ for storm peak H_S, we obtain corresponding 100 year design storm peak H_S, $x_{100\text{Omni}}$, by solving $q_{100\text{Omni}}=P(X_{\max\ 100\text{Omni}}\leq x_{100\text{Omni}})$. However, specification of $q_{100\text{Omni}}$ does not uniquely specify design storm peak H_S. Nevertheless, we can calculate monthly nonexceedance probabilities $q_{100\text{Omni}S_i}=P(X_{\max\ 100S_i}\leq x_{100\text{Omni}})$ corresponding to $x_{100\text{Omni}}$ for months $\{S_i\}_{i=1}^{12}$ and fix the value of the all-sector nonexceedance probability $\tilde{q}_{100\text{Omni}}=\Pi_{i=1}^{12}q_{100\text{Omni}S_i}$ for all designs considered to ensure consistency. Note that $\tilde{q}_{100\text{Omni}}\neq q_{100\text{Omni}}$ in general because of the influence of storms on more than 1 month due to seasonal dissipation. Equality is achieved when each storm event influences 1 month only (which is approximately the case in the current application because of the limited dissipation of storms with season). This suggests, as limiting cases, the opportunity to design to either the omniseasonal H_{S100} or design to equal seasonal nonexceedance probability. Note that intermediate choices are also possible (e.g., [1]) propose a risk-cost design criterion).

For design to omniseasonal H_{S100}, we design the median omniseasonal storm peak H_{S100}, $x_{100\text{Omni}}$, in all 12 months. Using this approach, since months exhibit different extremal behaviors, monthly nonexceedance probabilities will vary. The all-month nonexceedance probability will be $\tilde{q}_{100\text{Omni}}$, as defined above. For the design to equal monthly nonexceedance, we design to the same nonexceedance probability $q_{100S_i}=(\tilde{q}_{100\text{Omni}})^{1/12}$ in all 12 months $\{S_i\}_{i=1}^{12}$, thereby achieving the all-month nonexceedance probability $\tilde{q}_{100\text{Omni}}$, maintaining consistency with design to median omniseasonal storm peak H_S. Using this approach, since months exhibit different extremal behaviors, sector design values will vary.

Resulting seasonal design criteria are given in Table 1 for design to the median omniseasonal H_{S100} at location A using either the seasonal or constant models, either 50% or 80% variable threshold, and either of the two design approaches. Omniseasonal values are similar regardless of model and variable threshold choices in this case. Note that nonexceedance probabilities for September are around 0.7. If we were to design to equal monthly nonexceedance probability, design criteria obtained vary considerably with month. This has interesting implications for structures to be deployed for limited periods only. For example, assume that we could locate a structure for June only. Using the seasonal model (which is likely to be more accurate than the constant model for this month), a design value of around 8 m corresponds to a monthly nonexceedance probability of approximately 0.94, which is itself the equivalent monthly nonexceedance probability were we to design to the usual omniseasonal criterion of 12.5 m. In this sense, 12.5 m for a whole year is equivalent to 8 m for June only.

Table 2 summarizes the analysis for location B. Variation in design criteria between the seasonal and constant models is greater. Seasonal model estimates a September nonexceedance probability of around 0.55 compared with 0.76 for the constant model, reflecting differences between cumulants in Figs. 20 and 21 in the Appendix. Nevertheless, general conclusions are similar. For convenience, Table 3 gives median values of monthly and omniseasonal H_{S100} for various model and design options.

Table 2 Design values corresponding to median omniseasonal H_{S100} for location B, comparing variable threshold values and model forms

