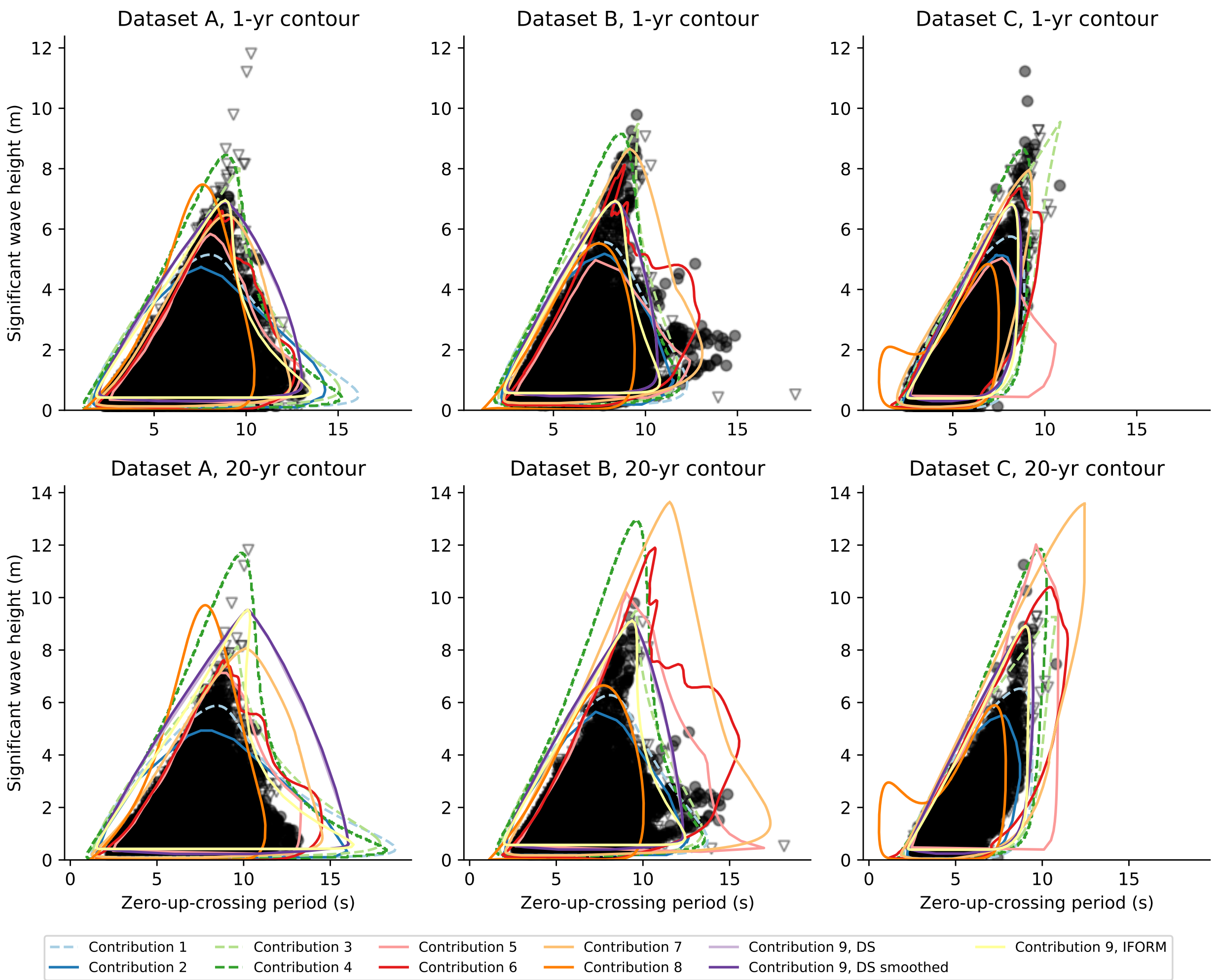
Graphical Abstract

A benchmarking exercise for environmental contours, preprint January 19, 2021

Andreas F. Haselsteiner,Ryan G. Coe,Lance Manuel,Wei Chai,Bernt Leira,Guilherme Clarindo,C. Guedes Soares,Ásta Hannesdóttir,Nikolay Dimitrov,Aljoscha Sander,Jan-Hendrik Ohlendorf,Klaus-Dieter Thoben,Guillaume de Hauteclocque,Ed Mackay,Philip Jonathan,Chi Qiao,Andrew Myers,Anna Rode,Arndt Hildebrandt,Boso Schmidt,Erik Vanem,Arne Bang Huseby



Highlights

A benchmarking exercise for environmental contours, preprint January 19, 2021

Andreas F. Haselsteiner,Ryan G. Coe,Lance Manuel,Wei Chai,Bernt Leira,Guilherme Clarindo,C. Guedes Soares,Ásta Hannesdóttir,Nikolay Dimitrov,Aljoscha Sander,Jan-Hendrik Ohlendorf,Klaus-Dieter Thoben,Guillaume de Hauteclocque,Ed Mackay,Philip Jonathan,Chi Qiao,Andrew Myers,Anna Rode,Arndt Hildebrandt,Boso Schmidt,Erik Vanem,Arne Bang Huseby

- Nine teams participated in a comparison of environmental contour methods
- Maximum values along some contours deviated by a factor of two
- Number of points outside the contours deviated by an order of magnitude among the submissions
- Differences due to the use of different models for the joint distribution were greater than differences due to different contour construction methods

A benchmarking exercise for environmental contours, preprint January 19, 2021

Andreas F. Haselsteiner[a,*], Ryan G. Coe[b], Lance Manuel[c], Wei Chai[d], Bernt Leira[e], Guilherme Clarindo[f], C. Guedes Soares[f], Ásta Hannesdóttir[g], Nikolay Dimitrov[g], Aljoscha Sander[a], Jan-Hendrik Ohlendorf[a], Klaus-Dieter Thoben[a], Guillaume de Hauteclocque[h], Ed Mackay[i], Philip Jonathan[j], Chi Qiao[k], Andrew Myers[k], Anna Rode[l], Arndt Hildebrandt[l], Boso Schmidt[l], Erik Vanem[m,n] and Arne Bang Huseby[n]

[a] *University of Bremen, Bremen, Germany*
[b] *Sandia National Labs, Albuquerque, NM, USA*
[c] *University of Texas at Austin, Austin, TX, USA*
[d] *School of Transportation, Wuhan University of Technology, Wuhan, China*
[e] *Department of Marine Technology, Norwegian University of Science and Technology, Trondheim, Norway*
[f] *Centre for Marine Technology and Ocean Engineering (CENTEC), Instituto Superior Técnico Lisboa, Universidade de Lisboa, Lisbon, Portugal*
[g] *Technical University of Denmark, Wind Energy Department, Roskilde, Denmark*
[h] *Bureau Veritas, Paris, France*
[i] *University of Exeter, Exeter, United Kingdom*
[j] *Shell Research Ltd, London, United Kingdom*
[j] *Department of Mathematics and Statistics, Lancaster University, United Kingdom*
[k] *Northeastern University, Boston, MA, USA*
[l] *University of Hannover, Hannover, Germany*
[m] *DNV-GL, Høvik, Norway*
[n] *University of Oslo, Oslo, Norway*

ARTICLE INFO

Keywords:
environmental contour
metocean extremes
joint distribution
extreme response
structural reliability

ABSTRACT

Environmental contours are used to simplify the process of design response analysis. A wide variety of contour methods exist; however, there have been a very limited number of comparisons of these methods to date. This paper is the output of an open benchmarking exercise, in which contributors developed contours based on their preferred methods and submitted them for a blind comparison study. The exercise had two components—one, focusing on the robustness of contour methods across different offshore sites and, the other, focusing on characterizing sampling uncertainty. Nine teams of researchers contributed to the benchmark. The analysis of the submitted contours highlighted significant differences between contours derived via different methods. For example, the highest wave height value along a contour varied by as much as a factor of two between some submissions while the number of metocean data points or observations that fell outside a contour deviated by an order of magnitude between the contributions (even for contours with a return period shorter than the duration of the record). These differences arose from both different joint distribution models and different contour construction methods, however, variability from joint distribution models appeared to be higher than variability from contour construction methods.

*Corresponding author

a.haselsteiner@uni-bremen.de (A.F. Haselsteiner); rcoe@sandia.gov (R.G. Coe); lmanuel@mail.utexas.edu (L. Manuel); chaiwei@whut.edu.cn (W. Chai); c.guedes.soares@centec.tecnico.ulisboa.pt (C. Guedes Soares); astah@dtu.dk (. Hannesdóttir); nkdi@dtu.dk (N. Dimitrov); guillaume.de-hauteclocque@bureauveritas.com (G. de Hauteclocque); e.mackay@exeter.ac.uk (E. Mackay); philip.jonathan@shell.com (P. Jonathan); atm@neu.edu (A. Myers); anna.rode@stud.uni-hannover.de (A. Rode); erik.vanem@dnvgl.com (E. Vanem)

ORCID(s): 0000-0002-9070-6933 (A.F. Haselsteiner); 0000-0003-0738-3772 (R.G. Coe); 0000-0002-0602-3014 (L. Manuel); 0000-0003-4847-4902 (W. Chai); 0000-0002-2084-0088 (B. Leira); 0000-0002-8570-4263 (C. Guedes Soares); 0000-0003-3399-4526 (. Hannesdóttir); 0000-0003-1325-4512 (N. Dimitrov); 0000-0001-8717-9688 (A. Sander); 0000-0002-0552-3404 (J. Ohlendorf); 0000-0002-5911-805X (K. Thoben); 0000-0001-7121-4231 (E. Mackay); 0000-0001-7651-9181 (P. Jonathan); 0000-0002-7566-8633 (C. Qiao); 0000-0002-5799-8022 (A. Hildebrandt); 0000-0002-9064-1615 (B. Schmidt); 0000-0002-0875-0389 (E. Vanem); 0000-0001-5990-0765 (A.B. Huseby)

1. Introduction

The environmental contour method is often used to aid in the design and analysis of marine structures. It is a simplified approach that derives extreme environmental conditions, which can be used to estimate the N-year structural response (with N corresponding to a target reliability for an ultimate limit state; for example $N = 50$ years). A more accurate approximation of the true N-yr structural response can be computed by integrating the product of the short-term response distribution, conditional on the environmental state, and the long-term joint distribution of the environment. This approach is usually referred to as "full long-term analysis" (see, for example, Guedes Soares (1993); Muliawan, Gao and Moan (2013)). Because a full long-term analysis can be computationally expensive, the environmental contour method is commonly used to get a first estimate of the response or as an entire replacement for performing a full long-term analysis.

It is important to understand that the typical two-dimensional contour method ignores response variability conditional on the sea state (or more generally the "environmental state") conditions and response computations are later made only for points on the contour that represent a subset of environmental conditions of interest for design. For each design condition on the environmental contour a number of stochastic time-domain simulations are run and the peak response in each simulation is recorded. The peak response in the environmental state is then defined as the median or – to indirectly account for short-term variability – a higher percentile (Baarholm, Haver and Økland, 2010; Muliawan et al., 2013) of the peak values over each simulation. Alternatively, short-term variability can also be accounted for by inflating the contour by using a higher return period or by applying an additional safety factor that is multiplied with the environmental load (NORSOK, 2007). Finally, the design value is taken as the largest response over all sea states along the contour. There are two limitations with this approach to design—first, many environmental states are not checked and, second, the short-term variability of the response is not fully assessed. Any full long-term response-based analysis (see, for example Vanem, Guo, Ross and Jonathan, 2020) makes neither of these approximations but can be prohibitively expensive.

Note that the environmental contour method is a special case of more general inverse reliability approaches that can expand the dimension space by one variable to include the response conditional on the environment as an additional variable. An advantage is that more environmental states can be checked than were included in the contour method but a disadvantage is that points on the hypersphere for the target reliability require quantiles of response conditional on the environment that are far from median levels and require considerable amount of simulations, albeit less than with the all sea states approach and, of course, a greater number of environmental states must be evaluated than with the environmental contour method. This has been demonstrated in the seminal work (Winterstein, Ude, Cornell, Bjerager and Haver, 1993) and explicit inclusion of response variability has been employed in the design of fixed and floating offshore wind turbines (Rendon and Manuel, 2014; Liu, Thomas, Goyal and Manuel, 2019). It is important to emphasize that contour methods and other inverse reliability methods are acknowledged as approximate methods; they were founded on principles of structural reliability with a view toward limiting computation that can be prohibitive with a full long-term analysis.

Deriving an environmental contour from a metocean dataset generally involves two steps: estimating the joint distribution of the environmental variables of interest, for example, wave height and period, and constructing the environmental contour based on that joint distribution (Figure 1). For both steps, various approaches have been proposed. The joint distribution can be estimated using different model structures such as global hierarchical models (for example, Mathisen and Bitner-Gregersen, 1990; Bitner-Gregersen, 2015; Horn, Bitner-Gregersen, Krokstad, Leira and Amdahl, 2018; Cheng, Svangstu, Moan and Gao, 2019), copula models (for example, Vanem, 2016; Fazeres-Ferradosa, Taveira-Pinto, Vanem, Reis and das Neves, 2018; Manuel, Nguyen, Canning, Coe, Eckert-Gallup and Martin, 2018; Zhang, Kim, Beer, Dai and Guedes Soares, 2018; Heredia-Zavoni and Montes-Iturrizaga, 2019; Lin, Dong and Tao, 2020), kernel density estimates (for example, Ferreira and Guedes Soares, 2002; Eckert-Gallup and Martin, 2016; Haselsteiner, Ohlendorf and Thoben, 2017a) or conditional extremes models (Jonathan, Flynn and Ewans, 2010; Jonathan, Ewans and Flynn, 2014). Another consideration relates to how the model parameters are estimated – even if the same model structure is used, parameter values estimated using, for example, maximum likelihood estimation or the method of moments, can strongly deviate from each other (Guedes Soares and Henriques, 1996; Vanem, 2015).

For the second step, namely construction of the contour, various methods have been proposed that differ in their definition of which regions in the variable space are considered exceedances (Figure 2). They include Haver's constant exceedance method (Haver, 1985, 1987), the inverse first-order reliability method (IFORM; Winterstein et al. (1993)), the inverse second-order reliability method (ISORM; Chai and Leira (2018)), inverse directional simulation

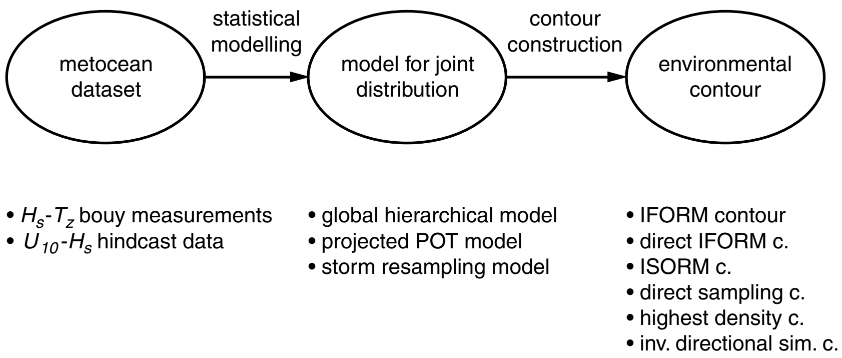


Figure 1: Deriving an environmental contour from a metocean dataset requires two consecutive steps: estimating the environment's long-term joint distribution ("statistical modelling") and constructing a contour based on that joint distribution ("contour construction"). In the lower part of the figure the types of metocean datasets, statistical models and environmental contours that have been used in this benchmark are shown.

(Dimitrov, 2020), the direct sampling contour method (Huseby, Vanem and Natvig, 2013), direct IFORM (Derbanne and de Hauteclocque, 2019), joint exceedance contour methods (Jonathan et al., 2014) and the highest density contour method (Haselsteiner, Ohlendorf, Wosniok and Thoben, 2017b). Due to contrasting definitions, even for the same underlying joint distribution, for any target exceedance probability, α, different constructed contours will result from each of the methods. Broadly, the contour construction methods can be classified by two criteria: i) whether the contour is constructed in the original (physical) variable space (as with the direct sampling, direct IFORM and highest density method) or in a standard normal space (as with IFORM, ISORM and inverse directional simulation) and ii) whether the contour definition is based on one or more regions in the variable space associated with the target (α) exceedance probability: In IFORM, direct IFORM and the direct sampling contour method, the failure surface is approximated as a hyperplane, which is equivalent to defining the contour exceedance probability as a marginal exceedance probability under a rotation of the axis. During contour construction, all axis rotations are considered such that there are many regions in the variable space that contain probability α, however, only one of these regions is assumed to match the failure surface of the structure of interest. In ISORM, inverse directional simulation and the highest density contour method, a single region that covers the complete variable space outside the contour, contains a probability of α. These two classes of contours can therefore be summarized as contours based on marginal exceedance probabililty and contours based on total exceedance probability. A study on the properties of these two classes of contours was presented by Mackay and Haselsteiner (2021).

