

Highlights

Quantitative comparison of environmental contour approaches

Guillaume de Hauteclocque, Ed Mackay, Erik Vanem

- Extreme responses from environmental contours compared to response-based analysis
- A large variation in the performance of the various contours was found
- Fitted models for the joint distributions contribute significantly to the errors
- Neglecting the serial correlation leads to significant overestimation
- Errors from the contour approximation itself are low for most response types

Quantitative comparison of environmental contour approaches

Guillaume de Hauteclocque^a, Ed Mackay^b and Erik Vanem^c

^aBureau Veritas, Paris, France

^bUniversity of Exeter, Penryn, United Kingdom

^cDNV-GL, Høvik, Norway

ARTICLE INFO

Keywords:

Long-term extreme response
Environmental contour
Response based analysis
Statistical model
Serial correlation

ABSTRACT

Environmental contours are a pragmatic and widespread method to estimate the long-term extreme response of marine structures. Over the years, a range of approaches have been proposed. A benchmarking study was recently conducted to compare the various methods using a common set of data. The current work extends this benchmark study by providing a quantitative assessment of the contours submitted to the exercise. The estimates of long-term responses from the contours were compared against a response-based analysis (RBA) for a wide range of responses. While some contour methods agreed well with estimates from the RBA (relative errors less than 10%), most methods were found to give large errors relative to the RBA. For the 1-year responses most methods showed a large positive bias, whilst both positive and negative biases were found for the 20-year responses. The reasons for the differences between the contours and RBA were explored. It was shown that the fitted statistical models accounted for a large portion of the error in some approaches, with both positive and negative biases of the order of 50% for some contributions, depending on the response type. Whilst for other methods, the statistical model gave accurate predictions for most responses, no models were able to capture all response behaviours for all locations. Secondly, most contour methods do not account for serial correlation in the data. It is shown that this introduces a significant positive bias into long-term response estimates, especially for lower return periods. The level of error introduced by the type of contour method is dependent on the assumption made about the shape of the failure region in the contour definition. For the predominantly unimodal response types considered, contours which approximate the failure region as having a linear boundary (IFORM and direct sampling contours), introduce relatively little error for most responses. However, for some responses, the direct sampling contours were found to introduce errors in the range 20–40%, depending on the variable space in which they are constructed. The ISORM and highest density contours were found to have a significant over-conservatism bias, which would be expected for the response types considered.

Nomenclature

H_s Significant wave height

T_z Zero up-crossing wave period

DIFORM Direct IFORM with de-clustering

DSCM Direct sampling contour method

EC Environmental contour


GHM Global hierarchical model

HDCM Highest density contour method

IDSCM Inverse directional simulation contour method

IFORM Inverse First Order Method

IID Independent and identically distributed

 guillaume.de-hauteclocque@bureauveritas.com (G.d. Hauteclocque); E.Mackay@exeter.ac.uk (E. Mackay); Erik.Vanem@dnvgl.com (E. Vanem)
ORCID(s): 0000-0001-7121-4231 (E. Mackay); 0000-0002-0875-0389 (E. Vanem)

ISORM	Inverse Second Order Method
LSQ	Least Square
MLE	Maximum Likelihood Estimate
POT	Peaks Over threshold
PPOTM	Projected peak over threshold model
RBA	Response Based Analysis
SRM	Storm resampling method
VBM	Vertical bending moment

1. Introduction

In recent years, the environmental contour (EC) method has become a popular approach for estimating the long-term extreme responses of marine structures. The EC method is widely recommended in design guidelines and standards [1–6], as it provides a practical, approximate method of estimating extreme responses, based on dynamic response calculations for a relatively small number of design conditions. One attractive feature of the EC method is that it effectively separates the probabilistic description of the environment from the structural design. Having established a set of environmental contours for a given environment, they may be used for a range of different structures and designs, and can be used to explore various design options. This is particularly useful in early stages of design.

The EC method generally involves three steps. First, a statistical model for the joint distribution of the metocean data is estimated. This statistical model is then used to construct contours. Finally, responses are calculated for a range of conditions along the N -year contour, and the largest response is taken as N -year extreme response. This procedure typically ignores the short-term variability in the response in a sea state, but there exists ways of adjusting contours to account for short-term variability in situations where this is important, see e.g. [5, 7, 8].

A wide variety of methods have been proposed, both for estimating the joint distribution of observations and for constructing contours from the joint distribution. Methods for estimating the joint distributions include global hierarchical models [9–15], kernel density estimates [16, 17], copula models [18–22], conditional extreme value models [23–25], and block-resampling methods [26]. Methods for constructing contours from the joint distribution include the inverse first-order reliability method (IFORM) [12, 27], the direct sampling contour method (DSCM) [28, 29], the direct IFORM (DIFORM) method [30], the highest density contour method (HDCM) [31], the inverse second-order reliability method (ISORM) [32] and inverse direct simulation contour method (IDSCM) [33]. It should be noted that the contours estimated by the various methods are defined differently, so they essentially represent different aspects of the joint distributions; they are not merely different approximations of the same thing.

It is widely acknowledged that the environmental contour method is a simple and approximate approach with several sources of uncertainties. Previous studies have investigated some of these uncertainties [34, 35], and also compared various aspects of different contour methods, see e.g. [36–41]. A comparison of contour-based methods and response-based methods for estimating long-term extreme ship responses is presented in [42]. Recently, an alternative approach to long-term extreme response assessment that does not rely on environmental contours was proposed in [43].

Due to the wide variety of methods available, a benchmarking exercise was recently proposed to compare contours constructed using different methods [44]. The contributions to the benchmarking exercise have recently been compiled and presented in [45]. There was a large variation in the contours derived using different methods, with the maximum value of significant wave height, H_s , along the 20-year contours differing by a factor of two for some methods. However, from the assessment metrics presented in [45] it was not clear which contours would give the most accurate estimates of long-term extreme responses. Moreover, it was not made clear how the different steps involved in the contour construction contribute to these differences and what the effect of the differences could be on typical structural responses.

The current paper aims to extend the analysis of the results presented in [45]. We present a quantitative assessment of the accuracy of long-term extreme responses derived from the various benchmark contributions, compared to a response-based analysis (RBA). A range of responses are considered, including the roll and vertical bending moment (VBM) response for eight ships and the tether tension for a tension leg platform (TLP). As the objective is to quantify

the relative error between various contour approaches and a reference estimate (RBA), simplified response models are used which are representative of realistic structures.

We then go on to explore the reasons for differences between the various contour methods and RBA. The differences are explained in terms of (a) the fit of the statistical model for the joint distribution of the data; (b) the method used to construct contours; and (c) assumptions made about serial correlation in the data. It is shown that each of these factors can significantly influence the results. Although the range of responses considered is, necessarily, limited, we attempt to draw conclusions from the analysis that have wider applicability.

The paper is organised as follows. A brief overview of the benchmarking exercise and the contributions submitted for comparison is presented in Section 2. The responses considered and the method used for RBA are described in Section 3. The results of the comparison between the contours and RBA is presented in Section 4 and the reasons for the differences are discussed in Section 5. Finally, conclusions and recommendations are presented in Section 6.

2. Overview of benchmarking exercise and contributions

The benchmarking exercise considered datasets for six locations. Three datasets contained 10 years of measurements of significant wave height, H_s , and zero up-crossing period, T_z . The other three contained 25 years of hindcast wind speed and H_s . In the current work, we restrict our attention to the three $H_s - T_z$ datasets. Dataset A was taken from NDBC buoy 44007, located off the coast of Maine, US; Dataset B was taken from NDBC buoy 41009, located off the coast of Florida, US; and Dataset C was taken from NDBC buoy 42001, located in the Gulf of Mexico.

Participants to the benchmarking exercise were asked to provide estimates of environmental contours for return periods of 1 year and 20 years for each of the three datasets. Nine contributions were made to the benchmarking exercise, details of which can be found in [45]. A summary of the contributions is presented Table 1, which lists the statistical model used for the sea state data and the method used to construct contours. No information was provided about the statistical model used contribution 7.

Of the nine contributions, seven used global hierarchical models (GHMs) for the sea state data. That is, an assumption is made about the form of the distribution of each variable, and the parameters of the distribution of one variable is modelled as conditional on the other, with a pre-assumed parametric form for the dependence function. In the GHMs the models are fitted to all observations, under the tacit assumption that they are independent and identically distributed (IID). The forms of distribution used and assumed parameter dependence models are listed in Table 2. Contributions 1, 2 and 9 all used the same form of GHM. Contributions 1 and 2 used the parameter values estimated for the baseline results, reported in [44], whereas contribution 9 used a different set of parameter estimates, reported in [46]. All other parameter values for the GHMs can be found in [45].

For contribution 5, a joint distribution model was not fitted to the data. Instead, contours are constructed using a number of univariate fits to data projected onto lines at various angles to the origin, as described in [30]. The statistical model is fitted in terms of $v_1 = H_s/s_1$ and $v_2 = H_s T_z/s_2$, where s_1 and s_2 are scale parameters, used to ensure the scale of the two variables is roughly equivalent. The scale parameters are defined as the 0.99 quantile of H_s and $H_s T_z$ respectively. The variables v_1 and v_2 are projected onto lines at various angles to the origin, $v_1 \cos(\alpha) + v_2 \sin(\alpha)$, and a threshold is set as the empirical return value at a return period of 3 months for dataset A and B and 4 months for dataset C. The generalised Pareto (GP) distribution is then fitted to declustered threshold exceedances for each angle, α . Contours are defined using the same approach as in the direct sampling contour method [28], with the quantile at each angle defined in terms of the POT fits. The contour method is referred to as the direct-IFORM (DIFORM) method.

For contribution 6, the statistical model was fitted using a multivariate block-maxima approach, described in [26]. The distribution of all data is recovered by simulating block-peak values from the fitted model and resampling and rescaling the measured blocks so that the peak values from the resampled blocks match the simulated peak values. The fit is made in terms of the significant steepness, $s = 2\pi H_s/gT_z^2$, (where g is acceleration due to gravity), and a distance variable which is orthogonal to steepness, defined as $d = (H_s^2 + T_z^2/2)^{1/2}$.

A variety of methods were used to construct contours from the data. Contribution 9 submitted three contours for each dataset, constructed from the same joint distribution, but using DSCM, smoothed DSCM and IFORM (these are labelled 9a, 9b and 9c from hereon). Contributions 1, 3 and 4 all used contours which are defined in terms of the probability that an observation falls anywhere outside the contour (referred to as the total exceedance probability). Whereas, the other contributions used contours defined in terms of the probability that an observation falls within a half-plane region, tangent to the contour (equivalent to a marginal exceedance probability under rotations of the axes). For a given joint distribution, contours defined in terms of total exceedance probability are significantly more

Table 1

List of contributions to the benchmarking exercise [45]

Contribution	Authors	Model for sea state data	Contour method
1	W. Chai, B. Leira	GHM	ISORM
2	C. Guilherme, C. Guedes Soares	GHM	DSCM
3	Á. Hannesdóttir, N. Dimitrov	GHM	IDSCM
4	A. F. Haselsteiner, A. Sander, J.-H. Ohlendorf, K.-D. Thoben	GHM	HDCM
5	G. de Hauteclocque	PPOTM	DIFORM
6	E. Mackay, P. Jonathan	SRM	IFORM
7	C. Qiao, A. Myers	-	IFORM
8	A. Rode, A. Hildebrandt, B. Schmidt	GHM	IFORM
9	E. Vanem, A. B. Huseby	GHM	IFORM and DSCM

Table 2

Global hierarchical models used in the benchmarking exercise. 2-P = 2-Parameter. 3-P = 3-Parameter. μ and σ are log-normal location and scale parameters. α and β are Weibull location and scale parameters. g is acceleration due to gravity. c_j are constants estimated from the data.

Contribution	Conditioning variable and model	Conditioned variable and model	Parameter dependence model
1, 2, 9	H_s , 3-P Weibull	T_z , 2-P Log-normal	$\mu = c_1 + c_2 h_s^{c_3}$ $\sigma = c_4 + c_5 \exp(c_6 h_s)$
3	H_s , 3-P Log-normal	T_z , 2-P Log-normal	$\mu = c_1 + c_2 h_s^{c_3}$ $\sigma = c_4 + c_5 \exp(c_6 h_s)$
4	H_s , 3-P Exponentiated Weibull	T_z , 2-P Log-normal	$\mu = \log(c_1 + c_2 \sqrt{h_s/g})$ $\sigma = c_3 + c_4/(1 + c_5 h_s)$
8	T_z , 2-P Weibull	H_s , 2-P Weibull	$\alpha = c_1 t_z^2 + c_2 t_z + c_3$ $\beta = c_4 t_z^2 + c_5 t_z + c_6$

conservative than contours defined in terms of marginal probabilities (see [47] for a discussion of this). Another key difference in the contour methods is the assumptions that are made about serial correlation in the data. The DIFORM method directly accounts for serial correlation, whereas all the other contour methods considered here, make the tacit assumption that hourly observations are independent. The effect of this is discussed further in Section 5.3.

Figure 1 shows the 1-year and 20-year contours for all contributions to the benchmarking exercise. There is a large scatter in the results for both the 1-year and 20-year contours. For the 20-year contours, the largest values of H_s on the contours differ by more than a factor of 2 between the highest and lowest contributions. There is a similar spread of values for T_z , which is clearly visible for the longer period values. For the short-period side of the contours, it is clearer to plot the contours in terms of H_s and significant steepness, as shown in Figure 2. From this figure, it is evident that many of the contributions to the benchmarking exercise give 20-year contours that contain conditions with steepnesses far in excess of the observed range, and which exceed physical limitations due to wave breaking (sea states with significant steepness in excess of 0.1 correspond to hurricane-strength winds and short fetches).

3. Responses used for assessment

The aim of the environmental approach is to provide estimates of long-term response values, at reasonable computational cost. Therefore, to evaluate the contour approaches in the benchmarking exercise, we compare the contours to estimates of long-term responses from a method that makes fewer approximations. The reference method used here for the long-term response assessment is response-based analysis (RBA). There is currently no commonly-agreed interpretation of what RBA involves. In the current work we use a peaks-over-threshold (POT) method, described in Section 3.2.