	Variable threshold set at 50%				Variable threshold set at 80%			
	Omniseasonal		Equal monthly NEP		Omniseasonal		Equal monthly NEP	
Month	Hs (m)	NEP	Hs (m)	NEP	Hs (m)	NEP	Hs (m)	NEP
				Seasonal model				
January	15.0	1.00	3.9	0.94	15.4	1.00	5.2	0.93
February	15.0	1.00	4.7	0.94	15.4	1.00	4.9	0.93
March	15.0	1.00	5.1	0.94	15.4	1.00	5.4	0.93
April	15.0	1.00	4.7	0.94	15.4	1.00	6.4	0.93
May	15.0	1.00	5.7	0.94	15.4	1.00	7.1	0.93
June	15.0	1.00	8.2	0.94	15.4	1.00	8.3	0.93
July	15.0	1.00	8.8	0.94	15.4	1.00	9.0	0.93
August	15.0	0.86	17.1	0.94	15.4	0.82	19.2	0.93
September	15.0	0.54	22.1	0.94	15.4	0.56	24.4	0.93
October	15.0	0.97	13.5	0.94	15.4	0.97	14.0	0.93
November	15.0	1.00	6.5	0.94	15.4	1.00	6.6	0.93
December	15.0	1.00	4.3	0.94	15.4	1.00	4.9	0.93
				Constant model				
January	14.2	1.00	5.0	0.94	15.7	1.00	5.8	0.94
February	14.2	1.00	5.7	0.94	15.7	1.00	4.7	0.94
March	14.2	1.00	6.0	0.94	15.7	1.00	5.3	0.94
April	14.2	1.00	5.3	0.94	15.7	1.00	6.9	0.94
May	14.2	1.00	8.7	0.94	15.7	0.99	10.3	0.94
June	14.2	0.94	14.3	0.94	15.7	0.96	14.2	0.94
July	14.2	0.94	14.3	0.94	15.7	0.92	16.7	0.94
August	14.2	0.87	16.0	0.94	15.7	0.86	19.4	0.94
September	14.2	0.76	17.7	0.94	15.7	0.77	21.4	0.94
October	14.2	0.84	16.6	0.94	15.7	0.85	19.1	0.94
November	14.2	0.97	12.9	0.94	15.7	0.97	13.4	0.94
December	14.2	1.00	8.5	0.94	15.7	0.99	8.2	0.94

5 Conclusions and Recommendations

In this work, we seek to estimate seasonal design criteria for two Gulf of Mexico locations using 105 years of GOMOS hindcast data for 78 grid points at each location. The rate of occurrence of storm events is modeled using a nonhomogeneous Poisson process. An extremal threshold, which varies with season, is used to characterize the changing extremal properties of storm peak H_S with season. Two different variable thresholds are used to illustrate the influence of threshold choice on design. To model both storm arrival rate and threshold exceedances, a high-order Fourier form for Poisson arrival rate and generalized Pareto shape and scale ensures that the seasonal extreme value model is sufficiently flexible to characterize variation in extremal behavior with storm season. A roughness penalty ensures that estimates for both storm arrival rate and generalized Pareto shape and scale are as smooth as possible consistent with the data within a maximum likelihood framework. Cross validation is used to estimate the appropriate roughness penalty. In estimating seasonal design conditions, the effects of storms whose peaks occur outside the seasonal interval of interest are accommodated by modeling seasonal dissipation of storms.

It is essential to capture covariate effects in extreme sea states when developing design criteria. Design criteria estimated by incorporating covariate effects such as season more adequately reflect underlying physical processes and will be different in general to those estimated ignoring covariates. The case for incorporating covariate effects is clear, unless it can be demonstrated using statistical tests that ignoring covariates is no less appropriate. The extent to which a particular covariate will influence extreme quantiles is not possible to anticipate in general prior to modeling. For current applications, estimates for monthly cumulative distribution functions of H_{S100} based on the seasonal model show more variability with season than those based on ignoring seasonal effects on extreme values. One consequence is that for temporary

Table 3 Median values for monthly and omniseasonal H_{S100} for different variable thresholds and model forms