The different definitions for the exceedance region are related to an important approximation of the environmental contour method: For a deterministic response, the method assumes that if a structure is designed using environmental conditions that have a joint exceedance probability of $\alpha = 1/(T_R \times n_{yr})$ (where T_R is the return period in years and n_{yr} is the number of environmental states per year), the resulting structure will have a probability of failure, p_f, that is less than, but close to α. Consequently, it is desired that the contour's exceedance region is a conservative approximation of the structure's failure region. Note, however, that the goal of conservatism in the approximation of the failure surface is not absolute: A slightly unconservative approximation can be compensated by choosing an appropriate safety factor, which is later multiplied with the design load. Studies on different constructed contours for the same joint distribution have been published by Leira (2008); Vanem and Bitner-Gregersen (2015); Vanem (2017); Huseby, Vanem and Eskeland (2017); Haselsteiner et al. (2017b); Chai and Leira (2018); Wang, Wang and Woo (2018); Vanem et al. (2020). A broad recent review on the environmental contour method was provided by Ross, Astrup, Bitner-Gregersen, Bunn, Feld, Gouldby, Huseby, Liu, Randell, Vanem and Jonathan (2019) and a recent effort presenting a comparison framework for environmental contours was published by Eckert, Martin, Coe, Seng, Stuart and Morrell (2021).

To provide a common basis to compare proposed environmental contour methods, a benchmarking exercise was proposed at the International Conference on Ocean, Offshore & Arctic Engineering (OMAE 2019; Haselsteiner, Coe, Manuel, Nguyen, Martin and Eckert-Gallup (2019)). This benchmark exercise involved two components—one focused on analyzing the robustness of contour methods across different sites ("Exercise 1") and the other focused on charac-

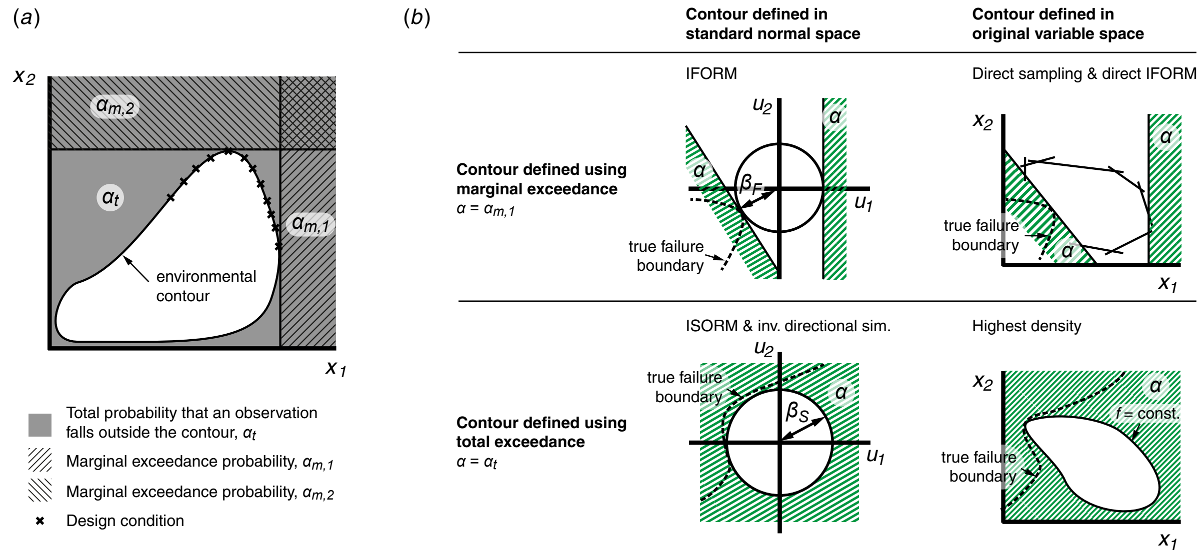


Figure 2: Contour construction methods differ in how the exceedance region associated with a probability α is defined. (a) An environmental contour, its associated marginal exceedance probabilities and its total exceedance probability. (b) Exceedance regions in the contour methods used in this benchmark. The shown failure surfaces are simple examples, illustrating that contours defined using marginal exceedance were proposed for convex failure regions and contours defined using total exceedance were proposed for non-convex failure regions. Hatched area = exceedance region, α=exceedance probability used to construct the contour, x_1 and x_2 = environmental variables in the original variable space, u_1 and u_2 = environmental variables transformed into standard normal space, β_F and β_S = reliability index of the inverse first- and second-order reliability method, f = probability density.

terizing sampling uncertainty ("Exercise 2"). Six datasets were provided, each involving two environmental variables. Three datasets comprised time series of significant wave height and zero-up-crossing period while three other datasets included wind speed and significant wave height data. Participants were asked to compute environmental contours with return periods of 1, 20 and 50 years. These two-dimensional contours of wave and wind variables were chosen for comparison because they represent common cases required in practice for design. The use of sea state contours is recommended in guidelines and standards such as DNV GL's recommended practice on environmental conditions and environmental loads (RP-C205; DNV GL, 2017) and NORSOK's standard on actions and action effects (N-003; NORSOK, 2007). The use of wind-wave contours is required when following IEC's standard on the design of offshore wind turbines (International Electrotechnical Commission, 2019). Studies on sea state environmental contours include the works of Haver (1985); Winterstein et al. (1993); Eckert-Gallup, Sallaberry, Dallman and Neary (2014); Velarde, Vanem, Kramhøft and Dalsgaard (2019) and studies on wind speed wave height contours have been published, for example, by Saranyasoontorn and Manuel (2006) and Karmakar, Bagbanci and Guedes Soares (2016). Some researchers combined these three variables to construct wave height, wave period and wind speed "contours", for example, Li, Gao and Moan (2015, 2016); Li, Yuan, Gao, Zhang and Tezdogan (2019); Vanem (2019).

Nine teams of researchers participated in the exercise and this paper presents the results. The benchmarking study was designed as an open, systematic comparison, allowing for both different models for the joint distribution and different methods for contour construction. The study's design did not aim to rank the appropriateness or accuracy/quality of any participant's submitted contours or methods used to derive them, but rather to provide a comparison of the wide range of contour methods that have been proposed in recent years.

The benchmark's results highlight and quantify some significant differences between contours derived via the different methods. For example, the highest wave height value along the contour varied by as much as a factor of two between some submissions while the number of metocean data points or observations that fell outside a contour deviated by an order of magnitude between the contributions (even for contours with a return period shorter than the duration of the record). While not covered by the originally scoped benchmark exercise proposal (Haselsteiner et al., 2019), the paper by de Hauteclocque, Mackay and Vanem provides further comparisons of the nine submitted contributions (de Hauteclocque, Mackay and Vanem, in preparation).

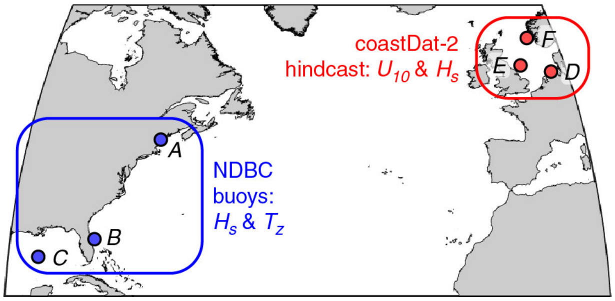


Figure 3: Locations of the six used datasets. Datasets A, B and C represent measurements from moored buoys. The data were downloaded from the website of the National Data Buoy Center (NDBC). Datasets D, E and F were obtained from the hindcast coastDat-2 (Groll and Weisse, 2017). Both types of datasets represent hourly time series. The provided buoy datasets cover 10 years and the provided hindcast datasets cover 25 years. Time periods of equal length were retained for evaluation.

The next section, section 2, describes the contour methods that were employed in the nine submitted contributions. It first provides a high-level overview and, then, a single sub-section for each contribution that offers additional details. Then, section 3 describes all the results using the methods and metrics that were stipulated in the paper that announced the exercise (Haselsteiner et al., 2019). Finally, section 4 contains a discussion and section 5 provides some conclusions.

2. Exercises and contributions

Nine teams submitted contributions for Exercise 1 and four of those teams submitted for Exercise 2. A high-level overview of the details of each contribution is presented in Table 2. In addition, Table 2 references code and stand-alone papers for each method, where available.

The two exercises were described in detail in the paper that announced the benchmark (Haselsteiner et al., 2019). They were based on six provided datasets (Figure 3). Three datasets, dataset A, B, and C, contained hourly buoy measurements from three sites off the US coast. They contained hourly time series of significant wave height, H_s, and zero-up-crossing period, T_z. The other three datasets, dataset D, E and F, were obtained from the hindcast coastDat-2 (Groll and Weisse, 2016, 2017) and covered three locations in the North Sea. They contained hourly time series of a 10-minute mean wind speed, U_{10}, at 10 m above sea level, and a 1-hour significant wave height, H_s. The organizers provided 10 years of data of sea state measurements (datasets A, B, and C) and 25 years of wind-wave data (datasets D, E and F) to the participants. For evaluation of the submitted contours, another 10 years of sea state data and 25 years of wind-wave data were retained by the organizers until after all the participants' results were received. To distinguish between the different parts of the datasets, we refer to the parts that were made available to the participants as "provided", to the parts that were not made available as "retained", and to the entire set as "full".

In Exercise 1, participants were required to compute 1-yr and a 20-yr sea state contours and 1-yr and 50-yr wind-wave contours. The exercise focused on the robustness of a contour method across different offshore sites. This robustness is mainly affected by the type of model used to represent the joint distributions. Three types of comparisons were outlined for Exercise 1 in the OMAE 2019 paper:

- plotting all submitted contours in an overlay for visual comparison,
- reporting the maximum values along the contour in each dimension and
- counting the number of points (measured environmental conditions) outside each contour.

To provide perspective to the number of points outside the contours, here, we also calculate the expected number of points outside a contour, assuming that measurement data represent independent observations (the impact of serial correlation in the data on this metric is discussed in section 4). This number varies with the contour's return period and the method that was used to construct the contour (Table 1). It is important not to overly weight the importance and reliability of this metric. Since environmental contours are generally used in engineering analyses in predicting extreme responses, only points that fall outside the contour that elicit large responses are generally of practical concern.

Table 1
Calculating the expected number of points outside a contour. Contour construction methods define the exceedance region that contains probability α differently such that the total exceedance probability outside the contour, α_t, varies (Figure 2). Assuming independent observations, the expected number of points outside a contour can be calculated as $E[n_{outside}] = n \times \alpha_t$; where α is the exceedance probability used to construct the contour, α_t is the probability that the contour is exceeded anywhere, n is the number of data points in the sample, χ^2_n is the chi-square distribution function, and Φ^{-1} is inverse normal distribution function.

Contour construction method	**α_t (total exceedance prob.)**	**$E[n_{outside}]$ for a 1-yr contour and a 50-yr hourly sample**
IFORM (Winterstein et al., 1993)	$1 - \chi^2_n([\Phi^{-1}(1-\alpha)]^2)$*	492
ISORM (Chai and Leira, 2018)	α	50
Inverse directional simulation (Dimitrov, 2020)	α	50
Direct sampling (Huseby et al., 2013)	ca. similar to IFORM*	ca. 492
Direct IFORM (Derbanne and de Hauteclocque, 2019)	ca. similar to IFORM*	ca. 492
Highest density (Haselsteiner et al., 2017b)	α	50

* See Mackay and Haselsteiner (2021) for additional background on this equation.

For this reason, in addition to reporting the number of points outside each contour alongside the analytically expected result, we also present the number of points outside each contour above a certain threshold.

Exercise 2 focused on characterizing sampling uncertainty. Participants were asked to sample 1-, 5- and 25-yr subsets from the provided dataset D and to compute environmental contours based on these subsets. For each of the three time periods, 1000 such subsets were required such that 1000 environmental contours per time period were calculated. To compare the contributions, it was described that an uncertainty overlay of these 1000 contours should be plotted and that confidence intervals should be calculated.

The following subsections describe the individual contributions. To keep overall paper length in balance, the subsections are relatively brief, however, for some contour methods further details are provided in individual papers.

2.1. Contour method 1

by Wei Chai and Bernt Leira

The environmental contours are constructed by application of the inverse second-order reliability method (ISORM). For calculation of the failure probability, the FORM approximation will underestimate the result for cases with a concave failure surface in the standard normal space (i.e. the U-space), see Fig. 4(b) for a two-dimensional example. Correspondingly, the IFORM contour will yield non-conservative results for design purposes.

In order to address such a shortcoming of the traditional IFORM contour, Chai and Leira (2018) proposed a specific second-order approximation to the failure surface in the U-space. Generally, the SORM approximation may provide better approximations to the failure probability than the FORM approach. The failure surface of the specific SORM approximation is assumed to be a circle in U-space for the two-dimensional case. Therefore, the estimated failure probability and the corresponding contour are generally conservative.

Similar to the development of n-dimensional contour for a given return period by the IFORM method, in the process of establishing the corresponding ISORM contours, an n-dimensional sphere with the radius β_S is first created, with the value of β_S being determined by the following equation:

$$1 - P_f = \int_{\sum_{i=1}^{n} u_i^2 \leq \beta_S^2} \phi_U(\mathbf{u}) d\mathbf{u} \tag{1}$$

where $\mathbf{u}$ represents an n-dimensional vector in the normalized U-space and $\phi_U(\mathbf{u})$ denotes the standard multivariate normal probability density function (PDF).

It is seen from Eq. (1) that, in the normalized U-space, the probability content outside the sphere with radius β_S is P_f. Moreover, the sum of n independent standard normal variables, follows a Chi-squared distribution with n degrees of freedom. Therefore, the radius β_S can be determined from the following equation:

$$\chi_n^2(\beta_S^2) = 1 - P_f \tag{2}$$

Table 2
Exercises contributions. Some participants provided open-source code to reproduce their results or wrote stand-alone papers describing their contributions. DIFORM: Direct IFORM with de-clustering, DSCM: direct sampling contour method, GHM: global hierarchical model, HDCM: highest density contour method, IDSCM: inverse directional simulation contour method, IFORM: inverse first-order reliability method, ISORM: inverse second-order reliability method, PPOTM: projected peak over threshold model, SRM: Storm resampling with non-stationary model for storm peaks.

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

[3] https://github.com/ec-benchmark-organizers/ec-benchmark/tree/master/participants-code/contribution-3
[4] https://github.com/ahaselsteiner/2020-paper-omae-hierarchical-models
[5] Haselsteiner et al. (2020)
[6] Mackay and Jonathan (2020)
[7] Vanem and Huseby (2020)

Subsequently, the n-dimensional sphere with radius β_S in the normalized U-space is transformed into the ISORM contour in the original parameter space by application of the Rosenblatt transformation if the joint distribution of environmental parameters is described by the conditional modelling approach. The Nataf transformation is applied if the marginal PDFs of the environmental parameters (in combination with corresponding correlation coefficients) are applied in order to describe the joint distribution of environmental parameters. In present exercise, we used the baseline joint distribution models that were provided in the paper that proposed this benchmark (Haselsteiner et al., 2019). They were established by following the conditional modelling approach.

2.2. Contour method 2

by Guilherme Clarindo and C. Guedes Soares

Participant 2 used the models for the joint distributions that were provided as baseline results by the benchmark organizers. However, while the baseline results then constructed IFORM contours, here, direct sampling contours were constructed.

The joint distribution of metocean variables was obtained by applying the conditional modelling approach (see Guedes Soares, Lopes and Costa (1988); Bitner-Gregersen and Haver (1991)). The model structure consists of a marginal PDF for H_s and a conditional PDF for co-variables T_z and U_{10}. The marginal distribution of H_s is assumed

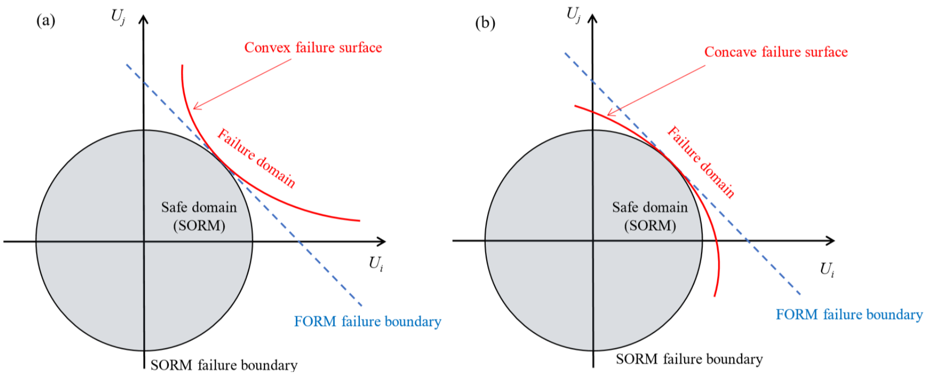


Figure 4: Illustration of failure probability approximated by the FORM and proposed specific SORM in the normalized U-space.

Table 3
Contribution 2 statistical scheme for datasets.