Quantitative comparison of environmental contour approaches

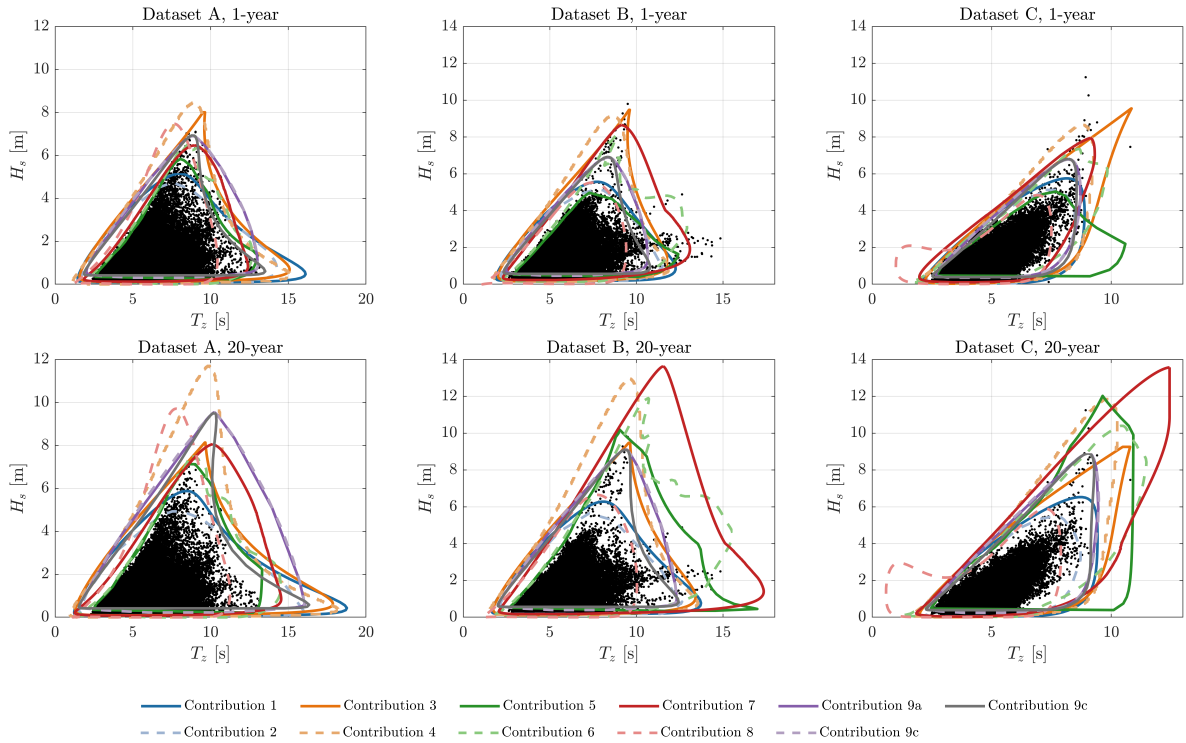


Figure 1: Environmental contours submitted to the benchmarking exercise and 10-year observed datasets (dots).

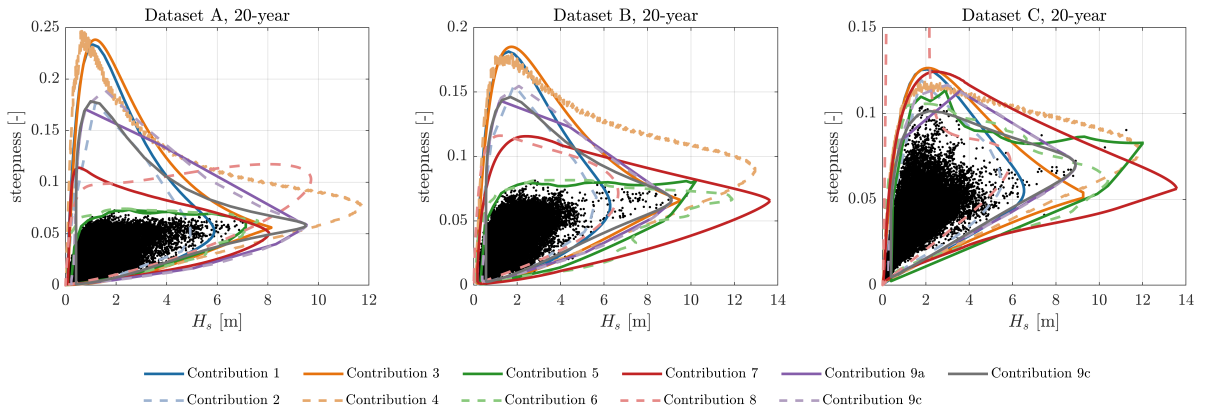


Figure 2: As previous figure, but plotted in terms of H_s and significant steepness.

To simplify the analysis, we do not consider the problem of accounting for short-term variability in the response. Instead, we consider the ‘significant response’ (defined in Section 3.1 below), which is assumed fixed for a given sea-state, and not random. Practical solutions for accounting for short-term variability in contour methods are discussed in e.g. [7, 8]. If the short-term response function is known, then short-term variability can also be accounted for straightforwardly in a POT approach using a Monte Carlo method (see e.g. [48, 49]). However, this is not considered further here. Since short-term variability is not accounted for in any of the contour methods considered here, neglecting this effect does not influence the relative comparison between the methods.

CodeName	TYPE	L (m)	B/L	D/L	T_r (s)	C_b
G03	Liquefied gas carrier	93.5	0.18	0.06	10.35	0.77
R05	Roll-on/Roll-off	148.0	0.17	0.05	7.48	0.66
T22	Tanker	164.2	0.17	0.07	11.90	0.76
B26	Bulk carrier	169.6	0.17	0.06	12.10	0.83
C19	Container-Ship	172.8	0.19	0.06	20.27	0.62
C03	Container-Ship	183.1	0.16	0.06	25.41	0.67
B30	Bulk carrier	211.2	0.15	0.06	15.59	0.83
B22	Bulk carrier	213.9	0.15	0.07	9.19	0.87

Table 3

Ships considered for the response models.

L is the length, B the beam, D the draft, T_r the roll resonance period and C_b the block coefficient (the ratio of the displaced volume of water to $L \times B \times D$)

3.1. Responses investigated

To be able to draw reasonably general conclusions, a variety of responses are investigated. We consider the vertical bending moment (VBM) in head-seas and roll response in beam-seas of eight ships, with dimension listed in Table 3. Ships with different lengths and resonance periods have been chosen to cover a wide range of characteristic periods, with the natural period for the roll response between 7.5 s and 25 s. Although, only two types of ship response are considered, roll is typical of many resonance driven responses, with a sharp-peaked narrow-band response. In contrast, VBM, is more excitation-driven, with a broader bandwidth. Thus the roll and VBM responses used here can be considered as representative of a wider range of responses of offshore structures.

The response amplitude operators (RAO)s for the ships, were calculated with using a 3D boundary element code, using the HydroStar software package [50]. The normalised VBM and roll RAOs are shown in Figure 3. The VBM and roll responses are linear in H_s . For both the RBA and the contour methods, we assume that waves are unidirectional and each sea state is represented by a JONSWAP spectrum with peak enhancement factor $\gamma = 1.5$. The response spectrum is then defined as $S_r(f) = S_w(f)RAO^2(f)$, where S_w is the wave spectrum and the ‘significant response’ R_s , is defined as $R_s = 4\sqrt{m_0}$, where m_0 is the zeroth moment of S_r (analogous to the definition of significant wave height). The significant responses, R_s , are calculated for each sea-state in the time-series for datasets A, B and C, and for all sea-states along the various environmental contours.

In addition, we consider a nonlinear tether tension response for a tension leg platform (TLP), described in Appendix A. Finally, we also consider H_s as a response variable, as wave elevation is itself a value of interest, for instance, for air-gap analysis.

It is acknowledged that the response models used here are approximate. For instance the hogging/sagging non-linearity in the VBM is not accounted for; and the linear roll response does not account for a quadratic damping term or for parametric roll. Since the aim of the current work is to compare various contour approaches to RBA, the precise values of the response functions are of less importance. The response models used here are therefore considered sufficiently realistic and the use of more complex nonlinear responses would not be expected to change the conclusions.

3.2. Response based analysis

The RBA method applied here, uses a peaks-over-threshold (POT) method, following standard textbook approach [51]. For a given response and dataset, the significant response R_s is calculated for each sea state in the time series, as described in the previous section. The response time series are then declustered to identify local maxima in the time series, separated by a minimum of 48 hours, which is assumed sufficient to ensure independence between adjacent peaks. The threshold is set as the empirical 3-month return value of the declustered response peaks. The generalized Pareto (GP) distribution is then fitted to the exceedance, using maximum likelihood estimation (MLE).

Return periods and return values can be defined in a number of ways which are asymptotically equivalent (see e.g. [52]). Here, we define a return value of variable X , denoted X_T , at return period T -years, in terms of the cumulative distribution function (CDF) of the maximum value of X in a T -year period, denoted $F_T(X)$. The return value is defined as the solution of

$$F_T(X_T) = \exp(-1) \approx 0.3679. \quad (1)$$

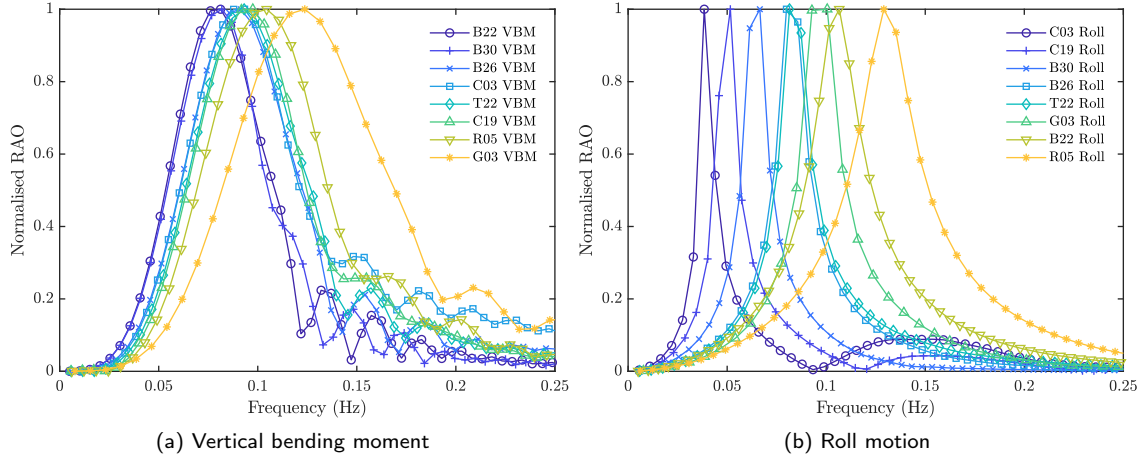


Figure 3: Normalised RAOs considered in the study.

This definition makes no assumption about what constitutes ‘independent events’ and is applicable to return periods less than one year. The definition can be shown to be asymptotically equivalent to defining return values as the solution of

$$F_c(X_T) = 1 - \frac{1}{MT}, \quad (2)$$

where $F_c(X)$ is the CDF of independent cluster-maxima of X , and M is the expected number of cluster-maxima per year [51, §4.3.3]. In the case that the cluster size is taken as one year, then $M = 1$, and this expression is equivalent to the definition in terms of the distribution of annual maxima for $T \gg 1$ year. In the case that it is assumed (erroneously) that hourly data are independent, then $M = 24 \times 365.25$ is the average number of hourly values per year. The latter case, considering hourly (or 3-hourly) observations as independent, is commonly used to construct contours (for all contributions to the benchmark exercise, other than contribution 5), so is also used here for reference.

Examples of the RBA for the VBM and roll responses of ship G03 are illustrated in Figures 4 and 5, respectively. Centred 95% confidence intervals (CI) have been calculated using the likelihood-delta method [51]. Iso-failure surfaces for each response are shown in Figure 6, where the failure value is taken as the 20-year response from the RBA.

Given the limited data available, there is considerable uncertainty associated with the 20-year return values. Due to this uncertainty, the RBA estimates of 20-year return values are used as a reference for comparison, and are not intended to be interpreted as the ‘true’ reference value. However, the confidence intervals for the 1-year values are much narrower, and provide a more certain reference for comparison with the contour estimates.

The RBA performed here is, relatively, easy for two reasons: the short-term variability is ignored, and the responses are calculated at almost no CPU cost. Actual responses relevant for design are likely to require time-consuming simulation which might not be computationally-feasible for several years of data (hence the need for simplified approach as environmental contour). As the focus of this work is to compare, relatively, response estimates from EC and RBA, the response functions used are considered sufficiently accurate for meaningful conclusions to be drawn.

4. Results

For each of the responses investigated, the maximum response over each individual contour is calculated and compared to the RBA, for the 1 and 20-year extreme response, respectively. For a given contour, dataset and response, the relative error compared to the RBA is defined as

$$\epsilon_i = \frac{R_{EC,i}}{R_{RBA,i}} - 1, \quad (3)$$

where $R_{EC,i}$ and $R_{RBA,i}$ are the response estimates from the EC and RBA, respectively. The full results for all contributions to the benchmarking exercise are listed in Tables 4 and 5, in Appendix B.

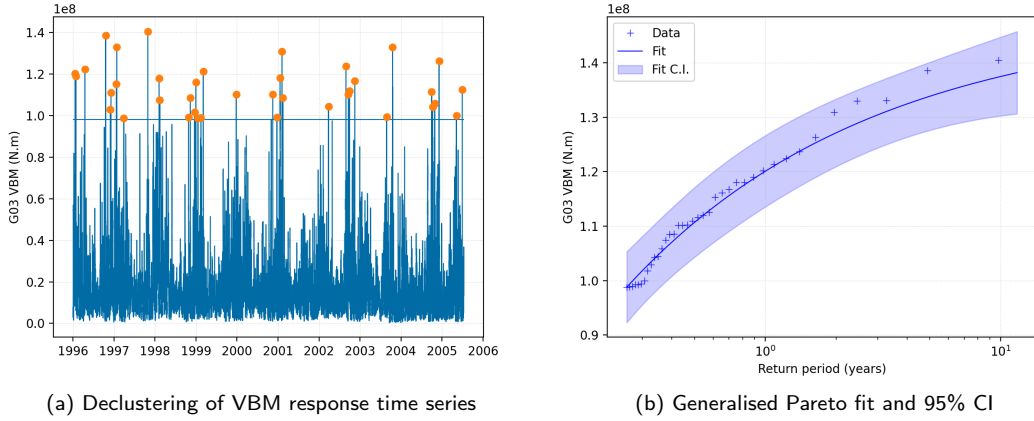


Figure 4: POT analysis of VBM significant response for ship G03 for dataset A.