Location	A	A	A	A	B	B	B	B
Threshold	50%	80%	50%	80%	50%	80%	50%	80%
Model	Seasonal	Seasonal	Constant	Constant	Seasonal	Seasonal	Constant	Constant
Omniseasonal	12.5	11.9	12.6	12.0	15.0	15.4	14.2	15.7
June	6.3	5.6	8.0	6.9	6.6	6.1	9.2	7.5
July	8.5	8.8	8.5	8.6	7.2	7.1	9.2	9.4
August	10.9	10.9	9.6	9.7	11.2	11.6	10.8	11.0
September	11.6	11.1	10.9	10.7	14.6	14.8	12.3	12.6
October	8.4	8.0	10.2	9.3	9.8	9.7	11.3	11.0
November	5.0	5.1	7.1	6.3	5.1	5.2	7.7	7.2
December	0	0	0	0	3	0	3.3	0

ocean structures, a materially smaller design value can achieve (for a period of deployment avoiding the worst months of the year) the same nonexceedance probability as a materially larger omniseasonal design value. Differences between seasonal and constant models are more pronounced at location B, suggesting that the extent of seasonal variation of extreme value parameters is dependent on the location in the Gulf of Mexico. It should be recalled that the constant model used here necessarily incorporates variation in Poisson rate and extreme value threshold with season. These effects are usually ignored in design practice. When the variation in extreme value parameters with covariates is suspected not to be smooth, then other functional forms may be useful alternatives to the Fourier representation. If natural cubic spline form is selected, the mechanics of the maximum likelihood methodology is similar to that for Fourier form. Indeed, the same methodology can be extended to encompass 2D spatial covariates using natural thin-plate splines (see Ref. [22]) or finite element L-splines (see Ref. [23]).

Site averaging for seasonal modeling is less advantageous than that for a directional analysis. That there is a spring period with effectively no storm activity suggests that the Fourier form, ideal for characterizing periodic forms, might be replaced by a spline model. Illustrations of the data in Sec. 2 suggest that modeling both directional and seasonal covariate effects is important. This is being addressed in current work. Another area for development is joint modeling of multiple dependent locations (or, e.g., of wind, wave, and current extremes) following the work of Heffernan and Tawn [16].

Acknowledgment

The authors acknowledge useful discussions with Vince Cardone, Idris Eckley, George Forristall, Carlos Guedes-Soares, John Heideman, Jonathan Tawn, and Michael Vogel, and thank Joost de Haan for his assistance with the data handling. The authors further acknowledge the support of Shell International Exploration and Production and Shell Research Ltd.

Appendix

This section provides supplementary illustrations for both locations and both 50% and 80% variable thresholds for comparison, as referred to in the main text. Figures 16–21, in order, show the extreme value shape γ and scale σ by season with 50% and 80% variable thresholds at locations A and B, respectively, the estimated Poisson rate μ by season with 50% and 80% variable thresholds at locations A and B, respectively, and cumulants for monthly and omniseasonal H_{S100} with 50% variable threshold at location B for the seasonal and constant models, respectively.

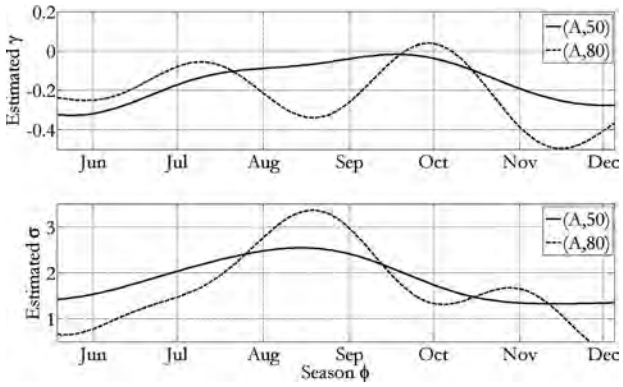

Fig. 16 Cumulative distribution function for omniseasonal H_{S100} with 50% and 80% variable thresholds at location A for the seasonal and constant models

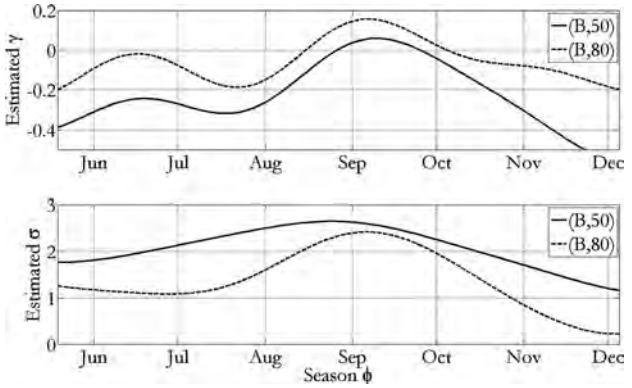

Fig. 17 Cumulative distribution function for omniseasonal H_{S100} with 50% and 80% variable thresholds at location B for the seasonal and constant models