Datasets	***A-C***		***D-F***	
Structure	Marginal	Conditional	Marginal	Conditoinal
Variables	H_s (m)	T_z (s)	H_s (m)	U_{10} (m/s)
Statistical model	Weibull 3p	Lognormal	Weibull 3p	Weibull 2p

to follow a three-parameter Weibull distribution in all dataset, while a conditional lognormal distribution is assumed for T_z in A, B and C, as, for example, in Lucas and Guedes Soares (2015). For the remaining dataset D, E and F, the U_{10} distribution is now conditioned by two parameters of a Weibull distribution and the statistical dependence is based on scale and shape parameters, whereas the lognormal distribution was conditioned by the mean and variance respectively.

The applied approach for contour construction is based on direct methods presented by Huseby et al. (2013); Huseby, Vanem and Natvig (2015) as an approach to establish environmental contours directly in the original space of the environmental variables based on Monte Carlo simulations of the fitted joint metocean observations, thus not requiring any transformation. The initial inaccuracies due to insufficient number of Monte Carlo samples can be improved by a reject sampling scheme explained by Huseby, Vanem and Natvig (2014). The understanding of this approach was extended to three dimensions by Vanem (2019).

Contours were constructed using 10° as the angular intervals, thus generating 35 points of intersection, which can be defined as coordinates. The number of samples generated by Monte Carlo simulation were five hundred thousand for each dataset.

2.3. Contour method 3

by Ásta Hannesdóttir and Nikolay Dimitrov

In this contribution the inverse directional simulation (IDS) was used to construct the environmental contours (Dimitrov, 2020). As shown in section 2.1 and in Figure 4, the classical IFORM approach is formulated to compute the probability of failure behind a linear limit state surface. This formulation is suitable for classical reliability analysis, where one assumes that the failure region is convex and the goal is to find a single design point (the most likely point of failure). Then it is sufficient to evaluate the probability behind the limit state surface in the close vicinity of the design point. However, for other types of problems, multiple points with equal return periods forming an entire contour or a segment (a part of contour) could be considered equally critical. In such situation it is required that the exceedance probability accounts for all events outside any part of the contour or segment. Using IFORM for such problems would lead to underestimation of the failure probability and a non-conservative result. The IDS method provides an exact solution for computing the total probability outside the return period contour or outside a contour segment of arbitrary size. This is achieved by replacing the linear IFORM failure boundary by a hyper-sphere in standard normal space. As with the IFORM, the reliability index (β) defines the radius of the sphere (or circle in 2D) which equals the L_2 norm of the variable vector $\mathbf{u}$:

$$\beta = \sqrt{\sum_i u_i}, \qquad i = 1, ..., n \tag{3}$$

Here u_i are the environmental variables in standard normal space and n is the number of dimensions in the variable space. Because the variables in U-space are independent and normally distributed, their sum follows by definition a chi-squared distribution. In a n-dimensional variable space, the radius (reliability index) may be defined by

$$\beta = \sqrt{\chi_n^{-1}(1 - P_f)} \tag{4}$$

where χ_n^{-1} is the inverse cumulative distribution function (CDF) of the chi-square distribution with n degrees of freedom and P_f is the probability of failure.

With the above formulation, it is straightforward to determine the radius of a contour segment where only exceedances within a certain range of variable combinations are relevant (for example, for a contour of wind speed vs. wave heights, we may be interested only in the segment where high wind speeds are combined with large wave heights). This procedure is described in Dimitrov (2020). For the case when the full contour length is considered, the IDS method is equivalent to the ISORM method (Chai and Leira (2018)).

Apart from how the reliability index is defined, the implementation of the IDS in the present study follows the general steps of the IFORM procedure as performed in for example Hannesdóttir, Kelly and Dimitrov (2019) for 10-minute mean- and standard deviation of wind speed measurements. The reliability index is estimated with equation 4 for exceedance probability corresponding to the entire space outside the contour. For datasets A, B and C we model the joint distribution assuming that the marginal distributions of both the significant wave height and the conditional distribution of zero up-crossing periods follow a log-normal distribution. For datasets D, E and F we model the joint distribution where we assume the marginal distribution of significant wave height to follow a 3-parameter Weibull distribution and the conditional distribution of wind speed to follow a 2-parameter Weibull distribution. The marginal distributions are fitted with the maximum likelihood method and the parameters of the conditional distribution are estimated with a least squares fit.

2.4. Contour method 4

by Andreas Haselsteiner, Aljoscha Sander, Jan-Hendrik Ohlendorf and Klaus-Dieter Thoben

We fitted novel types of global hierarchical models to the datasets and constructed highest density contours. Our model for the joint distribution of sea states assumes that the marginal distribution of significant wave height follows an exponentiated Weibull distribution and that zero-up-crossing period follows a conditional log-normal distribution. The model for the wind-wave joint distribution assumes that wind speed follows an exponentiated Weibull distribution and that significant wave height follows a conditional exponentiated Weibull distribution.

Recently, we analyzed how well the exponentiated Weibull distribution fits significant wave height data and we proposed a method that prioritizes high wave heights over low waves when the distribution's parameters are estimated (Haselsteiner and Thoben, 2020). This weighted least squares estimation method was used here as well. The dependence structure between the variables was designed to yield simple relationships between U_{10}, H_s and T_z that can be interpreted physically,

The models assume that the median significant wave height follows the relationship

$$\tilde{h}_s = c_1 + c_2 u_{10}^{c_3} \tag{5}$$

and that the median zero-up-crossing period follows the relationship

$$\tilde{t}_z = c_4 + c_5 \sqrt{h_s / g}, \tag{6}$$

where c_i are parameters that are estimated and g is Earth's gravity constant. Benefits of these dependence functions are that the parameter c_3 enables simple interpretation of how wave heights increase with wind speeds and that the dependence of $\tilde{t}_z$ on h_s is expressed in physically consistent units (unfortunately, the exponent c_3, which is estimated as a float, does not allow for physically consistent units in the relationship between $\tilde{h}_s$ and u_{10}). Additionally, the relationship between $\tilde{t}_z$ and h_s requires only two estimated parameters.

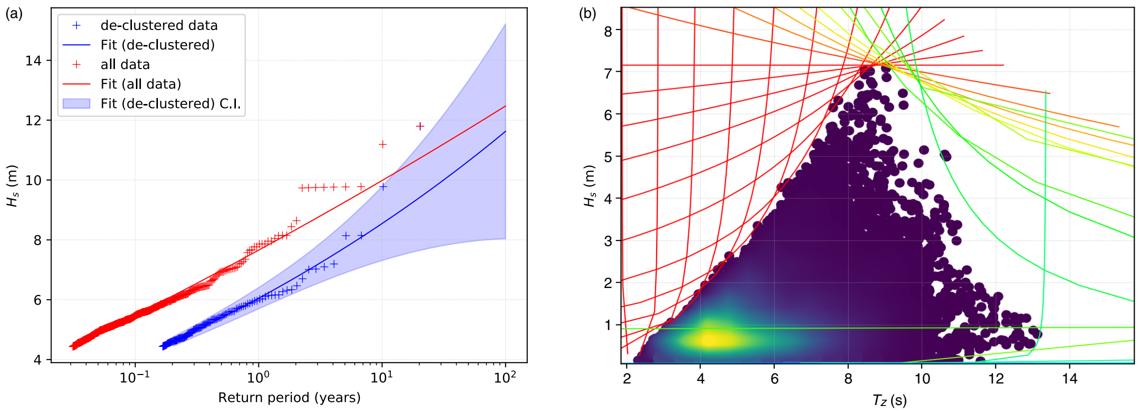


Figure 5: D-IFORM (contribution 5) illustration, Dataset A. (a) Effect of de-clustering. (b) Contour construction from tangents (transformed back in physical plane).

The complete biviarate models for $H_s - T_z$ and $U_{10} - H_s$ required the estimation of 8 and 10 parameters, respectively. These fitted joint distributions were used to construct highest density environmental contours. A highest density contour contains $1-\alpha$ probability content within the environmental contour and has constant probability density along its path (Figure 2). Thus, for a given joint distribution it leads to more conservative design conditions than IFORM or direct sampling contours (see Mackay and Haselsteiner (2021) for an analysis of the differences and application examples). Further details on this contribution are given in a conference paper (Haselsteiner et al., 2020).

2.5. Contour method 5

by Guillaume de Hauteclocque

The environmental contours are calculated using the method presented in Derbanne and de Hauteclocque (2019). Compared to all other approaches used in the benchmark, this method does not rely on the inference of a joint distribution. Extending the direct sampling method, the two original variables (H_s and T_z for instance) are projected on a search direction α (eq. 7) and univariate fit is performed on each $X(\alpha)$.

$$X(\alpha) = \bar{v}_1 \cos\alpha + \bar{v}_2 \sin\alpha \tag{7}$$

For each direction α, distribution parameters $\theta(\alpha)$ for $X(\alpha)$ are estimated (by MLE or other means). Those parameters are used to interpolate/extrapolate $X_{ext}(\alpha)$ at desired probability. The contour is then constructed from the tangents (or hyper-plane for $n > 3$) defined by α and $X_{ext}(\alpha)$ as with the tangent method described in Huseby et al. (2013) and illustrated on Figure 5b. In practice, the construction of a contour from tangents works better if data are convex. Thus a variable change is often necessary. Additionally, the data are also scaled so that the two variables have comparable dimension. $\bar{v}_1$ and $\bar{v}_2$ in equation 7 thus correspond to the transformed and scaled variables. In the current benchmark, $[H_s, T_z]$ is, for instance, transformed to $[H_s$, Steepness$]$.

One benefit of this direct IFORM approach is its ability to straightforwardly plug any state-of-the-art method with respect to the univariate fit. Here, this possibility is used to get rid of the "independent and identically distributed events" (IID) assumption, by using standard peak over threshold (POT) approach together with de-clustering. On each search direction α data are de-clustered, using threshold up and down crossing as cluster boundaries, with an additional minimum interval of 48 hours between clusters. Threshold exceedances are then fitted with generalized Pareto distribution (see Coles (2001) for further details). Figure 5a illustrates the de-clustering, effect on H_s (i.e. $\alpha = 0$.) for dataset A. The effect is dramatic for $T_R = 1$ yr, while, at $T_R = 20$ yrs, conclusion are not obvious due to large sampling uncertainties. Contour construction from tangents is presented on Figure 5b.

Note: Confidence intervals in Figure 5 are calculated using the delta method (Coles, 2001), and are not displayed when using all data because of the dependence of sea states (the effective number of independent observations is reduced, which strongly biases the uncertainty calculation by delta method).

2.6. Contour method 6

by Ed Mackay and Philip Jonathan

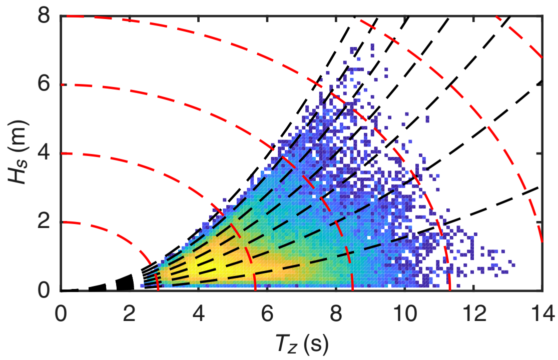


Figure 6: Joint occurrence of H_s and T_z for dataset A. Black dashed lines: constant significant steepness. Red dashed lines: constant distance $d = \sqrt{H_s^2 + T_z^2/2}$.

The joint distribution of environmental parameters was estimated using the storm resampling method presented in Mackay and Jonathan (2020) and contours were derived from the joint distribution using the standard IFORM method. In the block resampling method, it is assumed that the time series of environmental variables can be divided into blocks where the peak values in adjacent blocks can be considered independent. The peak values of each variable within the block are not required to coincide in time, but the blocks are assumed to be sufficiently short so that the peak values of each variable are related in some way. A model for the joint distribution of the peak values is then estimated. The distribution of all data is recovered by simulating block-peak values from the joint model and resampling and rescaling the measured blocks so that the peak values from the resampled blocks match the simulated peak values. As the data that the model is fitted to are approximately independent, this gives a better justification for the use of asymptotic extreme value models. The approach also has the advantage that much of the complex short-term dependence structure in the data is resampled rather than modelled explicitly.

The block resampling approach only preserves the distribution of the peak values in the block. For contours of H_s and T_z we are interested in both the maximum and minimum values of T_z for a given H_s. To get around this problem we work with the significant steepness, $s = 2\pi H_s / g T_z^2$, and a distance variable defined as $d = (H_s^2 + T_z^2/2)^{1/2}$. Lines of constant d are orthogonal to the lines of constant s in the H_s-T_z plane (see Figure 6). Moreover, the peak values of s and d correspond to the frontiers of interest for H_s-T_z contours. The marginal distributions of the block-peak values of s and d were modelled using a composite approach, with a kernel density model for the body and a generalized Pareto (GP) model for the tail. The joint distribution was also modelled using a composite approach, with a kernel density model for the body and the Heffernan and Tawn (2004) (HT) model for the tail.

A similar approach was used for estimating the marginal distributions of block-peak H_s and wind speed, U_{10}. However, for the joint distribution a piecewise-linear GP model was used for the distribution of H_s^{peak} conditional on U_{10}^{peak} for intermediate values of U_{10}^{peak} and the HT model was used for estimating the joint distribution for higher values of U_{10}^{peak}. Further details are provided in Mackay and Jonathan (2020).

2.7. Contour method 7

by Chi Qiao and Andrew Myers

When calculating environmental contours with a return period longer than the measurement period, the joint probability distribution must be extrapolated beyond the available data. Such an extrapolation brings two challenges for the global hierarchical models (expressed in a general form as $f_X \cdot f_{Y|X}$), where f_X indicates the distribution of the independent variable X and $f_{Y|X}$ indicates the conditional distribution of the dependent variable Y). First, tail extrapolation is required for both f_X and $f_{Y|X}$, but extreme value theory cannot be applied directly, since the IID assumption is not satisfied. Second, $f_{Y|X}$ needs to be parameterized as $f_{Y|X}(\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ represents the parameters (e.g., shape, scale, and location) of the distribution, and $\boldsymbol{\theta}$ needs to be fitted as a function of the independent variable, $\boldsymbol{\theta}(X)$, so that $f_{Y|X}$ is estimated at the tail of X where little or no data is available. It is challenging in practice to find such a prescribed distribution (for example, a Weibull distribution) for the dependent variable that is both accurate within its tail and shows a good fit of $\boldsymbol{\theta}(X)$.

Dataset	***A-C***		***D-F***	
Structure	Marginal	Conditional	Marginal	Conditional
Parameter	H_s	T_z	H_s	U_{10}
Distribution	2-parameter Weibull	2-parameter Weibull	2-parameter Weibull	2-parameter Weibull

Table 4
Models for the joint distribution that were used in Contribution 8.

A new framework is proposed to improve the performance of global hierarchical models in constructing the joint probability distribution. The framework is designed to be as flexible as possible to fit many types of environmental variables. A dedicated manuscript is submitted to this special issue to provide more details. In brief, two novel approaches are used to resolve the aforementioned challenges respectively. The first approach provides a method to extrapolate the tail of dependent measurements (for example, hourly sampled wind speed, which does not satisfy IID). This method makes use of the results from the extreme value theory and accounts for the dependency exhibited in the measurements. The resulting probability is a hybrid expression, with six parameters describing the tail (three for the generalized extreme value distribution, and three describing the dependency) and a prescribed distribution (or even empirical estimations) for the non-tail part. In the second approach, the conditional distribution with the hybrid expression is re-parameterized using a prescribed distribution, which avoids the need to fit a large number of parameters in the hybrid expression. Only the quantiles related to the desired return period are used for the re-parameterization, which allows almost any type of prescribed distribution to construct environmental contours regardless of the goodness-of-fit for irrelevant quantiles. As such, the optimal prescribed distribution for the dependent variable is selected based on the goodness-of-fit of $\boldsymbol{\theta}(X)$. The joint probability distribution is then used to construct the environmental contours presented in this paper via the IFORM approach.

2.8. Contour method 8

by Anna Rode, Arndt Hildebrandt and Boso Schmidt

The parameters of all statistical models were estimated with maximum likelihood estimation (MLE) and subsequently the contours were derived with the classical IFORM approach (Winterstein et al., 1993).

For the three datasets A-C the significant wave height H_s was opted as the marginal distribution, whereas the zero-up-crossing period T_z was modelled as the conditional environmental variable (Table 4). Both the marginal and the conditional distribution were fitted by a 2-parameter Weibull distribution. The stochastic parameters of the marginals were estimated as above mentioned with MLE. The dependence functions of the conditional distributions for T_z were estimated by using a linear regression of the previously classified Weibull parameters.

Datasets D-F provide data of the significant wave height H_s and the wind speed U_{10}. As with the datasets A-C, the environmental variable H_s is assumed to be the marginal distributed variable for these three hindcast datasets. In these cases, the wind speed is the conditional environmental variable. The environmental variable H_s and U_{10} were both assumed to follow a 2-parameter Weibull distribution. For constructing the environmental contours, the classical inverse first-order method (IFORM) as presented in Winterstein et al. (1993) was used.