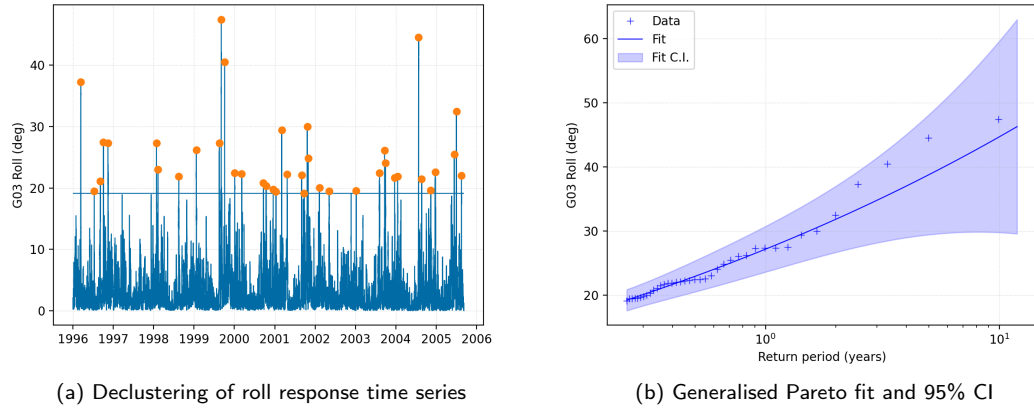


Figure 5: POT analysis of VBM significant response for ship G03 for dataset B.

To get a more concise overview, we summarise the results in terms of the mean error, $\mathbb{E}(\epsilon)$, root-mean-square error (RMSE), and coefficient of variation (CoV), defined as

$$\mathbb{E}(\epsilon) = \frac{1}{n} \sum_{i=1}^n \epsilon_i, \quad (4)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \epsilon_i^2}, \quad (5)$$

$$\text{CoV} = \frac{\sigma(\epsilon)}{\mathbb{E}(\epsilon + 1)}, \quad (6)$$

where $\sigma(\epsilon)$ denotes the standard deviation, and the summation is over all response types and the three datasets.

The summary errors statistics for the 1-year and 20-year responses are shown in Figure 7. Apart from contribution 2, all contributions show a positive bias at the 1-year level. Contributions 5 and 8 have only a small positive bias of 2% and 6%, respectively. In contrast, contributions 3, 4, 6, 7 and 9a-c have positive biases in the range 30-90% and correspondingly large RMSE. The biases are significantly reduced at the 20-year level, with contributions 1, 2, 3 and 8

Quantitative comparison of environmental contour approaches

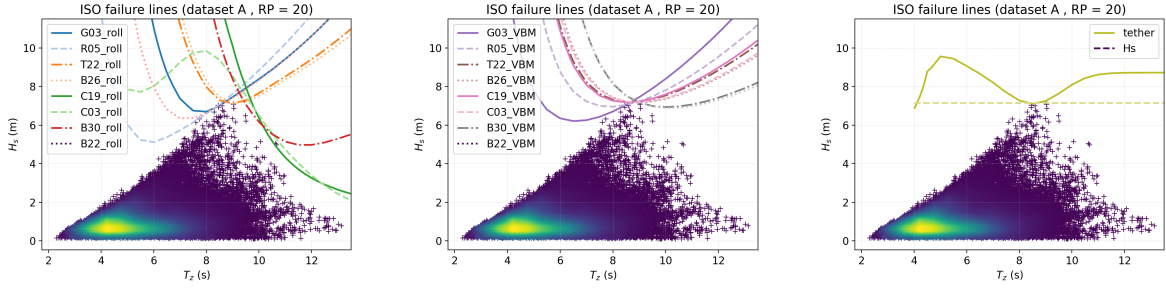


Figure 6: Iso-failure lines for each response for dataset A. Failure value is taken as 20-year response from the RBA.

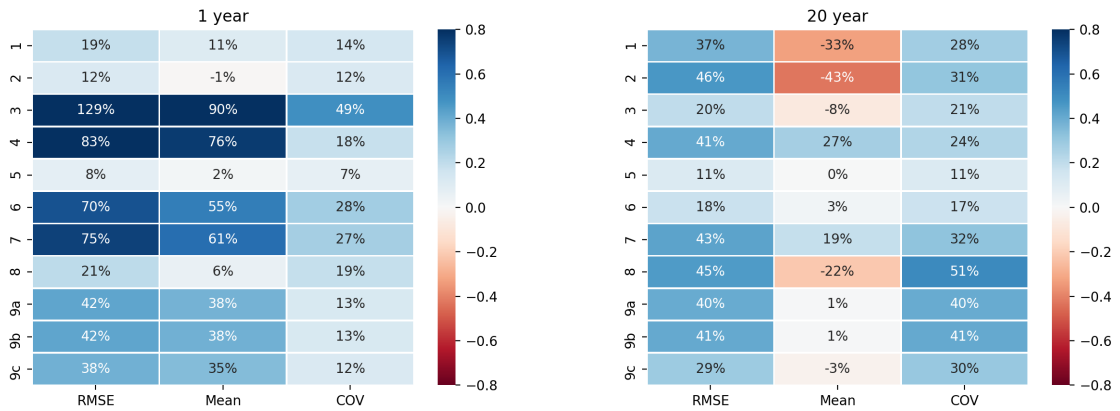


Figure 7: Errors statistics for the contributions to the benchmarking exercise.

exhibiting negative biases. Contributions 5, 6 and 9 have mean errors close to zero, while contributions 4 and 7 remain positively biased relative to the RBA. The CoV in the errors is increased compared to the 1-year values.

Given the relatively large confidence intervals for the 20-year return values from the RBA, some of the increase in the CoV of the errors at the 20-year level can be attributed to sampling uncertainties in the RBA. In an attempt to quantify this effect, Figure 8 shows the fraction of the cases for which the contour value lies within the 95% confidence interval for the RBA estimate. From Figures 7 and 8, it appears that, except for contribution 5, most of the contour fail to capture correctly the 1-year return values, for which the sampling uncertainties are small. With regard to 20-year return values, the comparison is more ambiguous, as the confidence interval associated with the reference RBA calculations are larger. However, a significant number of the contour estimates fall outside the 95% confidence interval.

5. Discussion

To explain the reasons for the large differences between the responses estimated from the various contours and the RBA, we consider three factors which influence the results. These are, (1) the fit of the statistical model to the data; (2) the contour method used; and (3) assumptions made about serial correlation in the data. In the following subsections, we isolate the influence of each of these factors, to determine how much they influence results.

5.1. Statistical model

In this section, we begin by making a qualitative assessment of the fit of the statistical models to the data, before considering how this affects the accuracy of response predictions. The fit of the models is assessed graphically in terms of the marginal distribution of H_s and the joint density functions. The fit of the models could also be assessed using various other diagnostic plots, such as binning the data into ranges of either H_s or T_z and looking at the fit

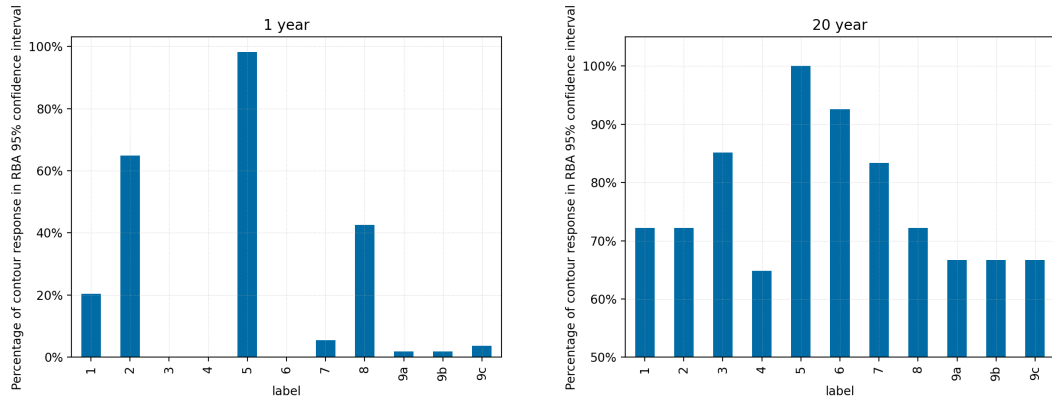


Figure 8: Percentage of contour response in RBA confidence interval

of the models through ‘slices’ of the joint distribution, or by examining the fit of model to data projected onto lines at various angles to the origin (using a similar idea to the direct sampling or direct IFORM contours). Alternatively, quantitative metrics of goodness-of-fit, such as the Kullback–Leibler divergence, could be calculated. It should be noted, however, that for extreme response estimation, it is the fit of the tails of the distributions that are important, and the overall goodness-of-fit measured by various metrics may not be directly relevant. Hence, for the purpose of explaining the errors in response predictions, which result from the fit of the statistical model, the diagnostic plots give a good indication of the reasons for discrepancies.

As noted in Section 2, no information on the statistical model for contribution 7 was provided. Furthermore, we consider contribution 5 separately to the other contributions, since this method only fits statistical models to declustered threshold-exceedances for projected data, whereas the other methods attempt to replicate the joint distribution of all observations. For contribution 5, the fits of the GP model for different projection angles, α , are shown in Figure 9. It can be seen that the fitted GP models are reasonably good match for the data for the three sites.

The return values of H_s for the observations and fitted models is shown in Figure 10 for the three sites (where return periods are calculated under the assumption of IID hourly observations). The model for contributions 1 & 2 significantly under-estimates the probability of large H_s for all three sites. For datasets B and C, the marginal model for H_s for contribution 8 is a poor fit to the data, significantly overestimating H_s at lower exceedance probabilities for dataset B and underestimating the occurrence of large H_s for dataset C. For datasets A and B, there is range of just over 1 m in the return values of H_s at the 1-year level for the other fitted models, and a spread of just over 2 m at the 1-year level for dataset C. The estimates of return values diverge significantly at higher return periods. For dataset A, there is a spread of 3.5 m at the 10-year level (ignoring contribution 1) and a spread of nearly 7m at the 100-year level. A similar range of results is observed for dataset C, with the range for dataset B being slightly smaller.

Apart from contribution 6, all the fitted statistical models shown in Figure 10 are based on global models (although for contribution 8, H_s is modelled as conditional on T_z). Figure 10 clearly illustrates the strong influence that the prior assumption about the form of the distribution can have on estimates of return values, even for return periods shorter than the length of the dataset. For dataset B there is a change in the gradient of the tail at a return period of around 10^{-1} , which indicates that the use of a simple parametric model may not be adequate. Both datasets B and C are located in regions influenced by tropical cyclones, so the change of gradient in the tail may be related to the extreme observations coming from a different population to the more common climatic conditions.

Figures 11 - 13 show scatter plots of observed H_s and T_z for the three datasets, overlaid with isodensity contours of the joint density functions for the fitted model. Contours are shown at probability densities of 10^{-n} for $n = 1, \dots, 6$. Note that these are not “environmental contours” as there is no fixed exceedance probability associated with them. Instead, the joint density function is shown instead of a particular type of environmental contour, as it can be used directly and unambiguously to assess the fit of the model. Although the empirical density is not shown, a qualitative visual assessment of the fit of the model can be made by noting that if the model is a good fit for the data, then the occurrence of observations between isodensity contours should be approximately constant. For the 10-year datasets

of observations at 1 hour intervals, the expected number of points in a square of size $\Delta t \times \Delta h$ is approximately $10 \times 365.25 \times 24 \times \bar{f}(t, h) \Delta t \Delta h$, where $\bar{f}(t, h)$ is the average density in the square. This equates to ≈ 8.7 observations in a 1×1 square where the average density is 10^{-4} or ≈ 3.5 observations in a 2×2 square where the average density is 10^{-5} . Due to serial-correlation in sea states, clustering of observations is expected in low density regions, meaning that the spread of observations will not be completely even. However, we can get an idea of systematic biases in the models by looking at how the density of observations compares to the isodensity contours.

As with the 1-year and 20-year environmental contours, shown in Figure 1, there is significant variation between the isodensity contours of the fitted models for each site. Some models are evidently in poor agreement with the data. For example, the model from contribution 8 at site B predicts occurrence of much steeper sea states than observed, and at site C, the model predicts the occurrence of relatively large H_s at very low T_z . A common feature of models for datasets A and B, is that all models, other than contribution 6, predict the occurrence of significantly steeper conditions than observed (this is evident from there being no observations in the high steepness side of the joint density contours). This is a result of the a priori assumption that the distribution of T_z conditional on H_s is log-normal (for contributions 1, 2, 4 and 9), which evidently is not a good fit to the data at low probability levels for datasets A and B, but appears to be a slightly better model for datasets C. For contribution 3, the model for the scale parameter of the conditional log-normal distribution becomes negative at high H_s , for all three sites, resulting in the joint density function not being defined above this level. So although the marginal model for H_s predicts a non-zero density above this level, the joint model is undefined when the log-normal scale parameter is negative, so the total probability of the joint distribution is less than one.

Contribution 6 appears to give the closest match to the data for steep sea states, which is likely related to fitting of the model to the data explicitly in terms of steepness and the distance variable (see Section 2). However, for datasets B and C it appears that it predicts more long-period sea states at large H_s than are observed. Of the other models, the fit of the joint distributions for the longer periods is rather mixed, with no model giving a consistently good match to the data at all sites.

To isolate how the fitted statistical models influence the accuracy of the response estimates, we compare responses calculated directly from the observations to those calculated from simulated data from the statistical models. For each of the statistical models, 1000 years of hourly observations were simulated and the VBM and roll responses for the eight vessels were calculated. The effect of the fit of the statistical models used in contribution 5 on the response estimates, cannot be decoupled from the contour approach in the same way as for the other contributions. We therefore do not consider contribution 5 in the following discussion. For contribution 3, when the log-normal shape parameter is zero or negative, the value of T_z has been set to e^μ (i.e. the mean value when $\sigma = 0$). Figures 14 and 15 show the percentage errors in return values of VBM and roll from the statistical models compared to those calculated directly from the observations, together with the percentage error in return values of H_s (again, calculated under the assumption that hourly observations are IID). It should be noted that there is significant sampling error in the largest empirical response values, so some level of scatter is expected on the right hand side of the plots.