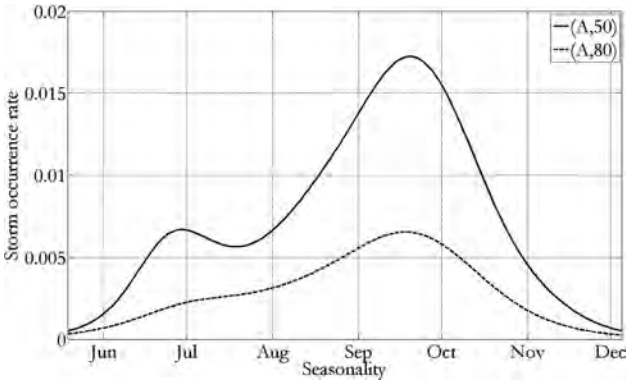

Fig. 18 Cumulative distribution functions for monthly and omniseasonal H_{S100} with 50% variable threshold at location A for the seasonal model

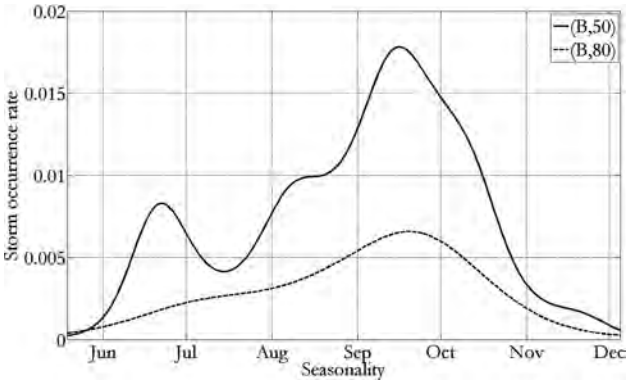

Fig. 19 Cumulative distribution functions for monthly and omniseasonal H_{S100} with 50% variable threshold at location A for the constant model

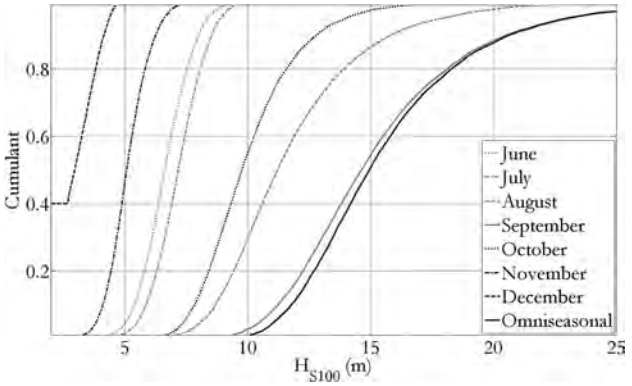

Fig. 20 Cumulants for monthly and omniseasonal H_{S100} with 50% variable threshold at location B and the seasonal model

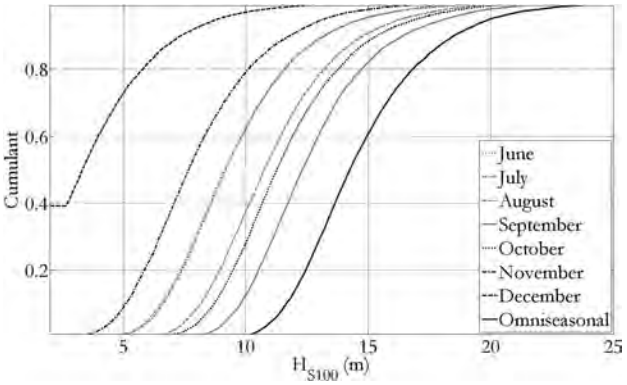

Fig. 21 Cumulants for monthly and omniseasonal H_{S100} with 50% variable threshold at location B and the constant model

References

[1] Jonathan, P., and Ewans, K. C., 2007, "The Effect of Directionality on Extreme Wave Design Criteria," Ocean Eng., **34**, pp. 1977–1994.