2.9. Contour method 9

by Erik Vanem and Arne Bang Huseby

A set of environmental contours were calculated in this contribution, as outlined in Vanem and Huseby (2020). All contours were based on fitting a conditional model to the data, as a product of a marginal model for the primary variable and a conditional model where the model parameters are modelled as parametric functions of the primary variable (Bitner-Gregersen, 2015; Horn et al., 2018). For all datasets the marginal model for H_s were the 3-parameter Weibull distribution. For the sea state data a conditional log-normal distribution was assumed for T_z and for the wind-wave data, a conditional (2-parameter) Weibull distribution was assumed for U_{10}. Based on the fitted joint models, environmental contours were calculated by two different approaches, i.e. the IFORM approach (Haver and Winterstein, 2009) and the direct sampling approach (Huseby et al., 2013, 2015). The main differences between these contour methods are that the IFORM approach includes a transformation to standard normal space and assumes a convex failure region in the transformed space (or rather, performs a linearization of the failure boundary in this space), whereas the direct sampling approach makes the same assumption in the original parameter space, see for example (Vanem and Bitner-Gregersen, 2015; Vanem, 2017).

The direct sampling contours were calculated based on a set of Monte Carlo simulations from the joint distributions, and there may be numerical uncertainties due to the Monte Carlo variance. This can be reduced by increasing the number of samples, and in this work, an efficient tail-sampling approach has been utilized in order to obtain a large effective number of samples with reasonable computational efforts, see for example Huseby et al. (2014); Vanem and Huseby (2018).

3. Results

This section presents the benchmark's results, using the previously described analysis methods. First the results of the sea state contours are reported (datasets A, B, C; subsection 3.1), then the wind-wave contours are presented (datasets D, E, F; subsection 3.2). Finally, the results for Exercise 2, the uncertainty characterization that was applied to dataset D, are reported (subsection 3.3). Appendix A lists the various joint models that were fitted by the participants.

3.1. Sea state contours

A set of overlays of the sea state contours are shown in Figure 7. In each of the six plots shown in Figure 7, the eleven contour contributions are overlayed with the full dataset (there are contributions from nine teams, but Contribution 9 contains three different contours that were calculated using the same joint distributions). The maximum H_s value as well as the maximum and minimum T_z value along each 1-yr and 20-yr contour is plotted in Figure 8. For reference, the empirical marginal 1-yr return value as well as the maximum and minimum of the full measured datasets are also plotted.

The maximum H_s values along the 1-yr and 20-yr contours strongly deviate between the contributions. For the 1-yr contours, as expected, among the contributions the highest H_s value is from a contour that is based on total exceedance (Contribution 4 in dataset A and Contribution 3 in datasets B and C). For the 20-yr contours, the highest H_s value is from a total exceedance contour in dataset A (Contribution 4) and from a marginal exceedance contour in datasets B and C (Contribution 7). In all datasets the highest maximum H_s value along the 20-yr contour is more than double the lowest maximum H_s value. For example, in dataset A the lowest value is about 5 m (Contribution 2) and the highest value is about 13 m (Contribution 4).

In dataset B and C the highest buoy-measured H_s value is within the range of the 20-yr contours's maximum values, but in dataset A the highest measured value is higher than all contour maxima. This effect is likely due to the much higher H_s maximum in the retained part of dataset A compared to the provided part (ca. 12 m and ca. 7 m, respectively). In general, differences due to the different joint models appear to be greater than differences due to different contour construction methods, as there is no consistent order between the maxima of marginal exceedance contours and total exceedance contours.

The counted data points outside the contours are presented in Table 5. As described in section 2, the expected number of data points outside the contours is different among the contributions, as different contour construction methods were used (see Table 1). However, for contributions that constructed ISORM, inverse directional simulation and highest density contours ("total exceedance contours"), the expected number of points outside the contour, $E[n_{outside}]$, is the same. For contributions that constructed IFORM, direct sampling and direct IFORM contours ("marginal exceedance contours") $E[n_{outside}]$ is similar. Note that if a statistical model was fitted to declustered data such that serial correlation is reduced, the environmental contour will have smaller dimensions and the number of expected data points outside the contour, $E[n_{outside}]$, will be higher for such a contour. Contribution 5 constructed direct IFORM contours after applying declustering.

For the first group, the total exceedance contours, the theoretical $E[n_{outside}]$ is 1 for a 20-yr contour, but between 13.7 and 114.0 points exceeded the constructed contours (average over datasets A, B, C). For the second group, the marginal exceedance contours, $E[n_{outside}]$ is exactly (IFORM) or approximately (direct sampling, direct IFORM) 11.5, but between 16.7 and 21966.3 points exceeded the constructed contours. Contribution 5 and 9, which have the highest points outside the 20-yr contour, contain many exceeding points at low H_s values, which are irrelevant for structural design. If these points are excluded by only counting sea states with a significant wave height greater than 1 m, between 7.7 and 280 points exceed the constructed marginal exceedance contours. In summary, in both contour classes the number of exceeding points varied by an order of magnitude among the submitted contributions.

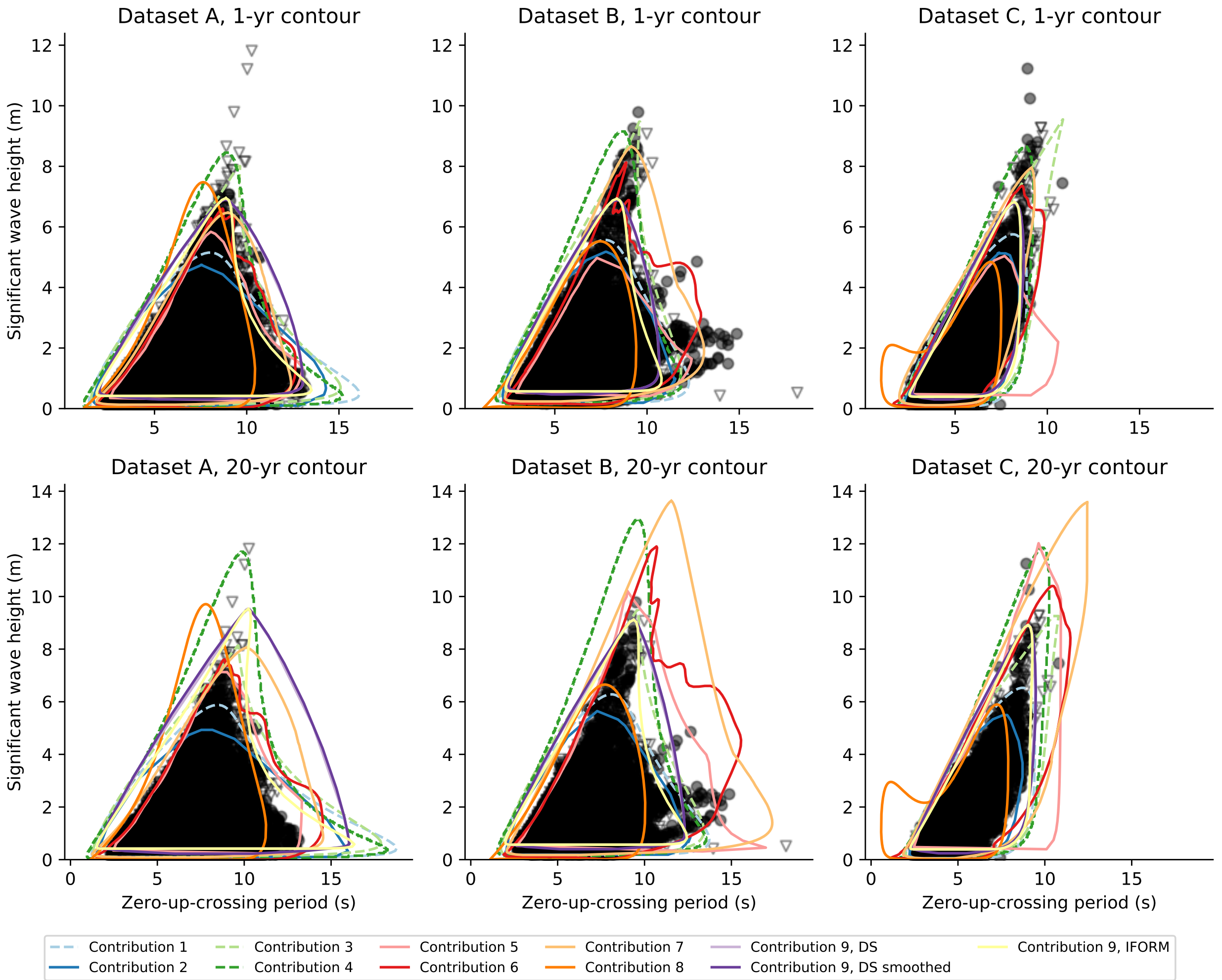


Figure 7: Exercise 1, datasets A, B, C: contour overlays. — = marginal exceedance contours, -- = total exceedance contours, ∘ = provided data (10 years), ▽ = retained data (10 years).

Table 5
Number of points outside in datasets A, B, C (full datasets). Reported is the average over the three datasets. For the 20-yr contour, values in the paranthesis are the points outside for dataset A, B and C individually.

Contr.	**1-yr contour**			**20-yr contour**		
	num. points outside	**expected points outside**	**points outside where $h_s > 1$ m**	**num. points outside**	**expected points outside**	**points outside where $h_s > 1$ m**
1	256.3	20	241	114.0 (153, 127, 62)	1	98.7 (117, 119, 60)
2	406.7	ca. 197	389.3	286.7 (437, 236, 187)	ca. 11.5	271.3 (401, 228, 185)
3	82.3	20	48.3	43.0 (68, 28, 33)	1	23.3 (14, 26, 30)
4	75	20	63	13.7 (1, 32, 8)	1	11.7 (1, 30, 4)
5	18295.3	ca. 197[1]	495.3	16764.0 (8503, 12573, 29216)	ca. 11.5[1]	24.7 (47, 24, 3)
6	154	197	115.3	16.7 (27, 14, 9)	11.5	13.3 (20, 12, 8)
7	281.3	197	63	35.3 (79, 4, 23)	11.5	7.7 (21, 0, 2)
8	762	197	638	322.0 (189, 368, 409)	11.5	280.0 (93, 339, 408)
9 DS	22127.3	ca. 197	130.7	21966.3 (22504, 22933, 20462)	ca. 11.5	21.7 (4, 43, 18)
9 DS s.	12031.7	ca. 197	126.7	6062.7 (5096, 7716, 5376)	ca. 11.5	21.0 (4, 42, 17)
9 IFORM	22207.7	197	233.7	21994.7 (22514, 22990, 20480)	11.5	47.7 (8, 100, 35)

1 Due to the applied declustering the expected number data points outside the contour will be higher than this number.

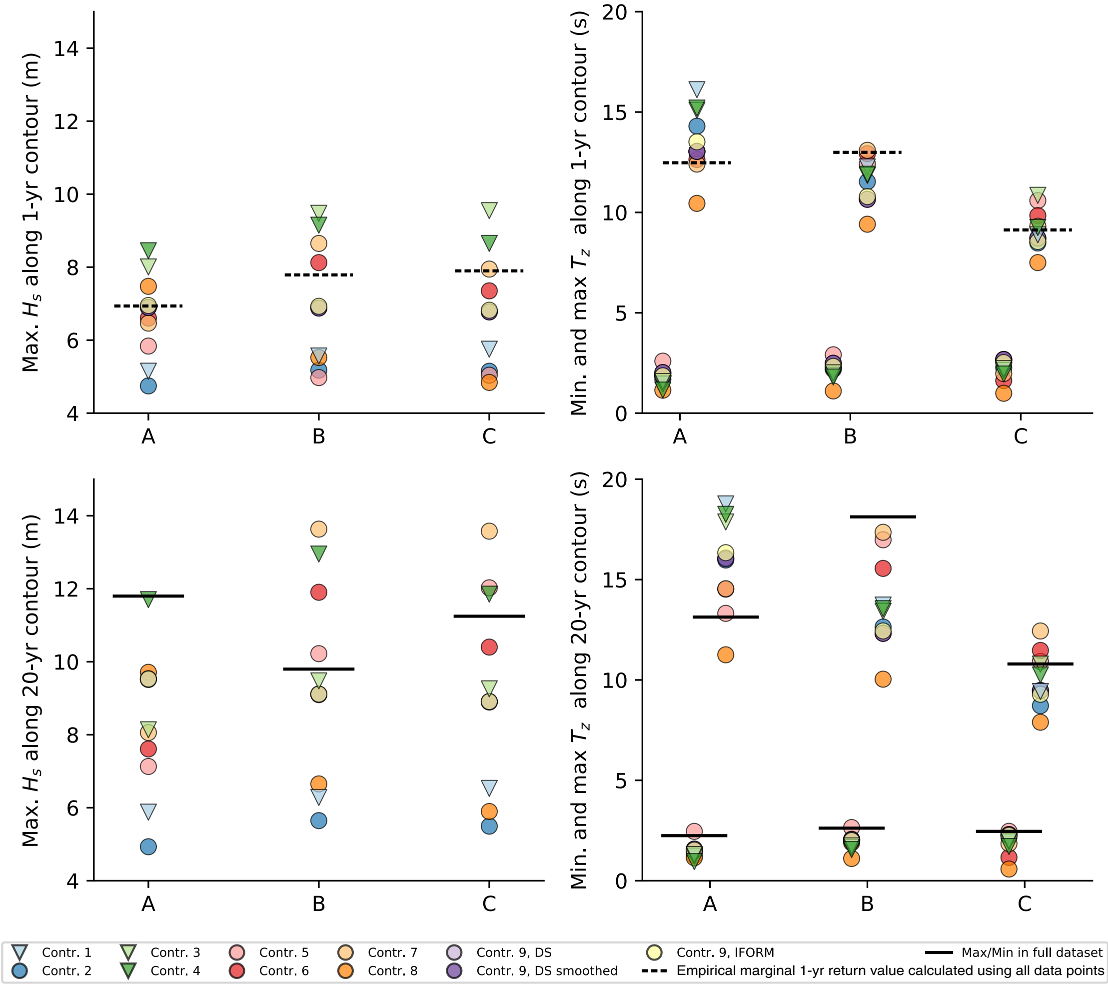


Figure 8: Exercise 1, maximum values along the contour for datasets A, B, C. For zero-up-crossing period, T_z, the minima are shown too (there is no horizontal line for the 1-yr minimum in the top right panel because the dashed lines in this figure represent the marginal 1-yr return value). ∘ = marginal exceedance contours (IFORM, direct sampling, direct IFORM), ▿ = total exceedance contours (ISORM, inverse directional simulation, highest density).

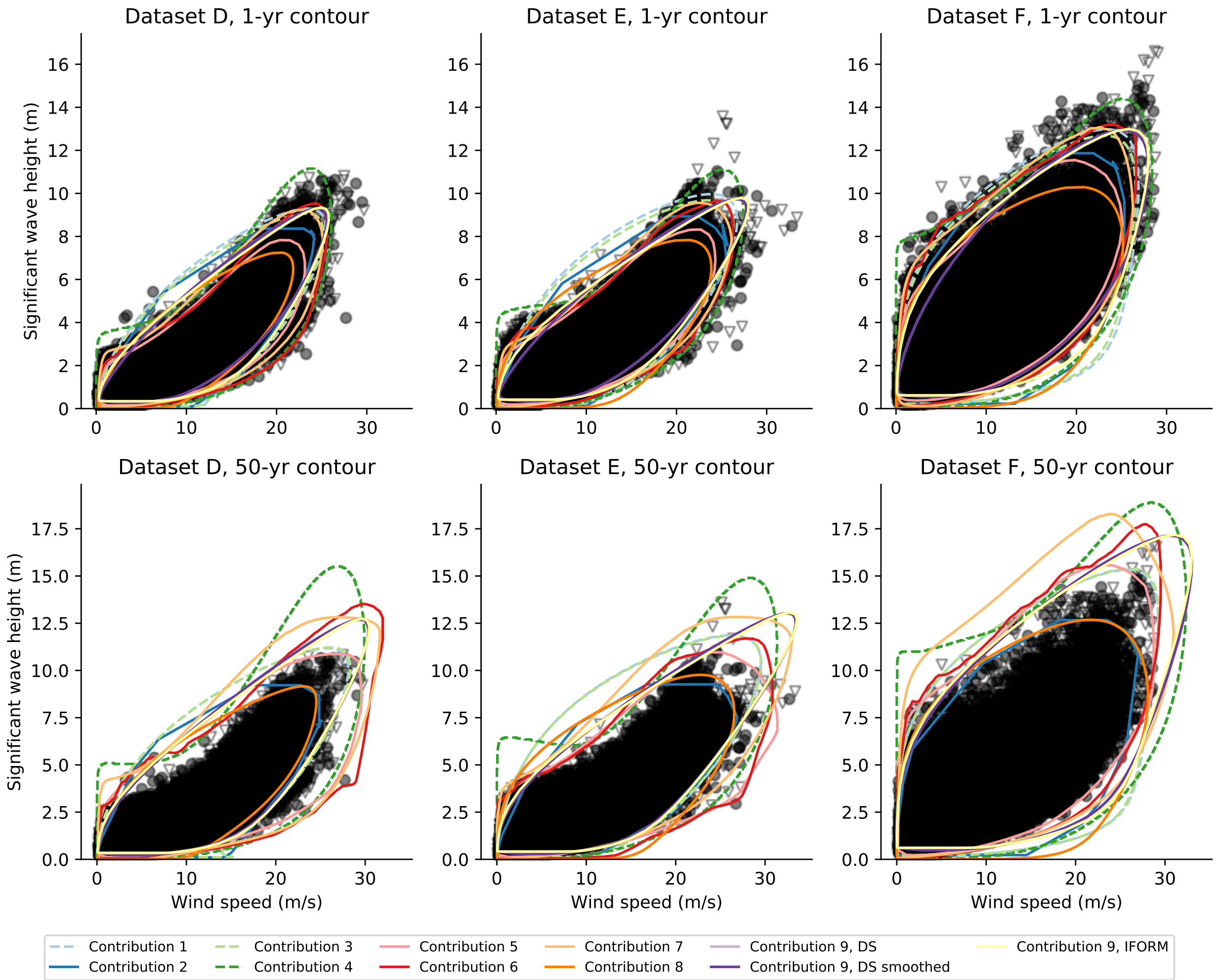


Figure 9: Exercise 1, datasets D, E, F: contour overlays. — = marginal exceedance contours, -- = total exceedance contours, ∘ = provided data (25 years), ▿ = retained data (25 years).