Since these response functions are linear in H_s , differences in the response return values scale linearly with differences in return values of H_s . If there was no bias in the marginal values of H_s from the fitted models, then models that predict a greater occurrence of long-period sea states than observed will lead to a positive bias in return values for the responses with higher natural periods (and vice versa for short periods). In reality, error due to biases in H_s and conditional T_z are combined, so the result of biases in T_z appear as a scatter about the biases in the return values of H_s alone. Some level of scatter in the return values from observations and fitted models would be expected due to finite sample size effects. However, systematic differences in the performance of each contribution are clearly evident.

The effect of biases in T_z have a lesser effect on the VBM response, since this is more broad-banded meaning that any errors are averaged over a range of periods to a certain extent. We therefore focus the discussion on the roll response ratios shown in Figure 15. There is a wide range in the responses predicted by various contributions, consistent with the differences in the statistical models, discussed above. The model for contributions 1 and 2, significantly underestimates responses for all three datasets, due to the underestimate in marginal H_s , and with a larger negative bias for the longer period responses. For contribution 3, there is a large bias in the longer period responses, with the ratio of return values in excess of 2 for dataset A at the 10-year level and in excess of 3 for dataset C at the 10-year level (note that the vertical scale has been cropped at 2 to provide a better visualisation for the other contributions). Similarly, there is a large scatter of response ratios for contribution 8, due to the poor match of the joint distribution to the observations.

Quantitative comparison of environmental contour approaches

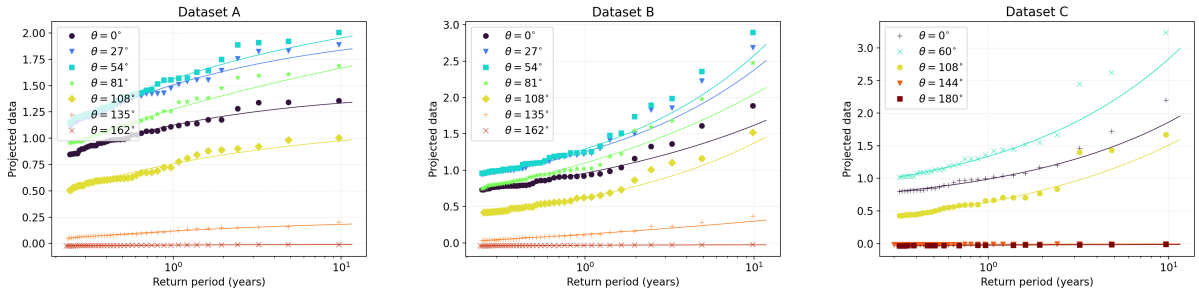


Figure 9: Fit of projected, declustered data for contribution 5.

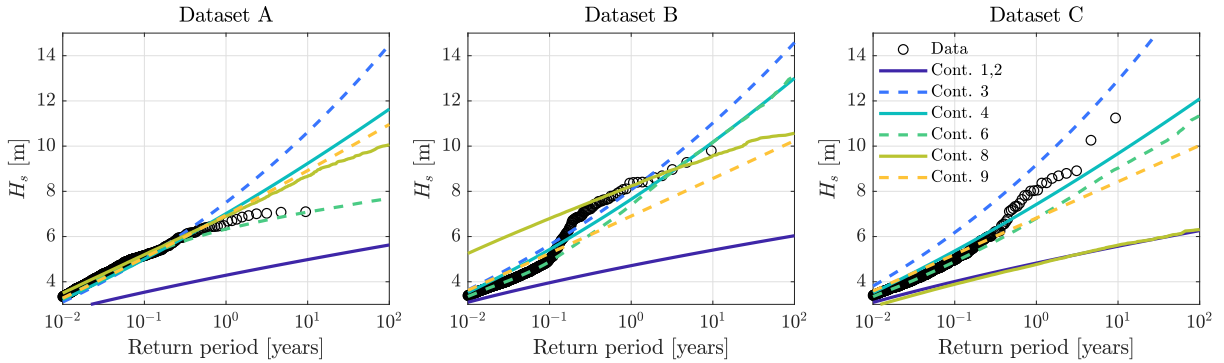


Figure 10: Comparison of exceedance probabilities for observed and modelled of H_s for the three sites.

The scatter is lower for the other contributions, although there are still significant differences for all models for at least one site, with differences around 50% in the return values at the 10-year level for some responses.

A summary of the error statistics from fitted statistical models compared to RBA under the assumption of IID hourly observations is shown in Figure 16. The RBA under the assumption of IID observation is performed in the same way as the one presented in section 3.2, except that all sea-state above threshold are considered (no declustering).

At the 1-year level, contribution 4 has the lowest bias and RMSE over all responses and sites, whilst at the 20-year level, contributions 6 and 9 both have a mean error of 0%, and contribution 6 has the lowest RMSE. The largest errors at both the 1- and 20-year levels are from contributions 1 and 3. The results indicate that the fit of the statistical model can introduce large biases into response estimates. Therefore, ensuring that the statistical model is a good fit to the data is of key importance for obtaining accurate response estimates.

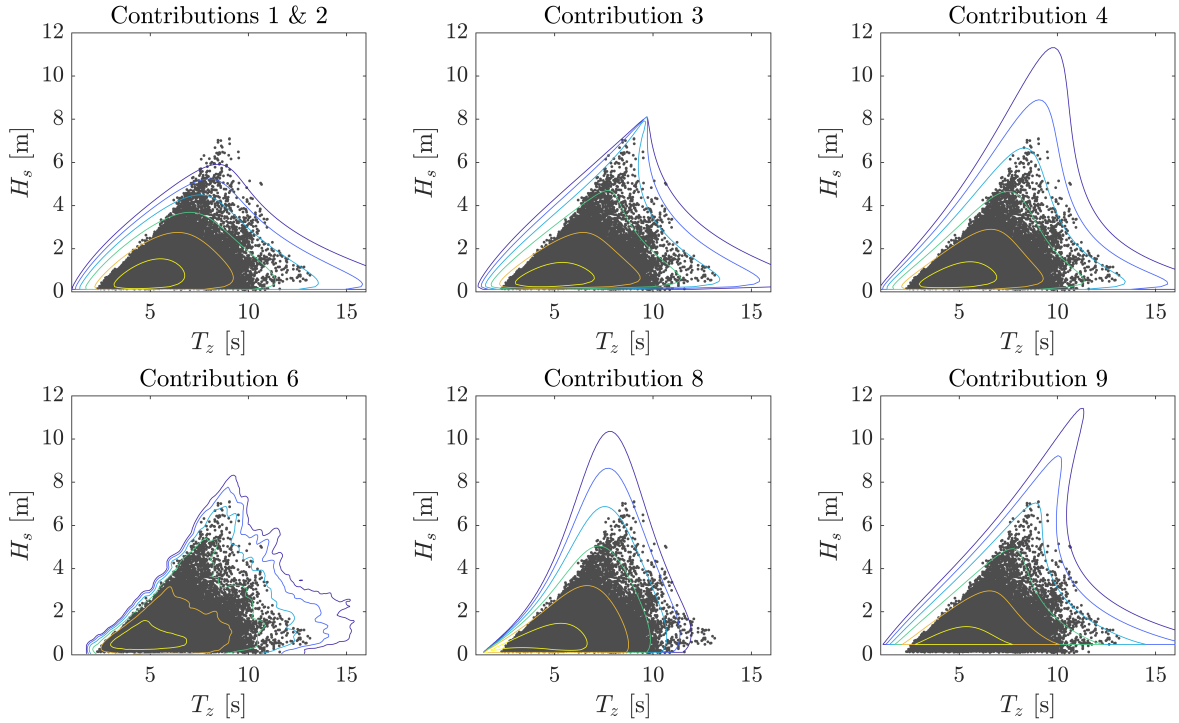


Figure 11: Scatter plot of observed H_s and T_z (dots), overlaid with isodensity contours of the fitted models for Dataset A. Contours shown at probability densities of 10^{-n} for $n = 1, \dots, 6$.

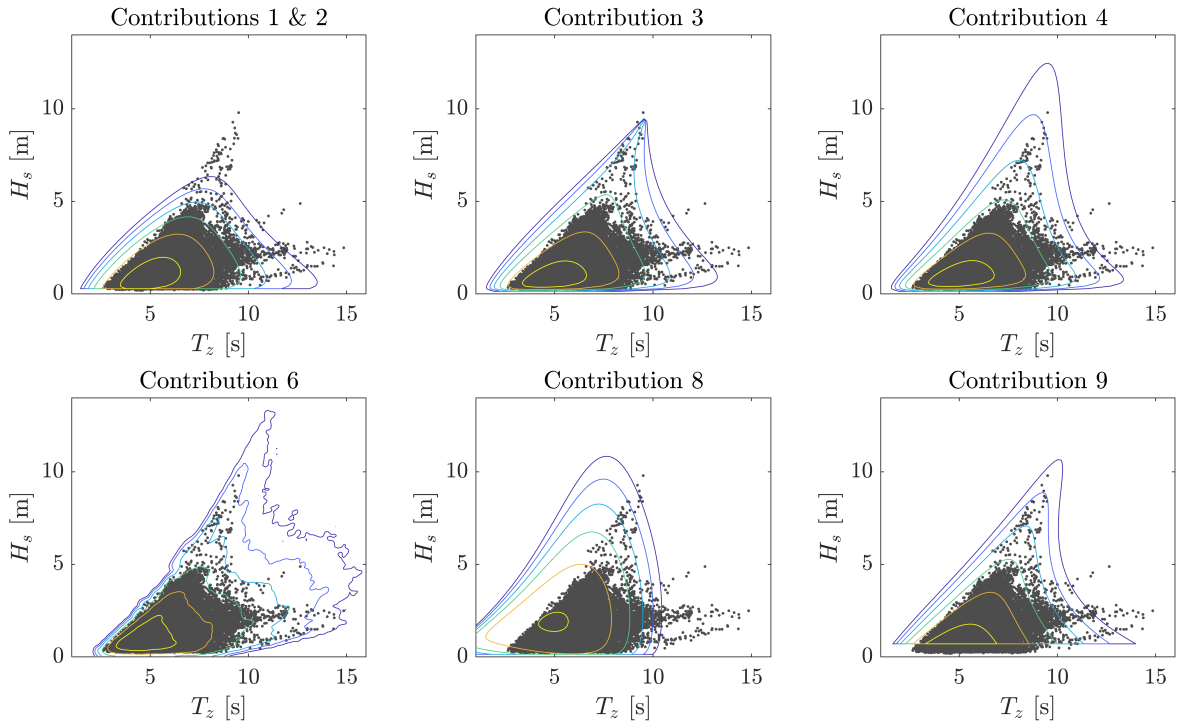


Figure 12: As previous figure, but for Dataset B.

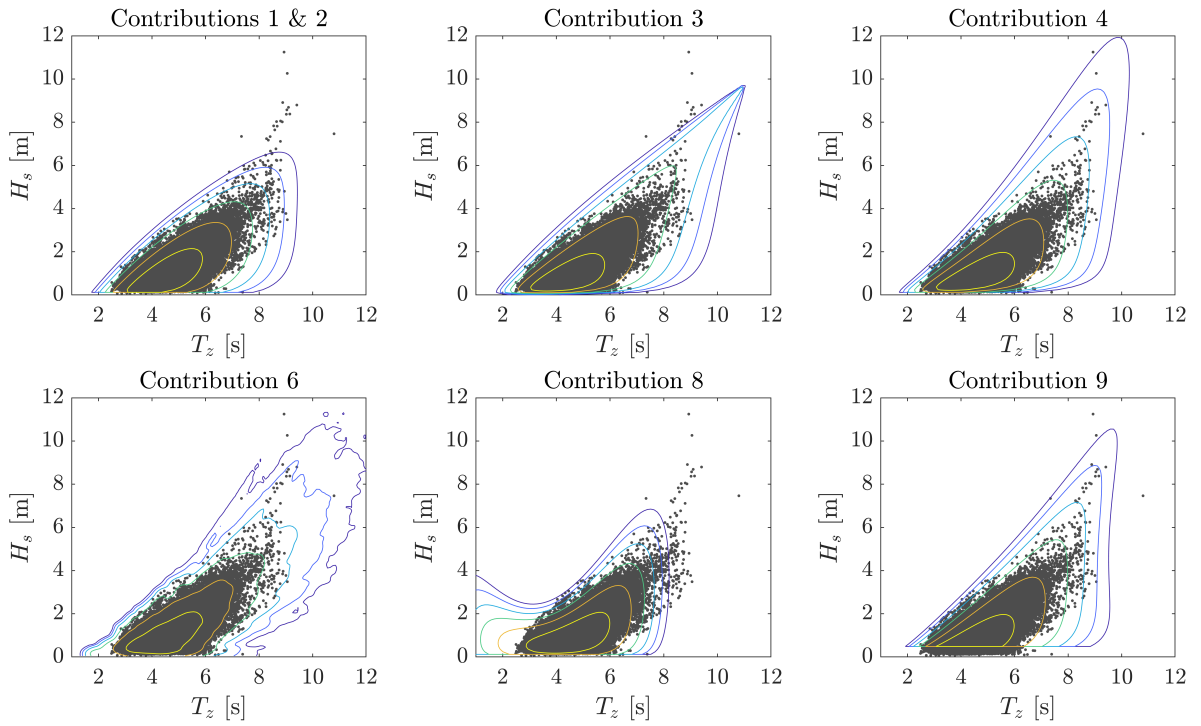


Figure 13: As previous figure, but for Dataset C.