[2] Ewans, K. C., and Jonathan, P., 2008, "The Effect of Directionality on Northern North Sea Extreme Wave Design Criteria," ASME J. Offshore Mech. Arct. Eng., **130**, p. 041604.

[3] Davison, A. C., and Smith, R. L., 1990, "Models for Exceedances Over High Thresholds," J. R. Stat. Soc. Ser. B (Methodol.), **52**, pp. 393–442.

[4] Robinson, M. E., and Tawn, J. A., 1997, "Statistics for Extreme Sea Currents," J. Appl. Stat., **46**, pp. 183–205.

[5] Anderson, C. W., Carter, D. J. T., and 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.

[6] Chavez-Demoulin, V., and Davison, A. C., 2005, "Generalized Additive Modelling of Sample Extremes," J. R. Stat. Soc., Ser. C, Appl. Stat., **54**, pp. 207–222.

[7] Coles, S., 2001, *An Introduction to Statistical Modelling of Extreme Values*, Springer, New York.

[8] Scotto, M. G., and Guedes-Soares, C., 2000, "Modelling the Long-Term Time Series of Significant Wave Height With Non-Linear Threshold Models," Coastal Eng., **40**, pp. 313–327.

[9] Coles, S. G., and Powell, E. A., 1996, "Bayesian Methods in Extreme Value Modelling: A Review and New Developments," Int. Statist. Rev., **64**, pp. 119–136.

[10] Scotto. M. G., and Guedes-Soares, C., 2007, "Bayesian Inference for Long-Term Prediction of Significant Wave Height," Coastal Eng., **54**, pp. 393–400.

[11] Coles, S. G., and Casson, E., 1998, "Extreme Value Modelling of Hurricane Wind Speeds," Struct. Safety, **20**, pp. 283–296.

[12] Casson, E., and Coles, S. G., 1999, "Spatial Regression Models for Extremes," Extremes, **1**, pp. 449–468.

[13] Coles, S. G., and Tawn, J. A., 1996, "A Bayesian Analysis of Extreme Rainfall Data," J. Appl. Stat., **45**, pp. 463–478.

[14] Coles, S. G., and Tawn, J. A., 2005, "Bayesian Modelling of Extreme Sea Surges on the UK East Coast," Philos. Trans. R. Soc. London, Ser. A, **363**, pp. 1387–1406.

[15] Ledford, A. W., and Tawn, J. A., 1997, "Modelling Dependence Within Joint Tail Regions," J. R. Stat. Soc. Ser. B (Methodol.), **59**, pp. 475–499.

[16] Heffernan, J. E., and Tawn, J. A., 2004, "A Conditional Approach for Multivariate Extreme Values," J. R. Stat. Soc. Ser. B (Stat. Methodol.), **66**, pp. 497–546.

[17] Guedes-Soares, C., and Scotto, M., 2001, "Modelling Uncertainty in Long-Term Predictions of Significant Wave Height," Ocean Eng., **28**, pp. 329–342.

[18] GOMOS, 2005, Oceanweather, Inc.

[19] Eastoe, E. F., 2007, "Statistical Models for Dependent and Non-Stationary Extreme Events," Ph. D. thesis, University of Lancaster, UK.

[20] Stiff, J. J., Roper, R. R., Cardone, V. J., Cox, A., and Lewis, D. R., 2006, "Metocean Criteria for Jack-Ups in the Gulf of Mexico," Offshore Technology Conference, Paper No. OTC17879.

[21] Embrechts, P., Klueppelberg, C., and Mikosch, T., 2003, *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag, Berlin.

[22] Jonathan, P., and Ewans, K. C., 2010, "A Spatiodirectional Model for Extreme Waves in the Gulf of Mexico," ASME J. Offshore Mech. Arct. Eng., **133**, p. 011601.

[23] Ramsay, T., 2002, "Spline Smoothing Over Difficult Regions," J. R. Stat. Soc. Ser. B (Stat. Methodol.), **64**, pp. 307–319.