3.2. Wind-wave contours

Overlays of all contours are plotted in Figure 9. The maximum values along the contours are plotted in Figure 10. Similar to the sea state contours, there is a wide variability among the contributions. The spread from the highest to the lowest maximum value is higher for the wave height values than for the wind speed values (ca. 40-50% versus ca. 20% of the empirical maximum for the 50-yr contours). For example, in dataset D the contributions have maximum wave height values between ca. 9 m and 15 m and maximum wind speed values between ca. 25 m/s and 32 m/s. The highest wave height maximum along the 50-yr contour is on a total exceedance contour in all datasets (Contribution 4), but the highest wind speed maximum is on marginal exceedance contours (Contributions 6 and 9). Consequently, similar to the sea state contours, differences due to selected model types for the joint distribution seem to be bigger than differences due to the selected contour construction methods.

The number of data points outside the wind-wave contours is presented in Table 6. For reference, the expected number of points outside the contour, if events were independent, is also reported. For total exceedance contours, the expected number of points outside the contour is $E[n_{outside}] = 1$ for a 50-yr contour and 50 years of data, but between 3.7 (Contribution 4) and 103.7 (Contribution 1) points exceeded the constructed contours (average over datasets D, E, F). For marginal exceedance contours, $E[n_{outside}]$ is exactly (IFORM) or approximatley 12 (direct sampling, direct IFORM), but between 8.3 (Contribution 6) and 15301 (Contribution 9 DS) points exceeded the constructed contours. Similar as with the sea state contours, the thousands of exceeding points in Contribution 9 are at low wave heights

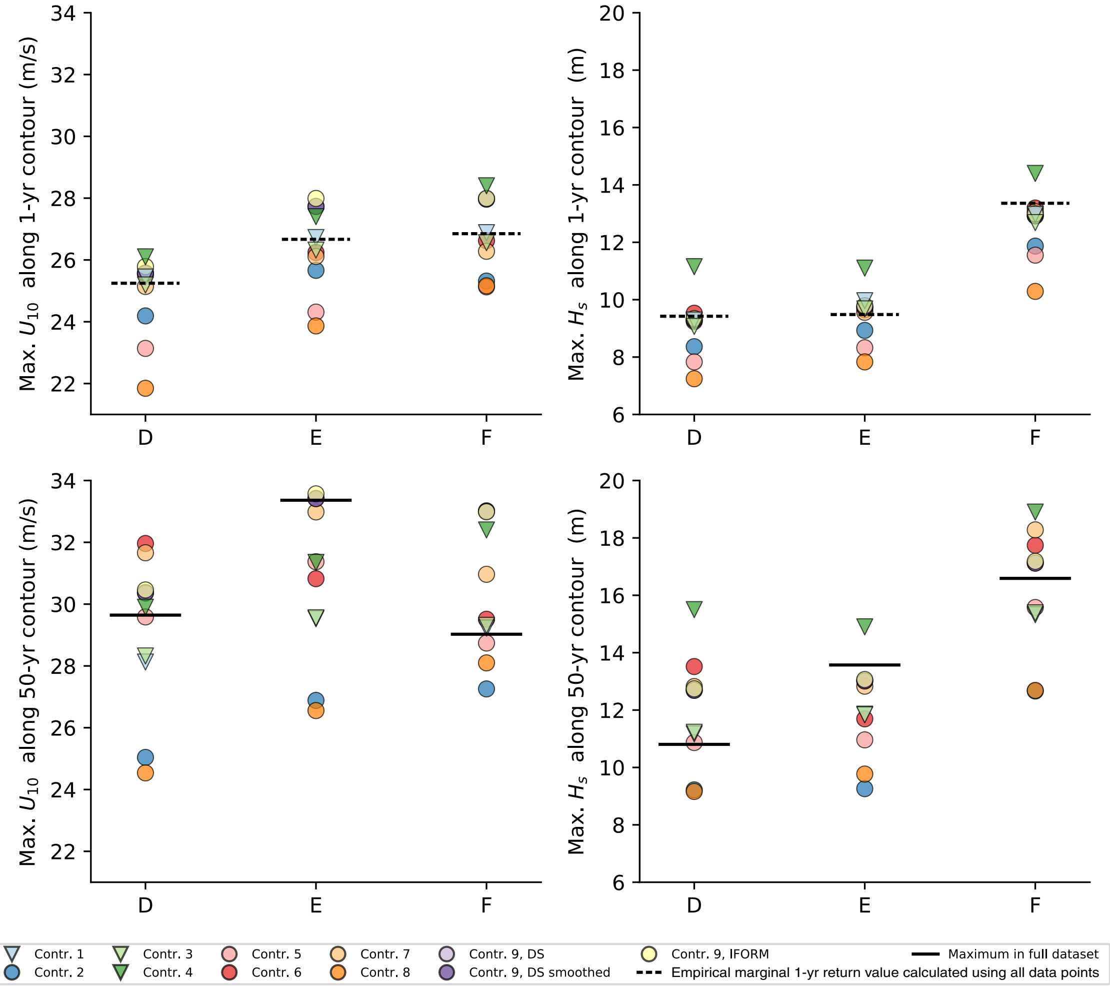


Figure 10: Exercise 1, maximum values along the contour for datasets D, E, F. $\circ$ = marginal exceedance contours (IFORM, direct sampling, direct IFORM), $\triangledown$ = total exceedance contours (ISORM, inverse directional simulation, highest density). The empirical 1-yr return value is calculated based on the empirical distribution of the full datasets, assuming that these events are independent and identically distributed.

and are therefore not relevant for structural design. If these points are excluded by only counting sea states with a significant wave height greater than 1 m, and wind speeds greater than 1 m/s only 103±45 points exceed the contour (mean ± standard deviation; n=3). In summary, similar to the sea state contours, the number of exceeding points vary by an order of magnitude among the submitted contributions (for both groups, marginal exceedance and total exceedance contours).

3.3. Uncertainty of the wind-wave contours

To better understand the level of sampling uncertainty in different contours, an exercise was proposed using a bootstrap sampling approach. The procedure was based on a study, first published by Gramstad, Vanem and Bitner-Gregersen (2018) and extended by Vanem, Gramstad and Bitner-Gregersen (2019), and was defined as follows (Haselsteiner et al., 2019):

1. Set the index, $i = 1$.
2. Resample Y (1, 5 or 25) years of data from dataset D (resulting in sample O_i).
3. Fit the model structure that you used in Exercise 1 to the sample O_i (resulting in the statistical model $\mathbf{X}_i$).

Table 6
Number of points outside in datasets D, E, F (provided and retained). Reported is the average over the three datasets. For the 50-yr contour, values in the paranthesis are the points outside for dataset D, E and F individually.

Contr.	1-yr contour			50-yr contour		
	num. points outside	expected points outside	points outside where $u_{10} > 1$ m/s and $h_s > 1$ m	num. points outside	expected points outside	points outside where $u_{10} > 1$ m/s and $h_s > 1$ m
1	424.3	50	268.3	103.7 (154, 86, 71)	1	60.3 (107, 63, 11)
2	1186.3	ca. 492	744	529.7 (514, 720, 355)	ca. 12	309.3 (304, 388, 236)
3	524.7	50	341.3	102.7 (154, 82, 72)	1	60.0 (106, 62, 12)
4	88	50	88	3.7 (3, 8, 0)	1	3.7 (3, 8, 0)
5	2235	ca. 492	860	156.0 (174, 112, 182)	ca. 12	14.7 (7, 18, 19)
6	270.7	492	247.7	8.3 (3, 17, 5)	12	8.3 (3, 17, 5)
7	371.3	492	258	8.7 (3, 16, 7)	12	6.7 (2, 16, 2)
8	1729.7	492	1442	262.0 (509, 110, 167)	12	238.7 (449, 108, 159)
9 DS	18191.7	ca. 492	1941.3	15301.0 (17460, 17226, 11217)	ca. 12	103.0 (146, 122, 41)
9 DS s.	12819.3	ca. 492	1952.3	5771.3 (5092, 6782, 5440)	ca. 12	103.0 (146, 122, 41)
9 IFORM	16267.3	492	686.7	15267.0 (17415, 17184, 11202)	12	86.3 (123, 107, 29)

4. Compute a 50-yr contour with the same method that you used in Exercise 1 based on the statistical model $\mathbf{X}_i$ (resulting in environmental contour C_i).
5. If $i < 1000$: Increase the index i and repeat steps 2-4.

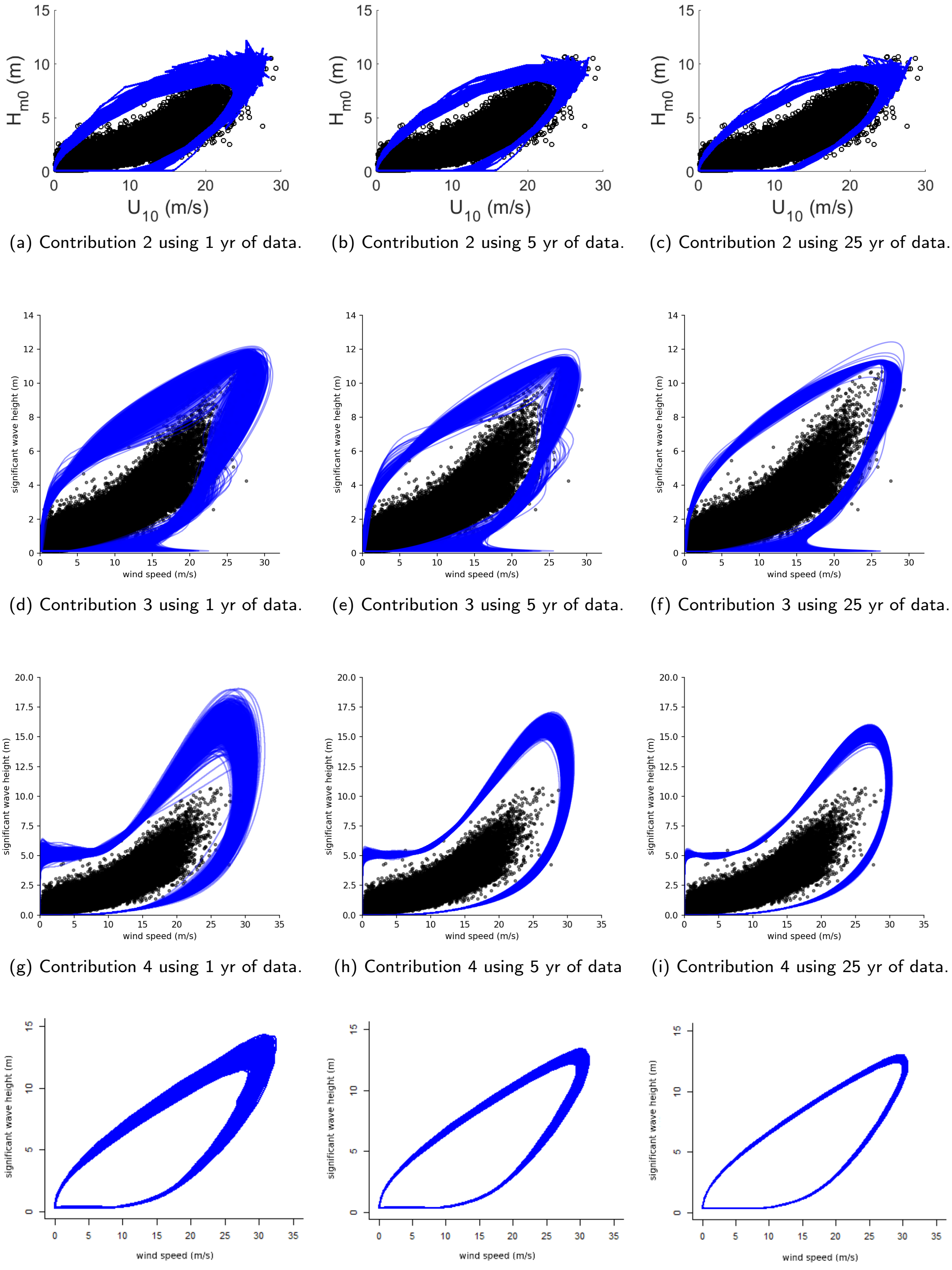
This procedure leads to 1000 different environmental contours. These contours were then used to compute a 95% confidence interval and a median contour. The methods for finding the confidence interval and median contour can be found in Haselsteiner et al. (2019). Note that the confidence intervals from this procedure are not a true reflection of the uncertainty due to a dataset's typical length because they neglect the serial correlation in the data. Thus, the results from this exercise should only be interpreted as comparative between contour methods rather than quantitative estimates of the sampling uncertainty. The effect of serial correlation will be further discussed in section 4. Further, contours that are based on total exceedance (Contributions 3 and 4) are expected to have greater uncertainty than contours that are based on marginal exceedance (Contributions 2 and 9), since for a given return period, total exceedance contours are extrapolating further in the tail of the distribution.

Overlay plots of all 1000 contours are presented in Figure 11. Each row in the figure relates to a specific contribution, from top to bottom these are Contributions 2, 3, 4, and 9 (some contributors chose not to participate in Exercise 2). The columns relate to the amount of data used in producing the contours. In the left-most column, for example, the 50-yr contours are produced using just 1 year of data. Moving to the right within the figure, more data is used in estimation of the contours. The overlay plots show that uncertainty decreases with increasing sample size in all contributions.

In Contribution 3, when 1 and 5 years of data are used, two different "modes" seem to be apparent: a major mode where the contours' upper parts have only strong curvature at the region of ca. ($u_{10} = 26$ m/s, $h_s = 11$ m) and a minor mode where the contours additionally have strong curvature at ca. ($u_{10} = 26$ m/s, $h_s = 6$ m). Possibly these different "modes" could also be present in the estimated parameter values. In the study by Vanem et al. (2019) on sampling uncertainty, the estimated parameter values clustered around two distinct values in a case where they resampled from a hindcast dataset and used maximum likelihood estimation.

Confidence intervals for all contributions are plotted in Figure 12. Similar to the overlay plots, they show how uncertainty decreases with increasing sample size. The confidence intervals for Contributions 2 and 9[1] were narrower than for Contributions 3 and 4. Contributions 2 and 9 constructed contours that belong to the group of marginal exceedance contours while Contributions 3 and 4 constructed total exceedance contours such that the latter group's contour required further extrapolation. This additional extrapolation should lead to greater variability among the 1000 contours. There are, however, many other differences between these four contribution such that it is not clear whether the different contour construction methods are the main reason for the different degrees of variability.

[1] In Contribution 9 the confidence intervals were computed with a different procedure than the one outlined in the paper that announced the benchmark (Haselsteiner et al., 2019) such that the results are not fully comparable. This procedure is described in the work by Vanem and Huseby (2020) and is considered to be only negligibly different from the procedure outlined in Haselsteiner et al. (2019).



(a) Contribution 2 using 1 yr of data.
(b) Contribution 2 using 5 yr of data.
(c) Contribution 2 using 25 yr of data.

(d) Contribution 3 using 1 yr of data.
(e) Contribution 3 using 5 yr of data.
(f) Contribution 3 using 25 yr of data.

(g) Contribution 4 using 1 yr of data.
(h) Contribution 4 using 5 yr of data
(i) Contribution 4 using 25 yr of data.

(j) Contribution 9 using 1 yr of data.
(k) Contribution 9 using 5 yr of data
(l) Contribution 9 using 25 yr of data.

Figure 11: Exercise 2, overlay plots of 1000 environmental contours based on resampling 1000 times from dataset D. As the sample's length increases uncertainty decreases.