Quantitative comparison of environmental contour approaches

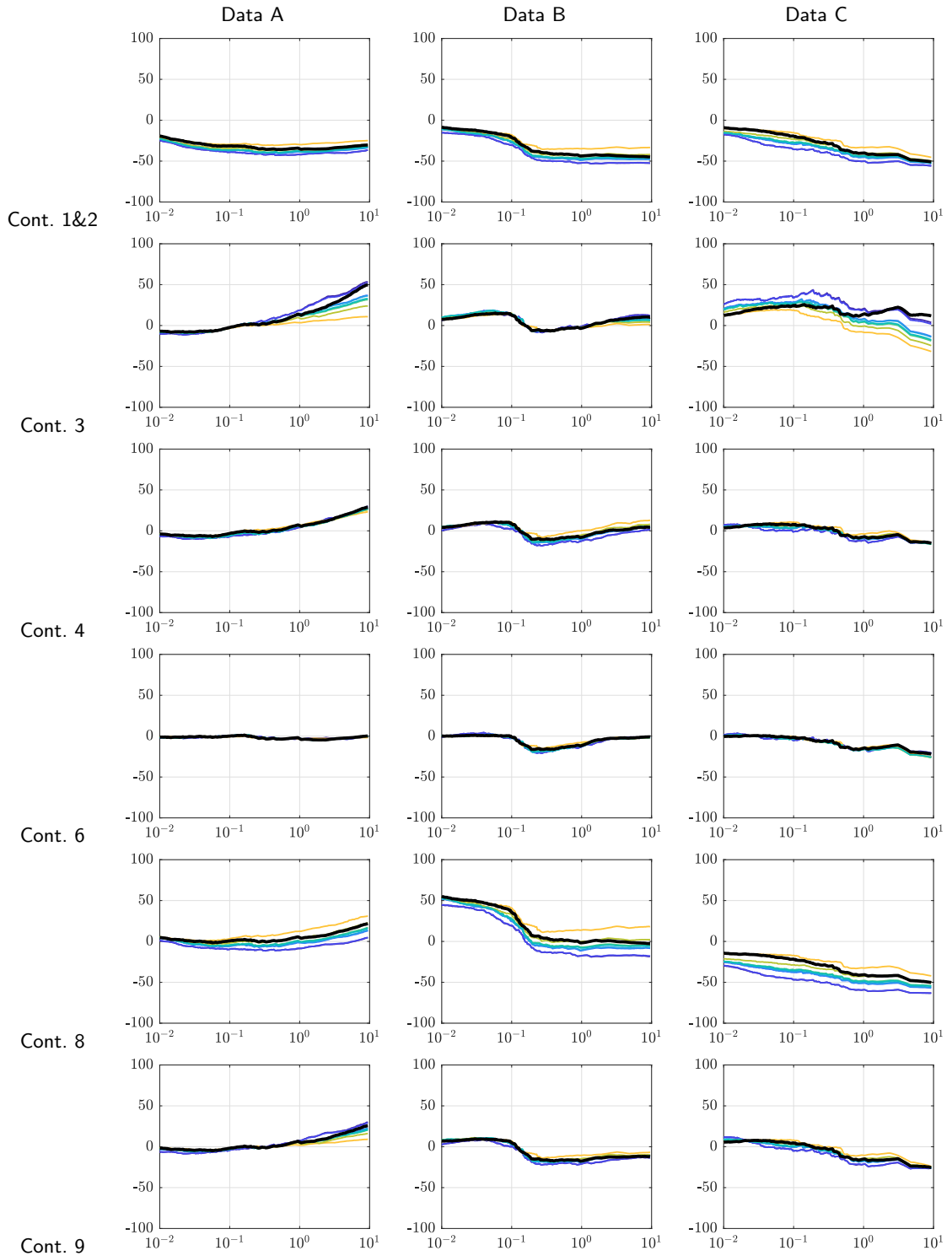


Figure 14: Percentage errors in return values of VBM from statistical models compared to observations against return period (years), for each of the 3 sites and each contribution. Colour of lines corresponds to RAOs shown in Figure 3a, with blue corresponding to response with highest natural period. Black lines indicate ratio of return values of H_s .

Quantitative comparison of environmental contour approaches

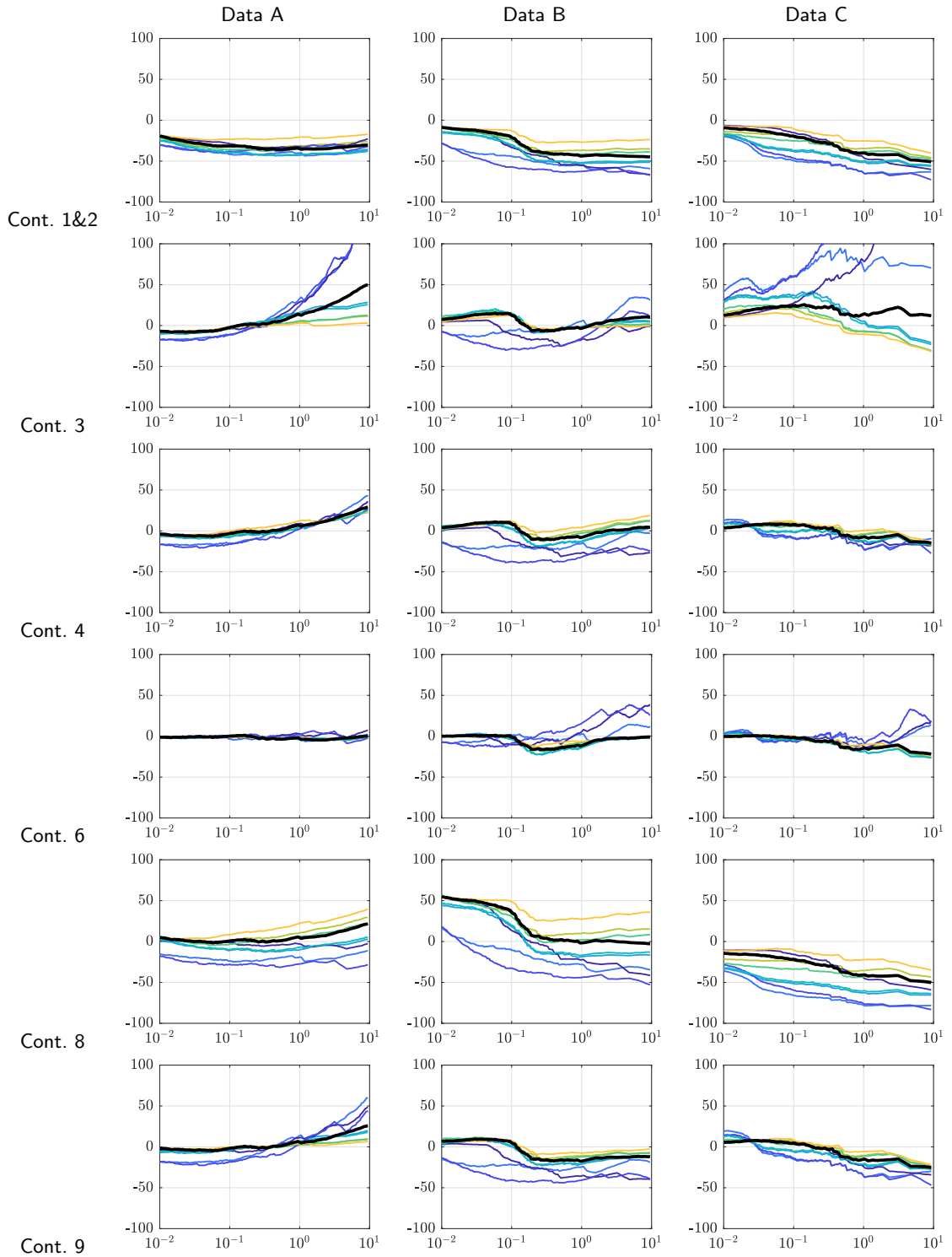


Figure 15: As previous figure, but for roll response.

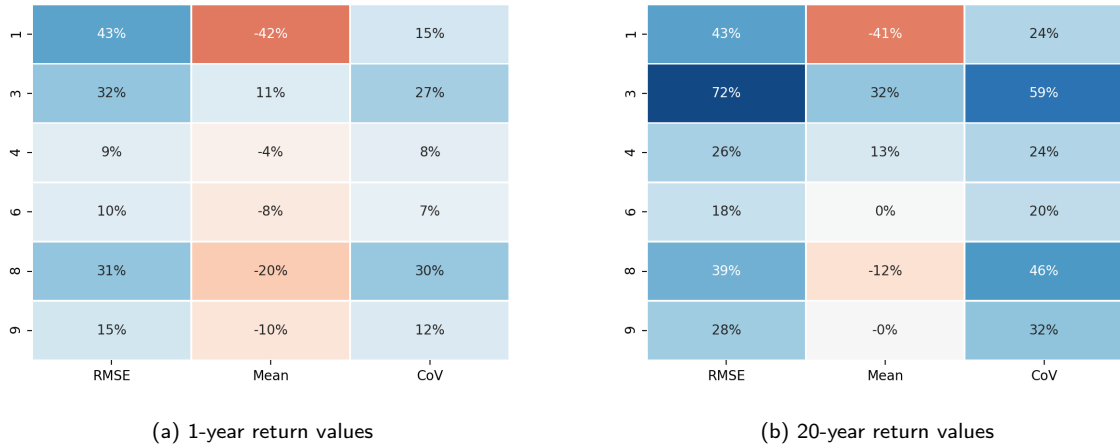


Figure 16: Error statistics from fitted statistical models compared to RBA under the assumption of IID hourly observations.

5.2. Contour construction

To isolate the effect of the contour construction method from other effects (errors in the statistical model and assumptions of independent observations), we compare response estimates from simulated data from a statistical model, under the assumption of IID hourly observations, to response estimates from various contours constructed from the same model. The statistical model user here is the model for dataset A, estimated in the baseline results [44]. We compare six types of contour, considered by the benchmark participants or in the literature:

- IFORM contour (linear failure surface in Gaussian space) : contribution 6, 7, 8, 9c
- IFORM calculated in (H_s, T_z) space : contribution 2, 9a
- IFORM calculated in $(H_s, steepness)$ space [53]
- IFORM calculated in $(H_s, H_s * T_z)$ space : contribution 5
- ISORM contour : contribution 1
- Highest density contour : contribution 4

The classic IFORM (in the gaussian space) is calculated analytically, while IFORM contours in physical space are numerically derived using Direct Sampling. The corresponding 20-year contours are shown in Figure 17. Responses return values estimated from these contours are compared to empirical return values estimated from 4000 years of simulated hourly data. The relative errors in the 1- and 20-year responses from the environmental contour approach compared to the direct Monte Carlo simulations are shown in Figure 18, and summarized in Figure 19. For most responses, the IFORM and direct sampling contours give relatively low errors, generally less than 5%. However, the direct sampling contour calculated in (H_s, T_z) space gives a large positive bias for the roll response of B30, C19 and C03, which have roll resonance periods of 15.6, 20.3 and 25.4s, respectively. The largest errors occur for C03, which has the longest period response, with a 47% positive bias in the 20-year return value from the contour. The errors are related to the shape of the joint distribution on the right-hand (long-period) side, which is concave (see Figure 11), whereas the direct sampling contour is always convex in the space in which it is defined. This causes the long-period response estimates from the direct sampling (H_s, T_z) contour to be positively biased.

Overall, the errors in the response estimates from the IFORM and direct sampling contours compared to the Monte Carlo simulations are small compared to the errors from the statistical model, discussed in the previous section.

The response estimates from the ISORM and HD contours exhibit positive biases for all responses, with slightly larger mean error of 19% for the 20-year ISORM estimates, compared to 16% for the 20-year HD contours estimate. Although ISORM and HD contours have the same total exceedance probability, ISORM contours are in fixed relation to the marginal quantiles of the conditioning variable used in the Rosenblatt transformation (H_s in this case), whereas

Quantitative comparison of environmental contour approaches

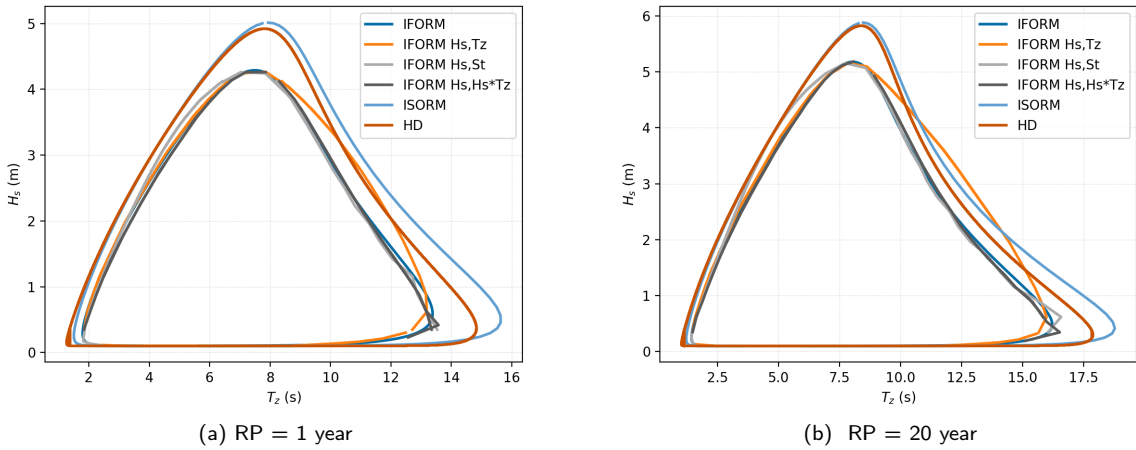


Figure 17: Contours calculated from the same joint distribution (IFORM contours in non-Gaussian space are calculated through direct sampling).

	IFORM	IFORM Hs,Tz	IFORM Hs,St	IFORM Hs,Hs*Tz	ISORM	HD		IFORM	IFORM Hs,Tz	IFORM Hs,St	IFORM Hs,Hs*Tz	ISORM	HD
G03_roll	4%	1%	3%	4%	22%	20%	G03_roll	3%	3%	3%	2%	15%	15%
R05_roll	3%	2%	5%	2%	15%	14%	R05_roll	2%	3%	6%	1%	11%	11%
T22_roll	4%	4%	2%	4%	27%	23%	T22_roll	2%	3%	2%	1%	19%	18%
B26_roll	4%	5%	2%	4%	28%	23%	B26_roll	2%	3%	2%	2%	20%	18%
C19_roll	2%	26%	-4%	0%	45%	26%	C19_roll	3%	42%	1%	3%	31%	19%
C03_roll	-7%	12%	-5%	-8%	52%	18%	C03_roll	4%	47%	-0%	-2%	51%	27%
B30_roll	2%	18%	-2%	4%	32%	22%	B30_roll	1%	22%	1%	4%	23%	18%
B22_roll	4%	3%	2%	3%	19%	18%	B22_roll	3%	1%	2%	2%	13%	13%
G03_VBM	2%	1%	2%	2%	17%	15%	G03_VBM	1%	1%	2%	1%	12%	11%
R05_VBM	2%	0%	1%	2%	20%	17%	R05_VBM	1%	1%	1%	1%	14%	13%
T22_VBM	2%	1%	0%	2%	22%	19%	T22_VBM	1%	-0%	-1%	0%	15%	14%
B26_VBM	2%	2%	1%	2%	23%	19%	B26_VBM	1%	1%	0%	-0%	16%	15%
C19_VBM	2%	1%	0%	2%	21%	19%	C19_VBM	1%	-0%	-1%	1%	15%	14%
C03_VBM	2%	1%	0%	2%	22%	19%	C03_VBM	1%	1%	0%	-0%	16%	15%
B30_VBM	2%	3%	0%	2%	25%	21%	B30_VBM	1%	2%	1%	1%	19%	17%
B22_VBM	2%	3%	-0%	2%	25%	21%	B22_VBM	1%	2%	1%	1%	19%	17%
tether	3%	2%	1%	2%	23%	20%	tether	2%	1%	1%	1%	17%	16%
Hs	-0%	-1%	-1%	-1%	17%	15%	Hs	-0%	-1%	-1%	-0%	13%	12%

(a) RP = 1 year

(b) RP = 20 year

Figure 18: Errors in response return values from contours relative to Monte Carlo simulations for each response (IFORM contours in non-Gaussian space are calculated through direct sampling).

for HD contours the relation with marginal return values is dependent on the shape of the joint distribution (see [47] for a discussion of this). In this case, as the gradient of the joint distribution is steeper for lower H_s , the HD contour is shifted toward the higher density region at the lower side of the marginal distribution of H_s and therefore the largest H_s on the HD contour is slightly lower than that for the ISORM contour, resulting in slightly lower biases in the estimated responses.

5.3. Serial correlation

5.3.1. Effect of serial correlation on benchmark results

Apart from contribution 5, all the contours submitted to the benchmarking exercise are based on a tacit assumption that hourly observations are independent. However, sea-state exhibit strong serial correlation, meaning that the assumption of independence is not correct. In extreme value theory, it is well-known that estimates of extremes

Quantitative comparison of environmental contour approaches

	RMSE	Mean	COV		RMSE	Mean	COV
IFORM -	3%	2%	2%	IFORM -	2%	2%	1%
IFORM Hs,Tz -	8%	5%	7%	IFORM Hs,Tz -	16%	7%	13%
IFORM Hs,St -	2%	0%	2%	IFORM Hs,St -	2%	1%	2%
IFORM Hs,Hs*Tz -	3%	2%	3%	IFORM Hs,Hs*Tz -	2%	1%	1%
ISORM -	27%	25%	8%	ISORM -	21%	19%	8%
HD -	20%	19%	3%	HD -	16%	16%	3%

(a) RP = 1 year

(b) RP = 20 years

Figure 19: Errors summary statistics for contour estimates relative to Monte Carlo simulations.

which neglect serial correlation and assume independence can lead to a positive bias in estimates of extremes (see e.g. Theorem 5.2 in [51]). The impact of neglecting serial correlation on estimates of extreme wave and crest heights was presented in [49] (this problem is directly analogous to the estimation of extreme responses, as the short-term wave or crest height distribution can be replaced with the short-term load distribution). In this section, we aim to quantify the error associated with assuming the sea states are IID, on estimates of extreme responses. Further discussion of this effect is provided in a separate paper [54].

To estimate the error resulting from assuming sea states are IID, we compare response estimates from RBA with and without declustering. The RBA with declustering is the one presented in section 3.2, where a threshold is fixed to the empirical 3 months values (based on decluster data). The RBA with IID assumption is the one used in section 5.1, where the threshold is set so that 50 events are retained. The GP distribution is then fitted to either (a) declustered threshold exceedances, or (b) all threshold exceedances. Return periods are calculated using (2) with M set as the expect number of either (a) declustered threshold exceedances, or (b) all threshold exceedances. The return values of H_s calculated from the two analyses are shown in Figure 20 for the three datasets.

For return periods of 1 year or less, the return values of H_s from the hourly data are significantly in excess of those from the declustered data. For dataset A, the difference at the 1-year level is over 1 m, whilst for datasets B and C the difference at the 1-year level is around 3 m. Note that by definition, the largest values in the declustered data is equal to the largest value of all observations. Therefore, the tails of the distributions coincide at the largest value. However, this is a finite sample size effect, and if a record length of more than 10 years was available, then the tails would not coincide exactly at the 10-year level. This is considered further in the next section. Here, we consider the effect of serial correlation on 1-year return values only.

Figure 21 shows the ratio between the declustered and IID 1-year return values for each response. For dataset A the IID return values are in the range 3-32% higher than the declustered values, whilst for datasets B and C, the effect is much larger, with over 100% relative errors in some cases. The difference between magnitude of the effect for the different sites and responses is related to the shape of the tail of the distribution, with longer-tailed distributions (a higher GP shape parameter) resulting in a larger bias from neglecting serial correlation [49].

Figure 22 compares the 1-year response estimates from the contours to responses from the RBA applied to hourly observations (assuming independence). The mean error is reduced for all contributions. For contribution 5, the mean error goes from 2% for the declustered RBA to -31% for the hourly RBA. Whereas for contribution 6, then mean error is reduced from 55% to 0%. This confirms that, for 1-year return values, assumptions made about serial correlation explain a large proportion of the mean errors observed in Figure 7. The effect for higher return periods is discussed in the next section.

5.3.2. Effect at higher return period with simulated time series

In order to estimate the effect of serial correlation at higher return periods, very long time series are required. Here, we use 1000 years of simulated time series, using the block-resampling method used to derive the joint distributions

Quantitative comparison of environmental contour approaches

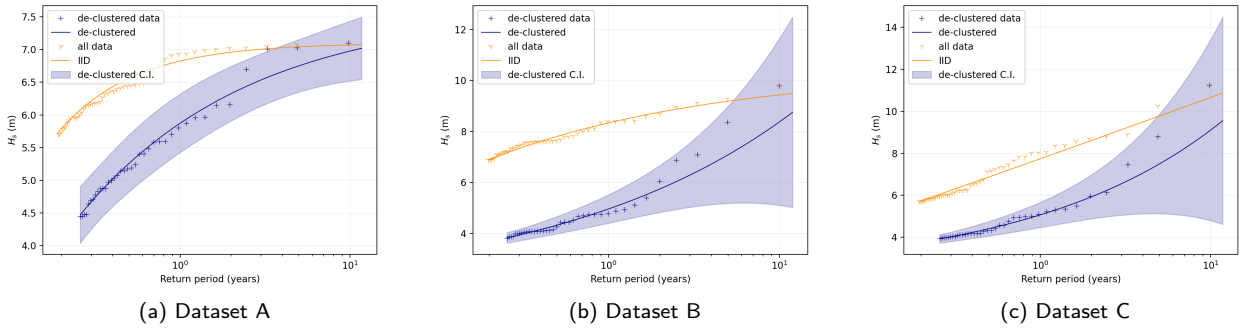


Figure 20: Return values of H_s estimated from declustered data and hourly observations under the assumption of independence.

G03_rol -	10%	62%	49%
R05_rol -	6%	33%	29%
T22_rol -	25%	105%	79%
B26_rol -	28%	108%	85%
C19_rol -	25%	81%	122%
C03_rol -	21%	89%	55%
B30_rol -	27%	93%	141%
B22_rol -	8%	53%	41%
G03_VBM -	11%	48%	37%
R05_VBM -	15%	65%	50%
T22_VBM -	19%	76%	57%
B26_VBM -	19%	79%	61%
C19_VBM -	18%	74%	56%
C03_VBM -	19%	79%	60%
B30_VBM -	25%	91%	72%
B22_VBM -	24%	91%	72%
tether	19%	81%	64%
Hs	16%	68%	53%
MEAN ERROR -	19%	76%	66%
	A	B	C

Figure 21: Effect of serial correlation on 1-year response

used for contribution 6, described in Section 2, with details available in [26]. Despite these not being ‘real’ time series, they are considered sufficiently realistic for the current purpose. As in the previous section, finite sample size effects mean that the largest value in the declustered data is equal to the largest hourly value. So this time we restrict our attention to return levels up to 100 years.

Figure 23 show the return values of H_s from the declustered data and calculated under the assumption of independent hourly values. For dataset A, there is relatively little difference in the 10-year and 100-year return levels, whereas for datasets B and C there is a difference of over 2 m in the 10- and 100-year return values.

The same procedure was applied to each response variable for the 1000-year time series, and the empirical return values and return periods were calculated using declustered data and hourly data. The results are summarised in Figure 24 and 25. The results are plotted in terms of the ratio of return values at a given return period and the ratio of return periods at a given declustered return period. For dataset A, the relative error in the return value is relatively small, whereas for datasets B and C, the effect is much larger, with mean errors in return values around 40% at the 10-year level and around 20-30% at the 100-year level. As mentioned above, this difference is related to the shape of the tail of the distribution.

It should be noted that the return value ratios shown here are for hourly observations, so may be expected to change if the time series discretization is modified. The effect of the time step will be investigated in future work.

Quantitative comparison of environmental contour approaches

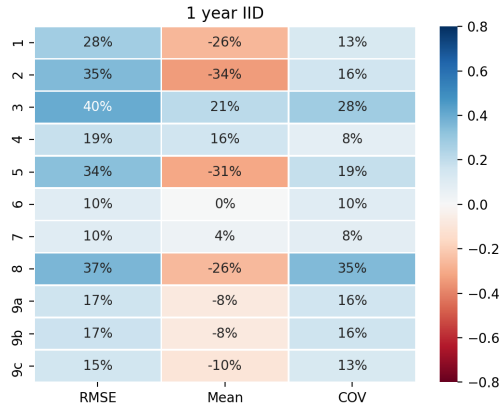


Figure 22: Errors statistics for contours compared to RBA under the assumption of independent hourly observations, RP = 1 year.

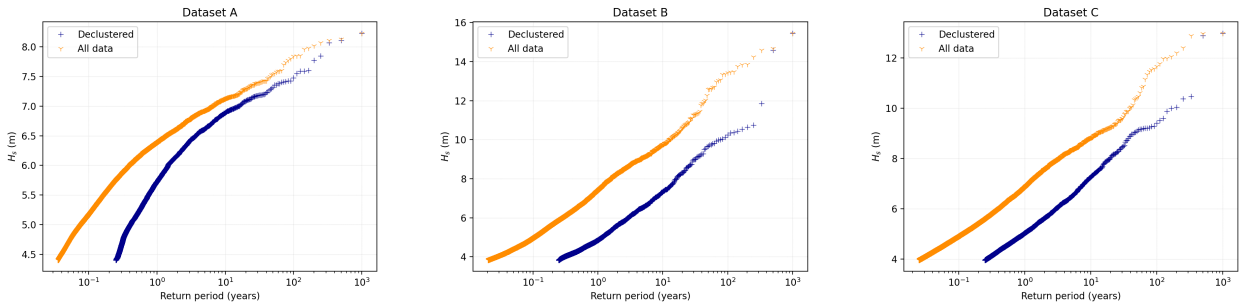
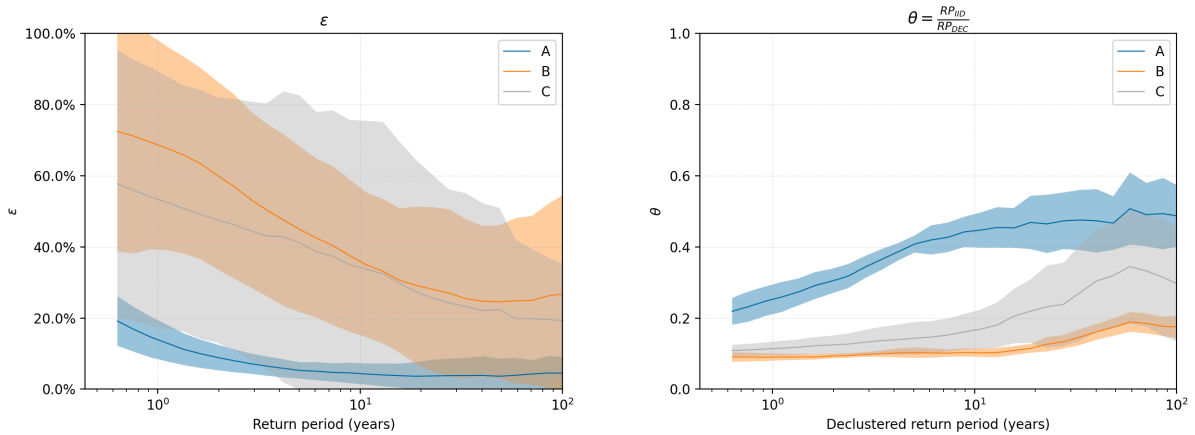


Figure 23: Return values of H_s for declustered data and under the assumption of independent hourly observations.



(a) Relative error in return value from hourly data to return value from declustered data (b) Ratio of return period from hourly data to return period from declustered data

Figure 24: Effect of serial correlation on errors in return values and return periods. Solid lines indicate mean values, shaded regions indicate plus/minus one standard deviation (taken over all responses).

Quantitative comparison of environmental contour approaches

G03_roll -	2%	11%	8%
R05_roll -	0%	8%	5%
T22_roll -	3%	20%	13%
B26_roll -	3%	22%	14%
C19_roll -	5%	76%	132%
C03_roll -	19%	93%	103%
B30_roll -	3%	57%	65%
B22_roll -	2%	12%	8%
G03_VBM -	2%	12%	7%
R05_VBM -	2%	17%	11%
T22_VBM -	3%	22%	12%
B26_VBM -	3%	25%	13%
C19_VBM -	3%	21%	12%
C03_VBM -	3%	24%	13%
B30_VBM -	3%	33%	18%
B22_VBM -	3%	34%	19%
tether -	3%	15%	10%
Hs -	3%	28%	15%
MEAN ERROR -	4%	29%	27%
	A	B	C

Figure 25: Effect of serial correlation on 20-year response, from extrapolated time series

6. Conclusions

In this work we have presented further analysis of the contributions submitted to the EC benchmarking exercise [44, 45]. The comparison of extreme responses calculated from the contours and RBA showed large errors in the estimates from the contours, with many showing strong positive biases in the range 30-90% in the 1-year return values. The results were more mixed at the 20-year level, with some contours exhibiting strong negative bias, while others had positive biases. The RMSE in the response estimates from the contours was in the range 11-46% at the 20-year level.

To investigate the reasons for the large differences in the results between the contour methods and various response types, the results were analysed further to isolate the influence of (1) the fit of the statistical model to the observations; (2) the method used to construct contours; and (3) assumptions made about serial correlation in the data.

The statistical models used for the benchmarking exercise varied in their fidelity to the data. Some models exhibited a poor match to observations, while others provided a close match to some datasets. None of the models used were a close match for all locations. A common feature of models which used a conditional log-normal model for wave period, was that they predicted unrealistically steep conditions sea states, with steepness far in excess of observed values. Comparisons of return values of responses calculated directly from simulated data from the statistical models with return values of responses calculated from the observations showed that a large portion of the errors in the response from the contours could be attributed to the fit of the statistical models. The errors were dependent on both the response function and the errors in the statistical model, with some models positively biased for some responses and negatively biased for others. The study highlights that ensuring the statistical model is a good fit to the data is an important factor in obtaining accurate response estimates from environmental contours.