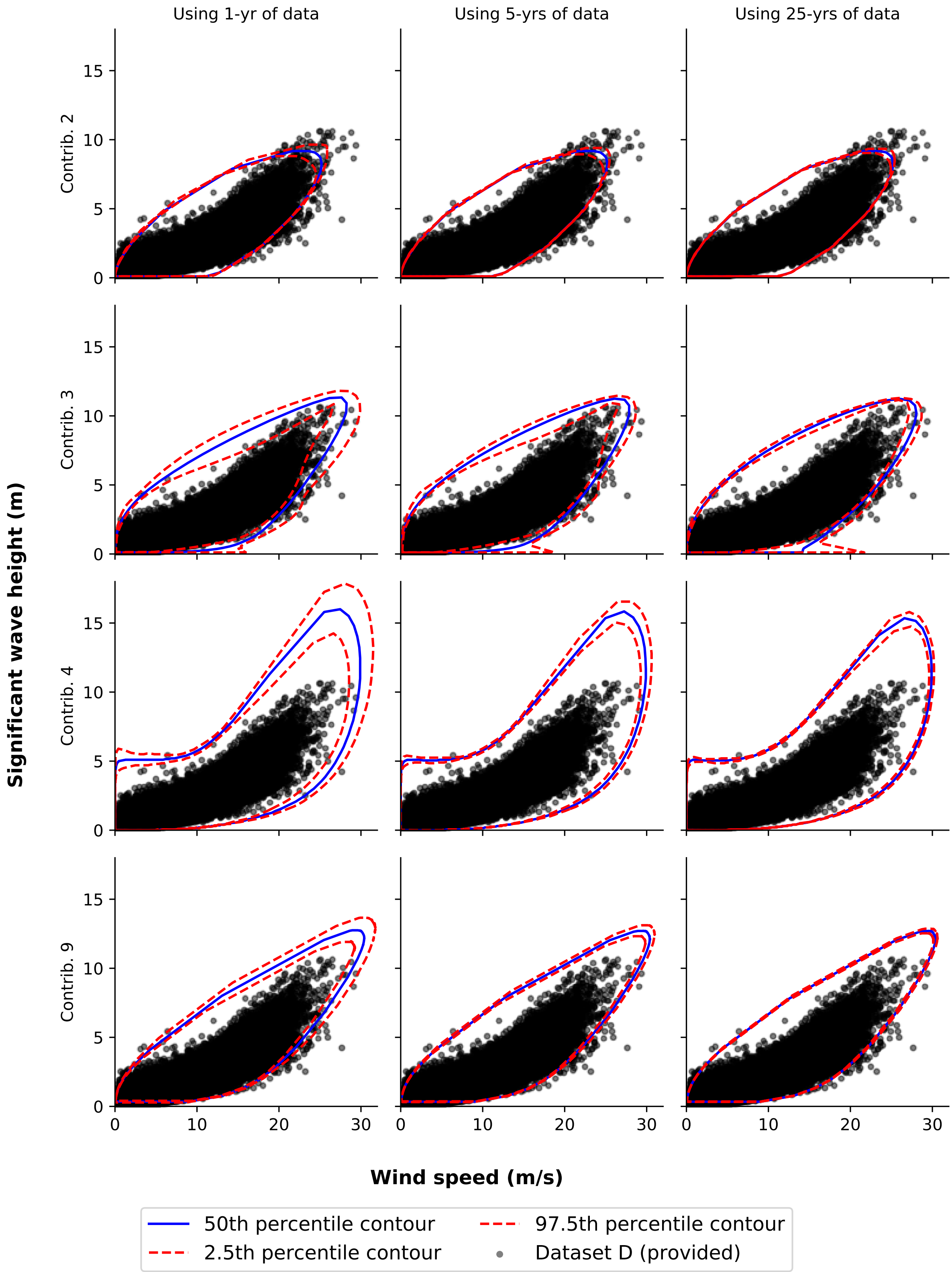


Figure 12: Exercise 2: Sampling uncertainty of the 50-yr contours. Each row of plots relates to a specific contribution (2, 3, 4, and 9); each column relates to the amount of data used to produce the 50-yr contours (1, 5, and 25 years).

4. Discussion

A common theme that affected our interpretation of the benchmark's results is the effect of the metocean data's autocorrelation or, as a synonym as we are dealing with time series, the effect of serial correlation. Thus subsection 4.1 will focus on how serial correlation affects return value estimates and measures of uncertainty. Subsection 4.2 will compare the different types of joint distribution models used in this study, focusing on extremal dependence. Then subsection 4.3 and subsection 4.4 will discuss the results of Exercise 1 and Exercise 2. Finally, we will suggest areas of future research (subsection 4.5).

4.1. Variability and serial correlation in the metocean datasets

This benchmark provided datasets comprising of hourly observations of metocean variables. Consecutive data points in such time series are not independent and identically distributed, but are strongly auto-correlated. There are multiple scales of variability and correlation in metocean conditions that correspond to different physical effects. These can be categorised as:

- short-term serial correlation of the order of hours to days related to passing weather systems,
- seasonal variability,
- inter-annual variability, related to longer-term climatic modes (for example NAO, ENSO, etc.) and
- longer-term, decadal-scale, climatic changes resulting from both anthropogenic influences and from naturally occurring climatic variations.

The full time series of significant wave height are shown in Figure 13. The seasonal variability is clearly observable because at these sites the highest H_s values typically occur in the winter months. Additionally, the time series show some rare storm events with much higher H_s than the highest storms in typical years. These rare storm events provide a challenge for estimating the tail of the probability distribution, because a metocean dataset might hold only one, two, or zero such events. Among the used datasets, the difference between the provided and retained part of the dataset was especially stark for dataset A. The maximum observed H_s in the provided portion was 7.1 m, whereas in the retained portion there were four storms where the peak H_s exceeded 8 m, with the largest storm peak H_s being 11.8 m.

The difference between the provided and the retained datasets can also be visualized by overlaying the empirical exceedance probability of observed storm peak H_s in both parts of the datasets (Figure 14). For this analysis, storm peaks are defined as a local maxima within a moving window of size five days. For datasets $\boldsymbol{B}$ and $\boldsymbol{C}$, the empirical distributions are similar for the provided and retained portions, indicating that a model that is a good fit to the provided data will also be a good fit for the retained data. However, for dataset A, the empirical distribution of the retained data has a significantly longer tail. In this case, a close fit to the tail of the provided data will underestimate the slope of the tail in the retained data.

The highest serial-correlation is in the short-term, in the order of hours to days related to passing weather systems (Figure 15). Because environmental contours were constructed based on bivariate distributions of sea states and wind-wave states, it is implicitly assumed that these distributions represent independent events. However, most contributions used the original, strongly auto-correlated time series to estimate these distributions, representing a strong modelling simplification. Some contributions used methods to avoid the serial correlation between observations by using a declustering method (Contribution 5) or by identifying independent storms (Contribution 6). While these approaches remove the most short-term serial correlation, some longer-term serial-correlation likely remains in the post-processed time series.

4.2. Types of joint models and extremal dependence

In this benchmark, participants submitted contributions based on three different kind of joint models: global hierarchical models, a projected peak over threshold model and a model based on storm resampling. These types of models make different assumptions on aspects like

- which kind of events the joint distribution represents (all environmental states or only storm peaks),
- whether and how the model accounts for serial correlation and
- how dependence is modelled (both, in the body and in the tail of the distribution).

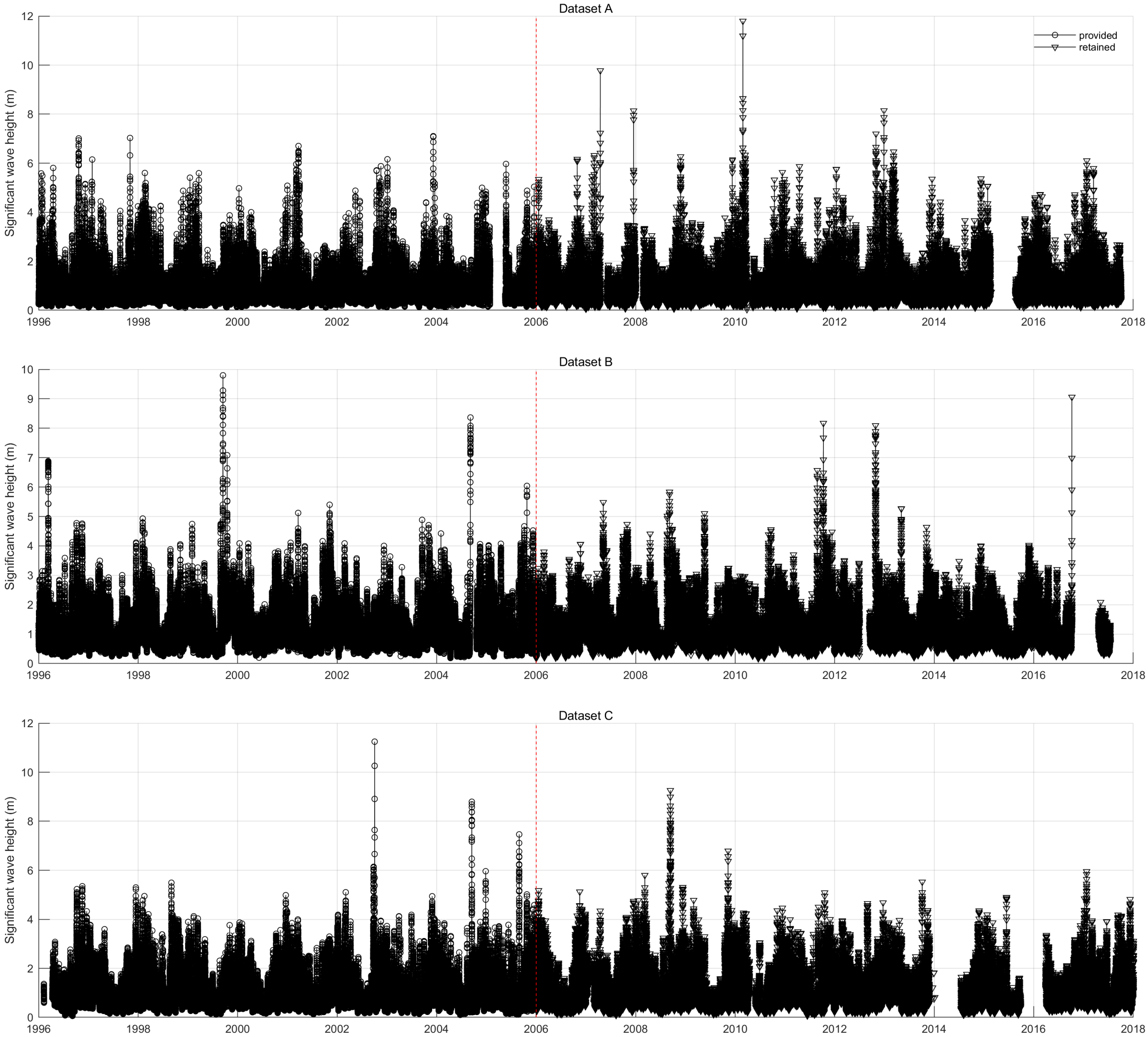


Figure 13: Full time series of significant wave height. The typical seasonal variability is visible, representing auto-correlation with a delay of one year. All datasets contain some storms where H_s strongly exceeds the typical annual maximum. The vertical dashed line denotes the seperation between the retained and provided data.

.

Because the environmental contour method deals with very high (and for some variables very low) quantiles, the model assumptions for the distribution's tail are especially important. Each of the models discussed in this work makes assumptions regarding the nature of dependence between variables in their joint tail. For variables X and Y, this extremal dependence can be quantified in terms of the upper tail dependence (see, Coles, Heffernan and Tawn (1999), who use the notation χ instead of λ for upper tail dependence),

$$\lambda = \lim_{u \to 1} \Pr(Y > F_Y^{-1}(u) | X > F_X^{-1}(u)). \tag{8}$$

λ quantifies how quickly Y becomes extreme when X becomes extreme. Often the value of λ for a particular model can be determined *a priori*. For example, a bivariate Gaussian distribution has $\lambda = 0$ regardless of the correlation between parameters (provided that the Gaussian correlation is <1). For some bivariate copula models, λ can be related

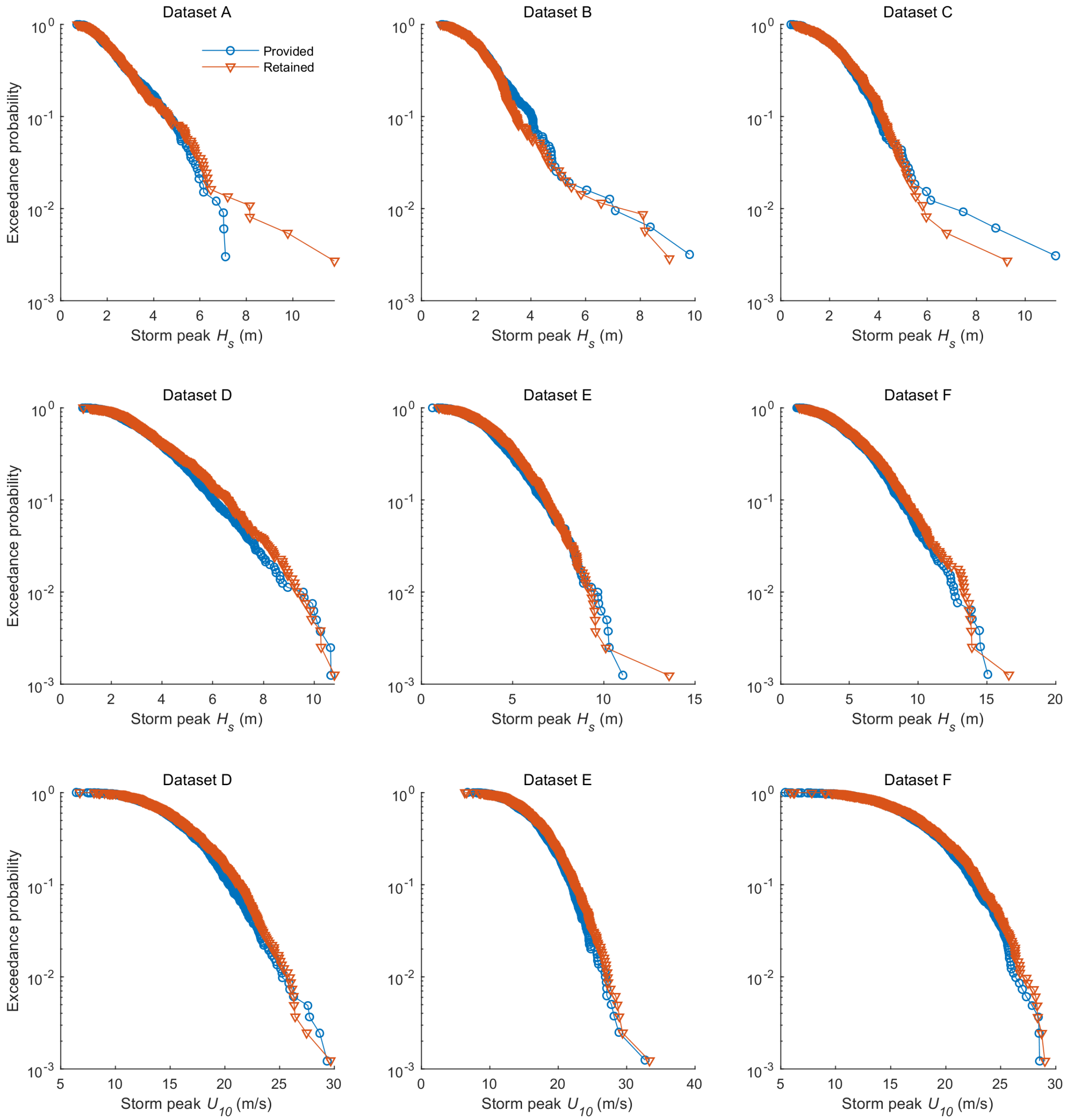


Figure 14: Empirical exceedance probability of observed storm peak H_s and U_{10}. Storm peaks were defined as a local maxima within a moving window of size five days.

to distributional summary statistics such as Kendall's tau, and to other model parameters. The conditional extremes model of Heffernan and Tawn (2004), which was used in Contribution 6, seeks explicitly to estimate whether the data exhibit asymptotic dependence (λ>0) or asymptotic independence (λ=0).

When deciding how to model the joint tail of a bivariate distribution, and hence attempting to estimate environmental contours corresponding to long return periods, the choice of bivariate distributional form and its extremal characteristics is of fundamental concern, and will influence the shape of the resulting contour estimate. Table 7 presents estimates $\hat{\lambda}$ for λ values of dataset D and for some joint distribution models fitted to it for probabilities of $u = 1 - \alpha$

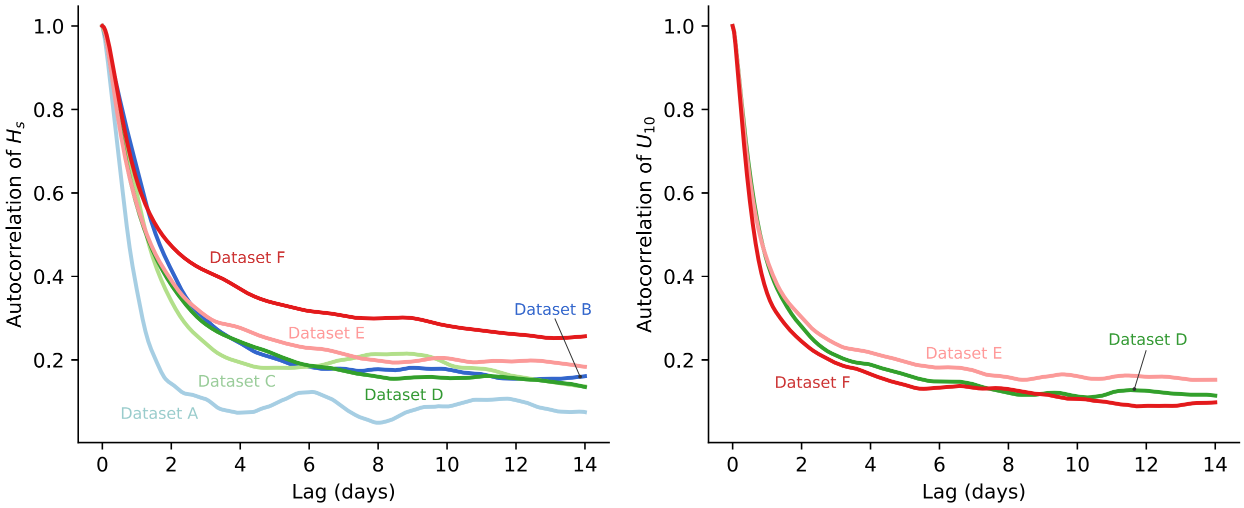


Figure 15: Autocorrelation function of the significant wave height, H_s, and wind speed, U_{10}, time series.