Response estimates from different types of contour calculated from the same joint distribution model showed that contours assuming a linear iso-failure surface (e.g. IFORM, direct sampling and direct IFORM) give similar response estimates, whereas ISORM and highest density give positive biases for the response functions considered. This was in line with expectations, as the response functions considered gave failure regions that were close to convex at the closest point to the design region, so the assumptions about the shape of the failure region made in the definition of ISORM and HD contours were overly-conservative for these cases. Comparisons of 1-year and 20-year response estimates from the contours and direct Monte Carlo simulations from the joint distribution showed relatively little error for the IFORM and direct sampling contours (of the order of a few percent).

Most environmental contours are defined under the tacit assumption that there is no serial correlation in the data and each sea state is independent. It was shown that neglecting serial correlation leads to a positive bias in estimates of return values, the size of the bias dependent on the shape of the tail of the response distribution. Furthermore, the bias decreases with return period. Contour methods which account for serial correlation in the data exhibit much closer agreement with the RBA.

For the responses considered here, the method that performed best overall was the direct IFORM method (contribution 5). This method combines the features which minimise the three main sources of error considered here (a statistical model which is in good agreement with the data; approximating the failure surface as linear; and accounting for serial correlation). This method provided response estimates with RMSE below 10% compared to the RBA. Given the dramatic reduction in calculation time for contour-based estimates compared to RBA, this is considered a good result. However, this is not the full picture, since the effect of short-term variability has not been considered here. Various methods to account for short-term variability in contour-based estimates have been proposed in the literature (e.g. [5, 7, 8]).

Although only a few types of response have been considered in this study, we expect the conclusions to have a wider applicability. That is, the fit of the statistical model can cause large positive or negative biases, depending on the response type and the errors in the statistical model. Secondly, neglecting serial correlation can introduce a significant positive bias to estimates from contours (which is combined with either positive or negative bias from the statistical model). Finally, on the contour construction itself, IFORM (or its derivative like direct sampling and direct IFORM) was found to show better accuracy than ISORM or highest density contour on the responses investigated in this paper. This is believed to also be the case for most typical responses of marine structures. It is of course possible to find responses for which IFORM provides a non-conservative estimate, but this under-estimation is expected to be slight compared to the over-estimation linked to the ISORM or highest density approach. Thus, unless the response of interest is expected to have a very concave ISO-failure surface, the IFORM approach is recommended. For very highly non-linear responses, for instance for floating wind turbines, involving active controllers and threshold effects,

the approximations made in the construction of the contours may be less appropriate, and will be considered in future work.

Data used for the current study

All primary data used on this study (wave datasets, contours and response transfer functions) can be found on the benchmark github repository [45], available at : <https://github.com/ec-benchmark-organizers/ec-benchmark>

Acknowledgement

EM was funded by the EPSRC Supergen Offshore Renewable Energy Hub [grant no: EP/S000747/1] Flexible Fund project “Improved Models for Multivariate Metocean Extremes (IMEX)”.

The authors would like to thank all the contributors to the environmental contour benchmark exercise [45], who have published the necessary data to perform this further analysis.

References

- [1] International Electrotechnical Commission, “Wind energy generation systems - Part 3-1: Design requirements for fixed offshore wind turbines,” Tech. Rep. IEC 61400-3-1, 2019.
- [2] International Electrotechnical Commission (IEC), *Marine energy - Wave, tidal and other water current converters - Part 2: Design requirements for marine energy systems. IEC TS 62600-2:2016*. 2016.
- [3] I. O. for Standardization (ISO), “Petroleum and natural gas industries—specific requirements for offshore structures—part 1: Metocean design and operating considerations: Iso 19901-1:2015,” tech. rep., 2015.
- [4] NORSOK, *Actions and action effects (N-003), Edition 3*. Norway, January 2017, 2017.
- [5] DNV GL, *Environmental Conditions and Environmental Loads*. DNV GL, september 2019 ed., 2019. DNVGL-RP-C205.
- [6] Bureau Veritas, “NI 638 : Guidance for Long-term Hydro-structure Calculations,” 2019.
- [7] S. R. Winterstein and K. Engebretsen, “Reliability-based prediction of design loads and responses for floating ocean structures,” in *OMAE 1998: 17 th International Conference on Offshore Mechanics and Arctic Engineering*, p. 1998, 1998.
- [8] Q. Derbanne, G. de Hauteclouque, and M. Dumont, *How to Account for Short-Term and Long-Term Variability in the Prediction of the 100 Years Response?*, vol. Volume 3A: Structures, Safety and Reliability of *International Conference on Offshore Mechanics and Arctic Engineering*. June 2017.
- [9] S. Haver, “Wave climate off northern Norway,” *Applied Ocean Research*, vol. 7, no. 2, pp. 85–92, 1985.
- [10] S. Haver, “On the joint distribution of heights and periods of sea waves,” *Ocean Engineering*, vol. 14, no. 5, pp. 359–376, 1987.
- [11] J. Mathisen and E. Bitner-Gregersen, “Joint distributions for significant wave height and wave zero-up-crossing period,” *Applied Ocean Research*, vol. 12, pp. 93–103, 1990.
- [12] S. R. Winterstein, T. C. Ude, C. A. Cornell, P. Bjerager, and S. Haver, “Environmental parameters for extreme response: Inverse FORM with omission factors,” in *Proceedings of the 6th International Conference on Structural Safety & Reliability (ICOSSAR)*, vol. 93, (Innsbruck, Austria), pp. 77–84, International Association for Structural Safety and Reliability, August 1993.
- [13] L. Li, Z. Gao, and T. Moan, “Joint environmental data at five European offshore sites for design of combined wind and wave energy devices,” *Journal of Offshore Mechanics and Arctic Engineering*, vol. 137, pp. 031901–1 to 031901–16, 2015.
- [14] E. M. Bitner-Gregersen, “Joint met-ocean description for design and operations of marine structures,” *Applied Ocean Research*, vol. 51, pp. 279 – 292, 2015.
- [15] J.-T. Horn, E. Bitner-Gregersen, J. Krokstad, B. J. Leira, and J. Amdahl, “A new combination of conditional environmental distributions,” *Applied Ocean Research*, vol. 73, pp. 17–26, 2018.
- [16] A. Eckert-Gallup and N. Martin, “Kernel density estimation (KDE) with adaptive bandwidth selection for environmental contours of extreme sea states,” in *OCEANS 2016 MTS/IEEE Monterey*, pp. 1–5, IEEE, 2016.
- [17] A. Haselsteiner, J.-H. Ohlendorf, and K.-D. Thoben, “Environmental contours based on kernel density estimation,” in *Proceedings of the 13th German Wind Energy Conference DEWEK 2017*, (Bremen, Germany), 10 2017.
- [18] R. Montes-Iturrizaga and E. Heredia-Zavoni, “Environmental contours using copulas,” *Applied Ocean Research*, vol. 52, pp. 125–139, 2015.
- [19] F. Silva-González, E. Heredia-Zavoni, and R. Montes-Iturrizaga, “Development of environmental contours using Nataf distribution model,” *Ocean Engineering*, vol. 58, pp. 27 – 34, 2013.
- [20] E. Vanem, “Joint statistical models for significant wave height and wave period in a changing climate,” *Marine Structures*, vol. 49, pp. 180–205, 2016.
- [21] R. Montes-Iturrizaga and E. Heredia-Zavoni, “Multivariate environmental contours using C-vine copulas,” *Ocean Engineering*, vol. 118, pp. 68–82, 2016.
- [22] T. Fazerer-Ferradosa, F. Taveira-Pinto, E. Vanem, M. T. Reis, and L. das Neves, “Asymmetric copula-based distribution models for met-ocean data in offshore wind engineering applications,” *Wind Engineering*, vol. 42, pp. 304–334, 2018.
- [23] P. Jonathan, J. Flynn, and K. Ewans, “Joint modelling of wave spectral parameters for extreme sea states,” *Ocean Engineering*, vol. 37, pp. 1070–1080, 2010.

- [24] P. Jonathan, K. Ewans, and J. Flynn, "On the estimation of ocean engineering design contours," *Journal of Offshore Mechanics and Arctic Engineering*, vol. 136, no. 4, pp. 41101–1 to 041101–8, 2014.
- [25] H. F. Hansen, D. Randell, A. R. Zeeberg, and P. Jonathan, "Directional–seasonal extreme value analysis of North Sea storm conditions," *Ocean Engineering*, vol. 195, p. 106665, 2020.
- [26] E. Mackay and P. Jonathan, "Estimation of environmental contours using a block resampling method," in *Proc. 39th International Conference on Ocean, Offshore and Arctic Engineering (OMAE 2020)*, American Society of Mechanical Engineers (ASME), 2020.
- [27] S. Haver and S. R. Winterstein, "Environmental contour lines: A method for estimating long term extremes by a short term analysis," *Transactions of the Society of Naval Architects and Marine Engineers*, vol. 116, pp. 116–127, 2009.
- [28] A. B. Huseby, E. Vanem, and B. Natvig, "A new approach to environmental contours for ocean engineering applications based on direct Monte Carlo simulations," *Ocean Engineering*, vol. 60, pp. 124–135, 2013.
- [29] A. B. Huseby, E. Vanem, and B. Natvig, "Alternative environmental contours for structural reliability analysis," *Structural Safety*, vol. 54, pp. 32 – 45, 2015.
- [30] Q. Derbanne and G. de Hauteclocque, "A new approach for environmental contour and multivariate de-clustering," vol. Volume 3: Structures, Safety, and Reliability of *International Conference on Offshore Mechanics and Arctic Engineering*, 06 2019.
- [31] A. F. Haselsteiner, J.-H. Ohlendorf, W. Wosniok, and K.-D. Thoben, "Deriving environmental contours from highest density regions," *Coastal Engineering*, vol. 123, pp. 42–51, 2017.
- [32] W. Chai and B. J. Leira, "Environmental contours based on inverse SORM," *Marine Structures*, vol. 60, pp. 34 – 51, 2018.
- [33] N. Dimitrov, "Inverse directional simulation: An environmental contour method providing an exact return period," *Journal of Physics: Conference Series*, 2020.
- [34] F. Silva-González, A. Vázquez-Hernández, L. Sagrilo, and R. Cuamatzi, "The effect of some uncertainties associated to the environmental contour lines definition on the extreme response of an FPSO under hurricane conditions," *Applied Ocean Research*, vol. 53, pp. 190–199, 2015.
- [35] E. Vanem, O. Gramstad, and E. M. Bitner-Gregersen, "A simulation study on the uncertainty of environmental contours due to sampling variability for different estimation methods," *Applied Ocean Research*, vol. 91, pp. 1–15, 2019.
- [36] C. Armstrong, C. Chin, I. Penesis, and Y. Drobyshevski, "Sensitivity of vessel response to environmental contours of extreme sea states," in *Proc. 34th International Conference on Ocean, Offshore and Arctic Engineering (OMAE 2015)*, American Society of Mechanical Engineers (ASME), May–June 2015.
- [37] E. Vanem, "Non-stationary extreme value models to account for trends and shifts in the extreme wave climate due to climate change," *Applied Ocean Research*, vol. 52, pp. 201–211, aug 2015.
- [38] L. Manuel, P. T. Nguyen, J. Canning, R. G. Coe, A. C. Eckert-Gallup, and N. Martin, "Alternative approaches to develop environmental contours from metocean data," *Journal of Ocean Engineering and Marine Energy*, vol. 4, no. 4, pp. 293–310, 2018.
- [39] A. B. Huseby, E. Vanem, and K. Eskeland, "Evaluating properties of environmental contours," in *Proc. ESREL 2017*, European Safety and Reliability Association (ESRA), June 2017.
- [40] E. Vanem, "A comparison study on the estimation of extreme structural response from different environmental contour methods," *Marine Structures*, vol. 56, pp. 137–162, 2017.
- [41] E. Ross, O. C. Astrup, E. Bitner-Gregersen, N. Bunn, G. Feld, B. Gouldby, A. Huseby, Y. Liu, D. Randell, E. Vanem, and P. Jonathan, "On Environmental Contours for Marine and Coastal Design," in *ASME 2019 38th International Conference on Ocean, Offshore and Arctic Engineering*, American Society of Mechanical Engineers Digital Collection, November 2019.
- [42] E. Vanem, B. Guo, E. Ross, and P. Jonathan, "Comparing different contour methods with response-based methods for extreme ship response analysis," *Marine Structures*, vol. 69, 2020.
- [43] O. Gramstad, C. Agrell, E. Bitner-Gregersen, B. Guo, E. Ruth, and E. Vanem, "Sequential sampling method using Gaussian process regression for estimating extreme structural response," *Marine Structures*, vol. 72, pp. 102780:1–14, 2020.
- [44] A. F. Haselsteiner, R. G. Coe, L. Manuel, P. T. T. Nguyen, N. Martin, and A. Eckert-Gallup, "A Benchmarking Exercise on Estimating Extreme Environmental Conditions: Methodology and Baseline Results," in *OMAE2019*, (Volume 3: Structures, Safety, and Reliability), June 2019.
- [45] A. Haselsteiner et al., "A benchmarking exercise for environmental contours," *Submitted to Ocean Engineering*, 2021.
- [46] E. Vanem and A. B. Huseby, "Environmental contours based on a direct sampling approach and the IFORM approach: Contribution to a benchmark study," in *Proc. 39th International Conference on Ocean, Offshore and Arctic Engineering (OMAE2020)*, American Society of Mechanical Engineers (ASME), August 2020.
- [47] E. Mackay and A. F. Haselsteiner, "Marginal and total exceedance probabilities of environmental contours," *Marine Structures*, vol. 75, 2021.
- [48] E. Mackay and L. Johanning, "A simple and robust method for calculating return periods of ocean waves," in *Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering - OMAE*, vol. 11B, American Society of Mechanical Engineers (ASME), 2018.
- [49] E. Mackay and L. Johanning, "Long-term distributions of individual wave and crest heights," *Ocean Engineering*, vol. 165, pp. 164–183, 2018.
- [50] X. Chen, "Hydrodynamics in offshore and naval applications - Part I," *Keynote lecture of 6th Intl. Conf. HydroDynamics, Perth (Australia)*, 2004.
- [51] S. Coles, *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics, London: Springer-Verlag, 2001.
- [52] P. Jonathan, D. Randell, J. Wadsworth, and J. Tawn, "Uncertainties in return values from extreme value analysis of peaks over threshold using the generalised Pareto distribution," *Ocean Engineering*, vol. 220, p. 107725, 2021.
- [53] Q. Derbanne and G. de Hauteclocque, *A New Approach for Environmental Contour and Multivariate De-Clustering*, vol. Volume 3: Structures, Safety, and Reliability of *International Conference on Offshore Mechanics and Arctic Engineering*, June 2019.
- [54] E. Mackay, G. de Hauteclocque, E. Vanem, and P. Jonathan, "The Effect of Serial Correlation in Environmental Conditions on Estimates of Extreme Events," *Submitted to this special issue*.
- [55] S. Winterstein, "Environmental Contours: Including the Effects of Directionality and other Sub-Populations," tech. rep., 2017.

A. Tether response model

The response model for the tether tension is taken from [55]. The tether tension σ_x is composed of first and second-order contribution, which are assumed uncorrelated:

$$\sigma_x^2 = \sigma_{x1}^2 + \sigma_{x2}^2, \quad (7)$$

where

$$\sigma_{x1} = L(T_p) * H_s, \quad (8)$$

$$\sigma_{x2} = Q(T_p) * H_s^2, \quad (9)$$

and T_p is the spectral peak period. The linear component is the sum of two terms:

$$L(T_p) = L_1(T_p) + L_2(T_p), \quad (10)$$

where

$$L_1(T_p) = \frac{L_\infty}{1 + \left(\frac{1}{L_0} - 1\right) \exp(-rT_p)}, \quad (11)$$

$$L_2(T_p) = b_0 \exp \left[- \left(\frac{T_p - 12}{\sigma} \right)^2 \right]. \quad (12)$$

The quadratic component is given by

$$Q(T_p) = \frac{Q_0}{\sqrt{[1 - (T_p/T_m)^2]^2 + (2\eta T_p/T_m)^2}} \quad (13)$$

For the Snorre TLP considered, the coefficients values are $L_\infty = 5.7$, $L_0 = 0.002$, $r = 1$, $b_0 = 1.3$, $\sigma = 2$, $Q_0 = 0.36$, $T_m = 5$ and $\eta = 0.1$.

B. Detailed results

Quantitative comparison of environmental contour approaches

Dataset	Response	1	2	3	4	5	6	7	8	9a	9b	9c	
A	G03 roll	-10%	-18%	17%	37%	1%	9%	6%	30%	13%	13%	13%	
	R05 roll	-2%	-7%	12%	33%	0%	6%	3%	39%	11%	12%	9%	
	T22 roll	-8%	-19%	43%	56%	1%	21%	20%	23%	28%	28%	29%	
	B26 roll	-7%	-17%	49%	60%	3%	23%	23%	23%	31%	31%	32%	
	C19 roll	25%	25%	72%	60%	1%	36%	63%	-11%	87%	90%	13%	
	C03 roll	30%	21%	68%	58%	-6%	21%	41%	16%	62%	66%	23%	
	B30 roll	-0%	-4%	85%	73%	3%	31%	49%	5%	58%	59%	31%	
	B22 roll	-10%	-16%	15%	33%	-1%	6%	3%	33%	10%	10%	10%	
	G03 VBM	-8%	-14%	18%	36%	1%	9%	6%	35%	13%	13%	13%	
	R05 VBM	-11%	-18%	27%	42%	1%	12%	9%	29%	16%	16%	17%	
	T22 VBM	-10%	-20%	38%	49%	2%	17%	14%	28%	22%	22%	23%	
	B26 VBM	-10%	-21%	41%	51%	1%	17%	15%	26%	23%	23%	24%	
	C19 VBM	-10%	-20%	36%	48%	2%	16%	13%	28%	21%	21%	22%	
	C03 VBM	-10%	-21%	41%	50%	1%	17%	15%	27%	23%	23%	24%	
	B30 VBM	-8%	-18%	56%	59%	2%	23%	23%	23%	30%	30%	31%	
	B22 VBM	-8%	-18%	58%	60%	2%	23%	24%	23%	31%	31%	31%	
	tether	-10%	-22%	30%	48%	1%	16%	13%	26%	21%	21%	21%	
	Hs	-12%	-19%	36%	44%	-1%	12%	10%	27%	18%	17%	18%	
	B	G03 roll	18%	10%	66%	85%	1%	57%	64%	16%	43%	43%	44%
		R05 roll	14%	6%	36%	58%	-1%	32%	37%	13%	25%	25%	24%
T22 roll		30%	19%	130%	125%	15%	104%	117%	22%	69%	69%	69%	
B26 roll		30%	20%	140%	128%	16%	109%	124%	22%	71%	71%	70%	
C19 roll		2%	-6%	118%	57%	-0%	165%	129%	-31%	20%	23%	-1%	
C03 roll		-3%	-10%	109%	66%	-11%	145%	96%	-4%	20%	20%	18%	
B30 roll		16%	7%	172%	109%	9%	91%	149%	-5%	54%	54%	37%	
B22 roll		16%	6%	58%	76%	5%	48%	55%	16%	36%	36%	36%	
G03 VBM		12%	2%	52%	68%	1%	43%	49%	12%	30%	30%	30%	
R05 VBM		14%	6%	78%	84%	0%	61%	69%	12%	40%	40%	41%	
T22 VBM		15%	7%	95%	94%	-2%	72%	82%	12%	46%	45%	46%	
B26 VBM		16%	7%	104%	97%	0%	76%	88%	12%	48%	48%	48%	
C19 VBM		15%	6%	93%	92%	-3%	70%	80%	12%	45%	45%	46%	
C03 VBM		16%	7%	102%	97%	-0%	75%	86%	12%	47%	47%	48%	
B30 VBM		19%	9%	130%	109%	6%	90%	107%	12%	56%	55%	55%	
B22 VBM		19%	9%	132%	109%	5%	90%	107%	12%	55%	55%	55%	
tether		20%	11%	93%	103%	2%	78%	87%	16%	52%	52%	53%	
Hs		12%	4%	91%	84%	0%	64%	74%	11%	39%	39%	40%	
C		G03 roll	17%	5%	43%	65%	3%	42%	47%	-9%	37%	37%	38%
		R05 roll	10%	3%	24%	44%	-0%	25%	29%	3%	23%	23%	23%
	T22 roll	28%	4%	89%	100%	2%	69%	84%	-25%	50%	50%	50%	
	B26 roll	31%	5%	101%	106%	2%	74%	93%	-26%	53%	53%	53%	
	C19 roll	48%	4%	576%	177%	42%	196%	177%	-36%	66%	66%	62%	
	C03 roll	8%	-3%	245%	79%	-6%	67%	74%	-3%	27%	26%	27%	
	B30 roll	63%	15%	426%	197%	23%	185%	192%	-39%	83%	83%	79%	
	B22 roll	14%	6%	37%	57%	1%	36%	40%	-1%	32%	32%	32%	
	G03 VBM	10%	3%	33%	53%	-1%	32%	36%	-1%	28%	28%	28%	
	R05 VBM	13%	1%	55%	67%	-0%	43%	50%	-8%	34%	34%	35%	
	T22 VBM	16%	1%	74%	76%	-1%	50%	60%	-13%	37%	37%	38%	
	B26 VBM	17%	1%	84%	79%	-2%	52%	65%	-16%	38%	38%	39%	
	C19 VBM	16%	1%	71%	75%	-1%	49%	59%	-12%	37%	37%	38%	
	C03 VBM	17%	1%	82%	79%	-2%	52%	64%	-15%	38%	38%	38%	
	B30 VBM	21%	-0%	119%	90%	-3%	61%	81%	-25%	42%	42%	42%	
	B22 VBM	21%	-0%	123%	90%	-4%	61%	81%	-25%	42%	41%	42%	
	tether	21%	5%	68%	84%	1%	56%	65%	-12%	43%	43%	44%	
	Hs	14%	2%	88%	71%	-1%	45%	57%	-4%	34%	34%	35%	

Table 4
Relative errors for RP = 1 year

Quantitative comparison of environmental contour approaches

Dataset	Response	1	2	3	4	5	6	7	8	9a	9b	9c	
A	G03 roll	-13%	-27%	3%	46%	1%	4%	1%	46%	14%	16%	13%	
	R05 roll	-6%	-15%	5%	43%	1%	3%	-0%	57%	14%	17%	10%	
	T22 roll	-18%	-36%	10%	58%	0%	7%	7%	24%	23%	23%	23%	
	B26 roll	-19%	-36%	12%	59%	0%	6%	8%	21%	25%	25%	25%	
	C19 roll	-0%	-3%	18%	98%	1%	29%	96%	-21%	142%	147%	72%	
	C03 roll	27%	8%	42%	86%	1%	43%	99%	-2%	178%	179%	61%	
	B30 roll	-21%	-28%	21%	93%	1%	6%	46%	-7%	69%	72%	66%	
	G02 roll	-12%	-23%	3%	47%	1%	5%	1%	51%	14%	16%	14%	
	G03 VBM	-13%	-24%	2%	47%	-0%	4%	0%	49%	14%	15%	13%	
	R05 VBM	-16%	-29%	7%	53%	0%	6%	4%	40%	21%	20%	20%	
	T22 VBM	-18%	-32%	10%	56%	-1%	6%	7%	32%	25%	25%	25%	
	B26 VBM	-18%	-33%	12%	59%	-1%	6%	9%	30%	28%	28%	28%	
	C19 VBM	-18%	-32%	10%	56%	-1%	6%	6%	33%	24%	24%	24%	
	C03 VBM	-19%	-33%	12%	59%	-1%	6%	9%	30%	28%	27%	27%	
	B30 VBM	-20%	-34%	17%	69%	-0%	5%	16%	20%	37%	37%	37%	
	B22 VBM	-20%	-34%	17%	70%	-1%	5%	17%	20%	39%	38%	38%	
	tether	-17%	-31%	5%	52%	0%	6%	3%	33%	17%	18%	17%	
	Hs	-18%	-31%	14%	64%	-0%	7%	13%	36%	34%	33%	33%	
	B	G03 roll	-30%	-39%	-11%	24%	3%	1%	6%	-25%	-9%	-10%	-10%
		R05 roll	-14%	-20%	-1%	35%	11%	5%	12%	-8%	-0%	1%	-2%
T22 roll		-48%	-55%	-21%	9%	-11%	-9%	-5%	-47%	-22%	-22%	-22%	
B26 roll		-47%	-55%	-18%	12%	-9%	-5%	-0%	-47%	-20%	-20%	-20%	
C19 roll		-57%	-62%	-31%	7%	-4%	56%	116%	-68%	-23%	-22%	-36%	
C03 roll		-78%	-81%	-60%	-42%	-40%	21%	22%	-79%	-60%	-59%	-63%	
B30 roll		-50%	-57%	-6%	35%	8%	50%	85%	-54%	-9%	-9%	-13%	
B22 roll		-31%	-36%	-14%	20%	-1%	-3%	2%	-25%	-13%	-13%	-14%	
G03 VBM		-24%	-29%	-5%	32%	9%	6%	12%	-18%	-5%	-5%	-5%	
R05 VBM		-36%	-44%	-12%	22%	-0%	3%	8%	-32%	-13%	-13%	-13%	
T22 VBM		-37%	-46%	-7%	27%	3%	10%	18%	-35%	-9%	-10%	-10%	
B26 VBM		-38%	-47%	-6%	28%	2%	12%	21%	-36%	-9%	-9%	-9%	
C19 VBM		-37%	-46%	-8%	26%	2%	9%	16%	-34%	-10%	-10%	-10%	
C03 VBM		-39%	-48%	-8%	26%	1%	10%	18%	-37%	-11%	-11%	-11%	
B30 VBM		-47%	-55%	-13%	18%	-9%	8%	20%	-47%	-17%	-17%	-17%	
B22 VBM		-48%	-56%	-14%	17%	-10%	8%	20%	-48%	-18%	-18%	-18%	
tether		-39%	-48%	-14%	19%	-2%	-3%	9%	-37%	-15%	-15%	-15%	
Hs		-37%	-43%	-5%	30%	2%	19%	37%	-33%	-9%	-9%	-9%	
C		G03 roll	-34%	-42%	-24%	4%	5%	-14%	-5%	-40%	-15%	-15%	-15%
		R05 roll	-26%	-30%	-19%	8%	7%	-12%	-2%	-25%	-10%	-10%	-11%
	T22 roll	-43%	-59%	-28%	-1%	1%	-18%	-9%	-61%	-22%	-23%	-22%	
	B26 roll	-46%	-62%	-31%	-5%	-3%	-20%	-11%	-65%	-26%	-26%	-26%	
	C19 roll	-73%	-86%	-20%	-23%	-10%	-8%	85%	-88%	-61%	-61%	-62%	
	C03 roll	-50%	-61%	19%	20%	32%	37%	201%	-59%	-30%	-30%	-30%	
	B30 roll	-80%	-89%	-56%	-52%	-51%	-53%	-33%	-92%	-72%	-72%	-72%	
	B22 roll	-34%	-41%	-25%	3%	4%	-15%	-6%	-35%	-16%	-16%	-16%	
	G03 VBM	-33%	-39%	-23%	6%	7%	-13%	-3%	-33%	-14%	-14%	-14%	
	R05 VBM	-38%	-47%	-24%	4%	7%	-13%	-2%	-44%	-17%	-17%	-17%	
	T22 VBM	-39%	-50%	-20%	7%	9%	-9%	3%	-49%	-17%	-17%	-17%	
	B26 VBM	-38%	-51%	-17%	10%	12%	-6%	8%	-50%	-16%	-16%	-16%	
	C19 VBM	-40%	-50%	-22%	5%	7%	-12%	1%	-49%	-18%	-19%	-19%	
	C03 VBM	-39%	-51%	-18%	8%	11%	-8%	6%	-50%	-17%	-17%	-17%	
	B30 VBM	-42%	-58%	-14%	10%	11%	-4%	16%	-59%	-19%	-20%	-20%	
	B22 VBM	-42%	-58%	-14%	10%	11%	-3%	18%	-60%	-20%	-20%	-20%	
	tether	-43%	-54%	-32%	-5%	-4%	-22%	-3%	-54%	-24%	-24%	-24%	
	Hs	-42%	-51%	-18%	6%	7%	-7%	21%	-47%	-21%	-21%	-21%	

Table 5
Relative errors for RP = 20 years