Joint distribution	$\hat{\lambda}(\alpha) = \Pr(H_s > \hat{F}^{-1}_{Hs}(1-\alpha) \vert U_{10} > \hat{F}^{-1}_{U_{10}}(1-\alpha)$		
	$\alpha = 10^{-2}$	$\alpha = 10^{-3}$	$\alpha = 10^{-4}$
Empirical distribution D_p	0.55	0.57	0.81
Empirical distribution D_r	0.85	0.89	0.57
Contribution 1 & 2 (baseline)	0.48	0.38	0.33
Contribution 4	0.53	0.40	0.28

Table 7
Upper tail dependence in dataset D and in some of the joint distribution models that were fitted to it. D_p = provided part of dataset D, D_r = retained part of dataset D.

with $\alpha = \{10^{-2}, 10^{-3}, 10^{-4}\}$. The provided part of dataset D has upper tail dependence of $\hat{\lambda} = 0.55$ at $\alpha = 10^{-2}$ and $\hat{\lambda} = 0.81$ at $\alpha = 10^{-4}$. The joint distribution models used by Contributions 1, 2 and 4 have lower values of upper tail dependence, ranging from $\hat{\lambda} = 0.53$ to $\hat{\lambda} = 0.28$ at $\alpha = 10^{-2}$ and $\alpha = 10^{-4}$, respectively. Note that, whereas upper tail dependence λ provides a means to quantify the extent of asymptotic dependence between variables, other statistics have been developed to summarise the extent of asymptotic independence ($\lambda = 0$, see Coles et al. (1999) for further details).

4.3. Discussion: exercise 1

The degree of variability amongst the presented contours is noticeable for all of the comparison cases. The maximum H_s values along the 20-yr sea state contours deviate circa by a factor of two (highest $H_s^{max} \approx 2\times$ lowest H_s^{max}). If, for example, we consider the contour for dataset A (see lower left-hand corner of Figure 7), the maximum significant wave height ranges from ca. 5 m to 12 m. This range in maxima can also be seen in Figure 8, which shows maxima of each contour. Consider that these different 20-yr contours are used to estimate the 20-yr extreme response of a marine structure. If the maximum H_s values deviate by a factor of two, responses that are linear functions of H_s will also deviate by factor of two. For example, if the motions of a vessel are analyzed in a sea keeping analysis, where often a linear relationship between the H_s and the vessel's movement is assumed, then design conditions from some contours would result in, for example, roll or pitch motions that are twice as high as the motions calculated using another contour. Some responses increases even more than linearly with H_s (see, for example, Vanem, 2017) such that responses could deviate even more than by a factor of two. Clearly, such differences can make a design either unreliable or strongly over-conservative.

The number of points outside of each contour presented in Tables 5 and 6 for sea state cases (datasets A, B, and C) and the wind-wave cases (datasets D, E, and F), respectively, provide a useful point of reference. While the number of points expected to fall outside of a contour can be estimated analytically when independent events are assumed, such a measure has multiple issues: (i) Hourly observations of metocean variables are not independent and identically distributed, but are strongly auto-correlated (Figure 15) and (ii) for the environmental contour method, it is only important whether a joint distribution model describes the true distribution at severe environmental conditions such that points outside the contour at severe conditions are of practical importance, while points at non-severe conditions are not.

The first issue, serial correlation, affects the metric in the sense that a single storm event might lead to one, two or n hourly observations that are outside a computed contour. Consequently, the observed number of points outside the contour is not the number of independent storm events, but represents a metric that could be interpreted as the number of independent storm events weighted by their duration.

The second issue, points outside the contour at non-severe regions, is well illustrated by looking, for example, at the results for Contribution 9 in Table 5. We can see that for the 20-yr contour there are 11.5 points expected outside of the contour, but the three contours from Contribution 9 have between 6×10^3 and 2.2×10^4 points outside of the 20-yr contour. Further inspection of the contours in Figure 7 reveals that the vast majority of the points falling outside of the contour are at low significant wave heights. In fact, as shown in the right-most column in Table 5, if we restrict the counting of points outside the contour to those above a threshold of $H_s > 1$ m, the results for all of the contributions are of more similar orders of magnitude. As these environmental contours are most often used in the analysis of extreme condition design load cases, it is reasonable to consider points falling outside of a contour with low significant waves heights to be immaterial. Thus, it can be a sensible model choice to accept low goodness of fit at low H_s values – and consequently many points outside the contour at this area – to increase the goodness of fit at the distribution's upper tail.

4.4. Discussion: exercise 2

The uncertainty study presented in Exercise 2 (see Figure 11 and Figure 12) showed how uncertainty decreases with the available sample size. As expected, we see relatively wide uncertainty bounds for the contours when the ratio of the return period to period of record is large. However, the sampling variance, even if only 1 year of data was used to construct the 50-yr contours, was smaller than the variance among different contour methods (compare Figure 11 and Figure 17).

Note that the kind of sampling uncertainty that was calculated for Exercise 2 ignores the auto-correlation of the environmental variable's time series. Because the points are sampled randomly from the provided dataset, each subsample of, for example, 1 year represents the full sample better than any consecutive 1-yr time period. Consequently, Exercise 2's sampling uncertainty is lower than if samples of consecutive time series would have been used. Figure 16 shows 25 contours that were computed by sampling 1 year of data randomly from the full dataset (top left panel) and 25 contours that were computed from 25 1-yr consecutive time series (top right panel). The 25 contours based on consecutive time series vary to a much greater degree. Some of them deviate up to 8 m/s wind speed and 8 m wave height from the contour that was calculated using all 25 year of data. The variability apparent when consecutive time series are used is more representative to the practical use of environmental contours, however, such an analysis can only be performed for very small sub-sample sizes because the overall sample size of buoy or hindcast data is usually limited to a duration in the order of 10 - 100 years. In Figure 16 the bottom right panel shows five contours that are constructed using consecutive 5-yr time series. The variability among these five contours is already much lower than the contour variability associated with 1-yr time series. Likely the variability would further decrease if longer consecutive time series were used. However, at some point the variability due to a changing climate could dwarf the inter-annual (and inter-n-year) variability associated with a theoretical stationary climate.

Exercise 2 only analyzed one kind of uncertainty. In general, the uncertainty of an environmental contour can be attributed to multiple components:

- uncertainty associated with the quality of the metocean dataset, for example, due to systematic biases in measurements or hindcast models ("dataset quality uncertainty"),
- uncertainty due to limited sample size and sampling variability ("sampling uncertainty"),
- uncertainty associated with choosing a model for the joint distribution (all statistical models that describe wave and wind can be considered to have some degree of model misspecification, the associated uncertainty contributes to overall "joint model uncertainty"),

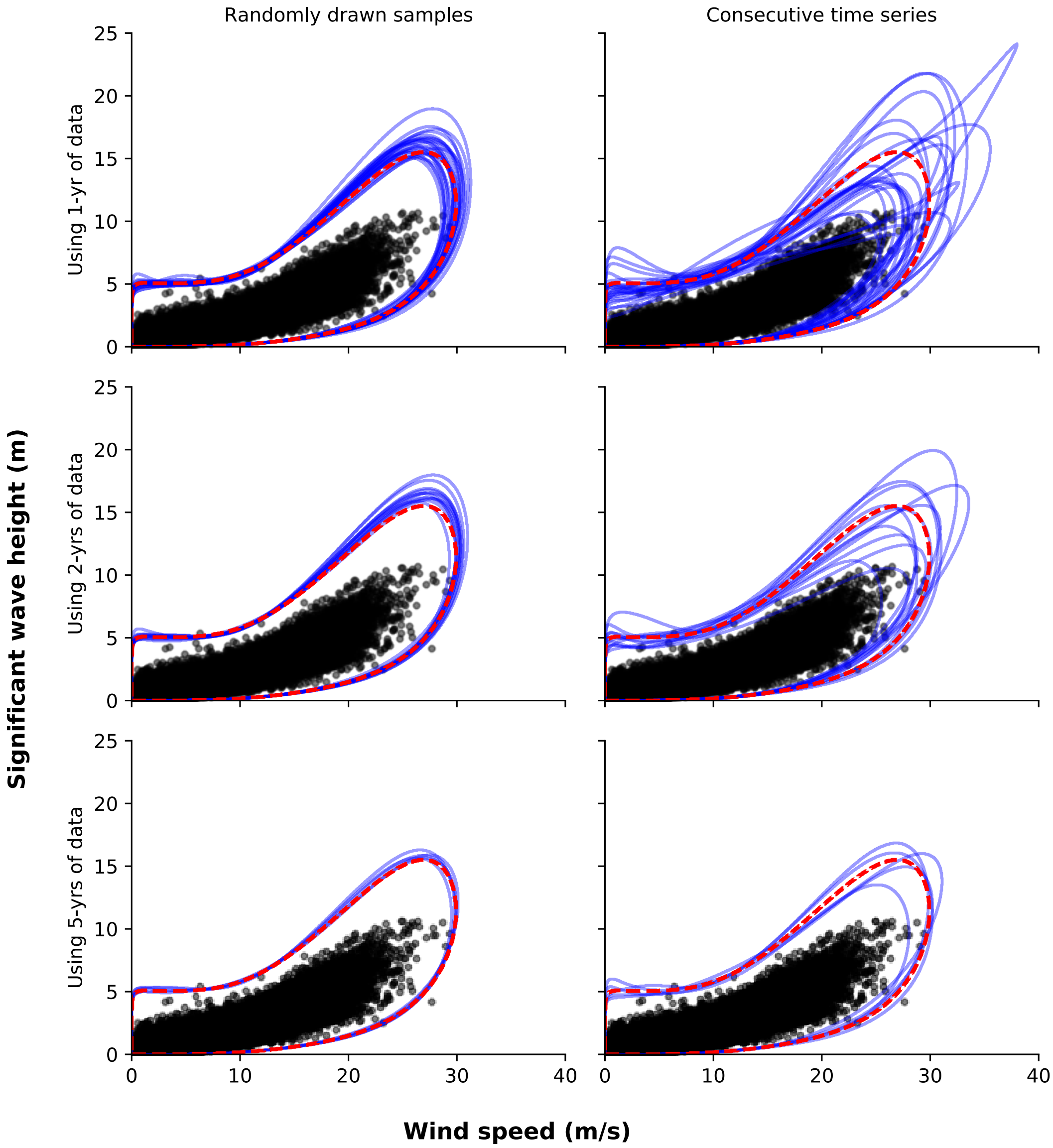


Figure 16: Effect of the metocean data's autocorrelation in Exercise 2. As an example contour method 4 is used. Left: 25 contours that are constructed by randomly sampling 1, 2 and 5 years of observations from dataset D as prescribed in Exercise 2. Right: 25 contours constructed from 1-yr, 2-yr and 5-yr consecutive sequences of dataset D. Continous lines = contours based on 1-yr, 2-yr and 5-yr datasets, dashed line = contour based on the full consecutive 25-yr time series.

- uncertainty associated with choosing a type of parameter estimation technique (for example, maximum likelihood estimation versus the method of moments or least squares estimation; contributes to "joint model uncertainty"),
- uncertainty associated with setting hyper-parameters in the parameter estimation technique (for example, the number of intervals that are used when data are binned; contributes to "joint model uncertainty"),
- uncertainty associated with how the climate will change ("climate uncertainty"),
- uncertainty associated with the numerical methods of a contour construction method, for example, due to the used numeric integration method or due to the applied Monte Carlo method ("numeric uncertainty") and

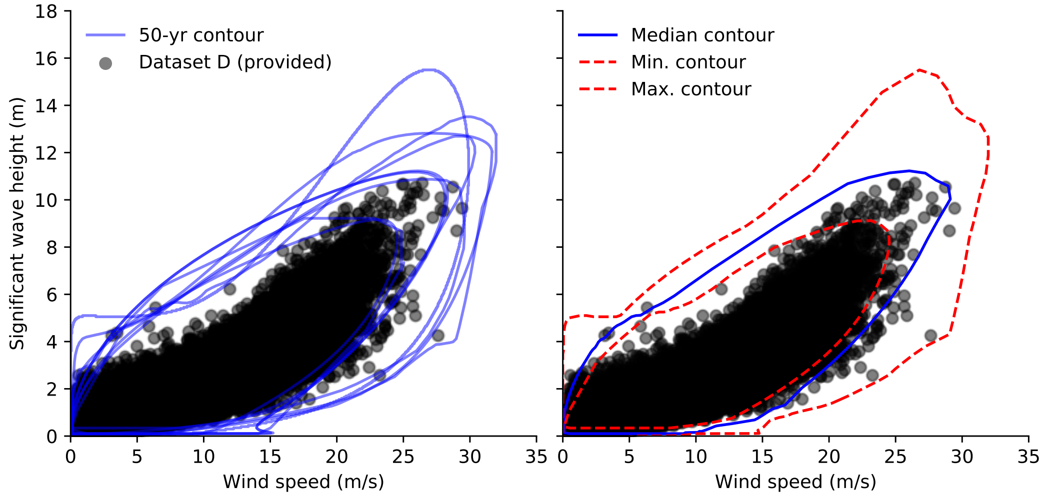


Figure 17: Illustration of the uncertainty associated with choosing a joint distribution model and a contour construction method. Left: 50-yr wind-wave contours that were submitted for Exercise 1. Right: Confidence intervals for these contours, computed using the same geometric algorithm that was proposed for Exercise 2 (0th percentile, 50th percentile and 100th percentile). For Contribution 9, only the unsmoothed direct sampling contour was used.

- uncertainty associated with which type of contour (IFORM, ISORM, ...) shall be constructed to approximate the failure surface to ensure that the contour's exceedance probability and the structure's probability of failure will match ("contour type uncertainty").

This study only directly analyzed the uncertainty due to limited sample size. Indirectly, it also provides insights into joint model uncertainty and contour type uncertainty (Figure 17) because many different models for the joint distribution have been fitted to the same datasets and different contour construction methods were applied. Note, however, that among the contributions joint model structures, parameter estimation techniques and contour construction methods were varied at the same time such that the variability among the contours cannot be pinned directly to one of these factors.

Other studies that analyzed uncertainties of environmental contours include the works of Silva-González, Vázquez-Hernández, Sagrilo and Cuamatzi (2015), Montes-Iturrizaga and Heredia-Zavoni (2017) and Vanem (2018), which focused on joint model uncertainty, the work of Vanem (2015) on climate uncertainty and the works of Gramstad et al. (2018) and Vanem et al. (2019) on sampling uncertainty. In summary, Exercise 2 examined the uncertainty due to the limited length of an available dataset. It allowed comparisons among different contour methods, however, because Exercise 2's methodology did not account for serial correlation, the degree of the "true, practical" sampling variability was not directly assessed. An additional analysis based on consecutive time series suggested that the practical sampling uncertainty is higher: The variability among contours derived from 1-yr to 5-yr consecutive time series was in the same order of magnitude as the variability among the nine submitted contributions (compare Figure 16 and Figure 17).

4.5. Areas for future research

As with any area of research, the improved understanding of environmental contours helps to highlight a number of areas in need of further consideration. As discussed in Section 4.1, serial correlation of sea states is a fundamental issue for environmental contours, in that most contour methods rely on an assumption of independence between samples. It is typical for sea state and weather data to be recorded hourly (or with intervals in the same order of magnitude, such as 10-minute mean wind speeds or 6-hour sea states). Nonetheless, it is well understood that weather patterns and sea states can persist for many hours and even days. From a system dynamics perspective, the ocean and atmosphere are dynamic systems, with inertia, and cannot shift instantly without some memory of their previous state. While some means of controlling for the serial correlation between samples have been proposed (see, for example, Contributions 5 and 6), more can be done to understand the implications of serial correlation of samples on environmental contours and to develop means of controlling for this effect.

Serial correlation can be categorized by the time scale, starting from short-term serial correlation caused by passing weather systems, to seasonal changes and climatic modes on the order of years or decades. Future research could explore the effect these different categories of serial correlation have on environmental contour methods. Short-term serial correlation could be assessed by using a block bootstrap technique, with block lengths of the order of one week. Similarly, the influence of inter-annual variability could be assessed by using block bootstraps with lengths of one year (as briefly explored in Figure 16). Assessing the influence of longer-term climatic variability is more challenging as there is usually insufficient data available to be able to accurately quantify these effects.

As discussed throughout this paper, we have not presented any definitive "correct" solution and performed a test of contours based on this solution. In the formulation of this benchmarking exercise, we considered a number of ways in which such a test might be constructed. One initial idea was to present data from known parametric distributions and construct contours from this "synthetic" data. While this approach does certainly have value, it can also be said that the distribution of environmental data varies: we observe dramatically different distributions for waves in the US Gulf Coast versus the North Atlantic. Another practical means of providing something closer to a definitive "correct" solution for environmental contours would be to utilize a very long climatic simulation dataset such as the 1200-yr dataset analyzed by Jones, Gibson and Shaffrey (2018) or the recently published 700-yr dataset (Song, Bao, Zhang, Shu, Song and Qiao, 2020) that represents a stationary preindustrial climate (Bao, Song and Qiao, 2020). In this way, it would be possible to more directly assess the agreement between a constructed contour and the data. Of course climatic simulations, and especially long simulations with durations such that they cannot be compared with historical measurements, also possess model uncertainty and model error. For the purpose of comparing contour methods, however, the errors in the model data could be assumed to be negligible and the effect of model error could be investigated separately.

Environmental contours are generally used within a larger engineering design workflow, in which the ultimate results are an estimate of a design response. Thus, it is logical that the most important test of an environmental contour is not the contour itself, but its ability to provide an accurate design response estimate. However, many design responses exist and they are typically specific to the structure. For example, one could consider maximum pitch angle in a container vessel, deck slamming in an offshore platform, tower bending moment in an offshore wind turbine, or mooring load in a wave energy converter. Thus, to avoid potentially favoring one engineering system over another, in this exercise a deliberate decision was made to consider only the environmental contours themselves. However, based on the discussions from amongst this paper's authors, a comparison has been made of design responses using these contours (de Hauteclocque et al., in preparation).

In this benchmarking exercise, contours of significant wave height and mean wind speed have been constructed. These two variables represent aggregate statistics of the changing water surface elevation and of wind fluctuations over a certain reference period. Here, significant wave height had a reference period of 1 hour, $H_{s,1h}$, and mean wind speed had a reference period of 10 minutes, U_{10}. Using a combination of 1-hr sea states and 10-min wind speed is typical for offshore wind turbine design (see for example the recommendations in IEC 61400-3-1; International Electrotechnical Commission (2019)). In general, however, combining variables with different reference periods in a joint distribution model raises the question what kind of joint environmental state the distribution represents. It is commonly assumed that the long-term evolution of environmental conditions can be considered as a sequence of stationary processes (Naess and Moan, 2013). That is, the random process associated with a certain environmental condition is assumed to be stationary for fixed time intervals of equal length. To be consistent with this assumption, one would assume that wind and wave represent a joint stationary process with a fixed length such that either wind speed should be converted to represent a 1-hr mean value, U_{1h}, or significant wave height should be converted to represent the intensity of a 10-min sea state, $H_{s,10min}$. While there exist recommendations on how to convert wind speeds with different reference periods using factors (International Electrotechnical Commission, 2019), future research could explore how conversion factors change with wind conditions and how such conversions change the dependence structure of the wind-wave joint distribution. Additionally, if different reference periods are used, they need be handled properly when time-domain simulations are conducted. The current standard on the design of offshore wind turbines, IEC 61400-3-1 (International Electrotechnical Commission, 2019, pp. 64-65), allows designers to either perform one continuous 1-hr simulation or six 10-min simulations and assumes that the maximum response values in these two options are similar.

5. Concluding remarks

Benchmark exercises were defined that sought to allow comparisons of alternative approaches for constructing environmental contours for different metocean datasets. Three datasets comprised 10 years of National Data Buoy Center (NDBC) buoy wave measurements gathered from three locations along the eastern coast of the United States, off the coast of Maine (A), off the coast of Florida (B), and in the Gulf of Mexico (C). These data consisted of hourly significant wave height and zero-up-crossing period data. Participants were asked to derive 1- and 20-yr environmental contours of these two variables. An additional set of three coastDat-2 hindcast datasets comprised 25 years of near-surface 10-minute wind speed and significant wave height values for offshore sites close to Germany (D), the United Kingdom (E), and Norway (F), respectively. Participants were asked to derive 1- and 50-yr environmental contours of these two variables. The development of contours as defined for these six different datasets was the focus of Exercise 1 of this benchmark study. A second exercise, referred to as Exercise 2, focused on characterizing uncertainty in the constructed contours. A total of nine teams offered contributions for Exercise 1 and four of these teams also contributed to Exercise 2.

With respect to Exercise 1, differences in the contours provided by the contributing teams resulted mostly due to different joint distribution models employed for the metocean variables and not as much due to the different methods for contour construction. Additionally, given that the amount of probability content outside some types of environmental contours is known, a discussion on the expected and observed number of “points” outside the contours was presented. While this quantity is not a universally agreed-upon metric of performance of any contour construction and indeed points lying outside a contour are problematic in some regions of the two-dimensional metocean space and not so much in others, this issue was merely remarked upon and it was generally found that points outside derived contours varied by an order of magnitude among the participants.

With regard to Exercise 2, as expected, it was found that constructed contour uncertainty decreased with increase in sample size (or amount of data made available). The uncertainty in contours arises from multiple sources including selecting a type of model for the joint distribution, selecting a parameter estimation technique, selecting a contour construction method as well as the metocean dataset’s finite sample size and climate patterns that are likely changing over time.

Several areas for further exploration have been outlined in a section of this article. The assumption of independent and identically distributed metocean data samples requires some reflection and perhaps some consideration in future contour construction efforts. Another issue worthy of exploration is that related to the use of contours—namely, in design. Accuracy of a contour construction approach may well be different depending on the application and associated limit state or performance functions involved.

Contributions

AFH, RGC and LM, along with the other authors of Haselsteiner et al. (2019), conceived the benchmarking study; AFH, RGC, and LM led the consolidation of results and writing of this paper; participating groups of this benchmarking exercise, listed as co-authors on this article, optionally presented results of their work and some of their work appears in this article as well.

Acknowledgments

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

A. Parameter values of the fitted models

The parameter values of the fitted joint models of each contribution are listed in tables. Table 8, Table 9, Table 10, Table 13 and Table 14 list the parameter values of contributions that used global hierarchical models. Table 11 and Table 12 list the parameter values of Contribution 5’s fitted generalized pareto distributions at various angles.

Table 8
Contribution 1 & 2: Used joint distribution models. Their parameters were estimated in Haselsteiner et al. (2019) to provide baseline results.

Dataset	Significant wave height			Zero-up-crossing period, log-normal distribution					
	Translated Weibull distribution			$\mu_{tz}(h_s) = c_1 + c_2 h_s^{c_3}$			$\sigma_{tz}(h_s) = c_4 + c_5 \exp(c_6 h_s)$		
	α (scale)	β (shape)	γ (location)	c_1	c_2	c_3	c_4	c_5	c_6
A	0.944	1.48	0.0981	1.47	0.214	0.641	0.00	0.308	-0.250
B	1.14	1.60	0.188	1.41	0.234	0.581	0.00	0.241	-0.200
C	1.16	1.56	0.0566	1.24	0.300	0.600	0.00	0.155	-0.161
	Significant wave height			**Wind speed, 2-p. Weibull distribution**					
	Translated Weibull distribution			$\alpha_u(h_s) = c_7 + c_8 h_s^{c_9}$			$\beta_u(h_s) = c_{10} + c_{11} h_s^{c_{12}}$		
	α (scale)	β (shape)	γ (location)	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}
D	1.58	1.41	0.102	0.00	7.58	0.520	0.00	3.89	0.497
E	1.86	1.49	0.122	0.00	7.40	0.525	0.00	3.89	0.398
F	2.57	1.55	0.225	0.00	5.77	0.561	1.97	0.279	1.27

Table 9
Contribution 3: Fitted joint distributions. The 3-parameter log-normal distribution was fitted using the Python software package scipy and its probability density function reads[1] $f(x|\mu,\sigma,\gamma) = \frac{1}{(x-\gamma)\sigma\sqrt{2\pi}} \exp\left(\frac{-[\ln(x-\gamma)-\mu]^2}{2\sigma^2}\right)$.

Dataset	Significant wave height			Zero-up-crossing period, 2-p. log-normal distribution					
	3-p. log-normal distribution			$\mu_{tz}(h_s) = c_1 + c_2 h_s^{c_3}$			$\sigma_{tz}(h_s) = c_4 + c_5 \exp(c_6 h_s)$		
	e^μ (scale)	σ (shape)	γ (location)	c_1	c_2	c_3	c_4	c_5	c_6
A	0.717	0.635	0.0634	1.38	0.302	0.517	-0.191	0.484	-0.114
B	0.972	0.572	0.0633	1.30	0.339	0.464	-0.293	0.519	-0.0596
C	0.937	0.620	-0.0318	1.15	0.389	0.512	-0.958	1.11	-0.0150
	Significant wave height			**Wind speed, 2-p. Weibull distribution**					
	Translated Weibull distribution			$\alpha_u(h_s) = c_7 + c_8 h_s^{c_9}$			$\beta_u(h_s) = c_{10} + c_{11} h_s^{c_{12}}$		
	α (scale)	β (shape)	γ (location)	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}
D	1.58	1.41	0.102	0.00	7.53	0.525	0.00	3.86	0.502
E	1.86	1.49	0.122	0.00	7.33	0.530	0.00	3.85	0.404
F	2.57	1.57	0.2248	0.00	5.71	0.566	1.94	0.292	1.26

[1] https://nbviewer.jupyter.org/url/xweb.geos.ed.ac.uk/~jsteven5/blog/lognormal_distributions.ipynb

Table 10
Contribution 4: Fitted joint distributions.

Dataset	Significant wave height			Zero-up-crossing period, log-normal distribution					
	Exponentiated Weibull distribution			$\mu_{tz}(h_s) = \ln\left(c_1 + c_2\sqrt{\frac{h_s}{g}}\right)$			$\sigma_{tz}(h_s) = c_3 + \frac{c_4}{1 + c_5 h_s}$		
	α (scale)	β (shape)	δ (shape)	c_1	c_2		c_3	c_4	c_5
A	0.207	0.684	7.79	3.62	5.77		0	0.324	0.404
B	0.0988	0.584	36.6	3.54	5.31		0	0.241	0.256
C	0.227	0.697	9.85	2.71	6.51		0.0109	0.147	0.236
	Wind speed			**Sig. wave height, exp. Weibull distribution with $\delta = 5$**					
	Exponentiated Weibull distribution			$\alpha_{hs}(v) = (c_6 + c_7 v^{c_8})/2.0445^{1/\beta_{hs}(v)}$			$\beta_{hs}(v) = c_9 + \frac{c_{10}}{1 + e^{-c_{11}(v - c_{12})}}$		
	α (scale)	β (shape)	δ (shape)	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}
D	10.0	2.42	0.761	0.488	0.0114	2.03	0.714	1.70	0.304	8.77
E	10.8	2.48	0.683	0.617	0.0174	1.87	0.724	2.01	0.309	9.59
F	11.5	2.56	0.534	1.09	0.0251	1.80	0.726	1.89	0.194	13.4

Table 11
Contribution 5 for datasets A and B: Fitted generalized Pareto distributions at various angles. A variable change has been used, $v_1 = H_s$ and $v_2 = H_sT_z$. Additionally, the variables were scaled, s_1 = scale factor for transform variable 1. s_2 = scale factor for transform variable 2. n_e = number of threshold exceedances per year.

Dataset A					Dataset B				
$s_1 = 5.232,\ s_2 = 42.903$					$s_1 = 5.192,\ s_2 = 42.122$				
Angle (deg)	Threshold	Shape	Scale	n_e	Angle (deg)	Threshold	Shape	Scale	n_e
-90.000000	-0.0172	-0.1335	0.0032	4.1994	-90.000000	-0.0291	-0.4002	0.0055	3.9994
-83.571429	-0.0127	-0.2297	0.0029	4.4993	-85.714286	-0.0242	-0.4230	0.0049	4.0994
-77.142857	-0.0068	-0.1300	0.0016	4.3993	-81.428571	-0.0188	-0.3995	0.0039	4.0994
-70.714286	0.0002	-0.6519	0.0020	4.0994	-77.142857	-0.0131	-0.3043	0.0028	4.0994
-64.285714	0.0128	-0.2958	0.0019	4.2993	-72.857143	-0.0068	-0.3090	0.0024	4.1994
-57.857143	0.0347	-0.6208	0.0041	4.0994	-68.571429	0.0010	-0.2072	0.0018	3.8994
-51.428571	0.0688	0.0376	0.0033	4.0994	-60.000000	0.0253	0.0228	0.0019	4.4994
-45.000000	0.1249	-0.1544	0.0068	3.9994	-55.714286	0.0452	-0.0474	0.0031	4.4994
-38.571429	0.2123	-0.4959	0.0216	3.8994	-51.428571	0.0721	-0.1159	0.0059	3.9994
-32.142857	0.3263	-0.4578	0.0467	3.9994	-47.142857	0.1094	-0.0755	0.0067	3.9994
-25.714286	0.4425	-0.4412	0.0849	3.9994	-42.857143	0.1545	-0.1772	0.0121	4.0994
-19.285714	0.5565	-0.3964	0.1226	3.9994	-38.571429	0.2094	0.0091	0.0170	4.1994
-12.857143	0.6622	-0.4241	0.1733	3.8994	-34.285714	0.2699	-0.0271	0.0324	3.9994
-6.428571	0.7584	-0.3979	0.2103	3.9994	0.000000	0.7304	0.3072	0.1269	3.9994
0.000000	0.8434	-0.4166	0.2586	3.9994	13.500000	0.8577	0.3355	0.1583	3.8994
13.500000	1.0078	-0.3572	0.3019	3.9994	27.000000	0.9394	0.4155	0.1609	3.9994
27.000000	1.1092	-0.3586	0.3530	3.9994	40.500000	0.9743	0.5260	0.1457	4.0994
40.500000	1.1596	-0.3267	0.3643	3.9994	54.000000	0.9548	0.4478	0.1682	3.9994
54.000000	1.1509	-0.2794	0.3457	3.9994	67.500000	0.8772	0.4001	0.1768	3.9994
67.500000	1.0792	-0.2290	0.3089	3.9994	81.000000	0.7485	0.2713	0.2005	3.9994
81.000000	0.9491	-0.1525	0.2511	3.9994	94.500000	0.5919	0.3077	0.1629	3.9994
108.000000	0.5064	-0.3332	0.2192	3.9994	108.000000	0.4164	0.4125	0.1090	3.8994
121.500000	0.2724	-0.0597	0.0983	3.9994	121.500000	0.2188	0.5171	0.0625	3.9994
135.000000	0.0488	-0.2796	0.0568	4.0994	135.000000	0.0316	0.1326	0.0553	3.9994
148.500000	-0.0103	-0.3956	0.0041	3.9994	148.500000	-0.0244	0.3415	0.0043	4.2994
162.000000	-0.0228	-0.3171	0.0056	4.1994	162.000000	-0.0403	-0.5340	0.0080	4.0994
175.500000	-0.0306	-0.3086	0.0062	3.8994	175.500000	-0.0513	-0.3806	0.0075	4.0994
189.000000	-0.0366	-0.4625	0.0079	3.8994	189.000000	-0.0589	-0.1053	0.0057	4.2994
202.500000	-0.0395	-0.3103	0.0070	4.0994	202.500000	-0.0642	-0.4043	0.0103	3.9994
216.000000	-0.0407	-0.4262	0.0085	4.1994	216.000000	-0.0643	-0.3623	0.0098	3.9994
229.500000	-0.0383	-0.3675	0.0079	4.0994	229.500000	-0.0597	-0.1953	0.0072	3.7995
243.000000	-0.0334	-0.5277	0.0087	3.6994	243.000000	-0.0532	-0.2990	0.0078	3.8994
256.500000	-0.0265	-0.3223	0.0059	4.0994	256.500000	-0.0436	-0.5041	0.0090	3.9994

Table 12
Contribution 5 for datasets D and E: Fitted generalized Pareto distributions at various angles. A variable change has been used, $v_1 = H_s/(\{10 + 20 \cdot [1 - \cos(U_{10} \cdot \pi/60)^2]\})$ and $v_2 = U_{10}$. Additionally, the variables were scaled, s_1 = scale factor for transform variable 1. s_2 = scale factor for transform variable 2. For brevity only every second angle is reported (for example 0° and 12°, but not 6° and 18°). n_e = number of threshold exceedances per year.

Dataset D					Dataset E			
$s_1 = 0.305,\ s_2 = 21.916$					$s_1 = 0.323,\ s_2 = 22.992$			
Angle (deg)	Threshold	Shape	Scale	n_e	Threshold	Shape	Scale	n_e
0.0	1.0062	-0.3960	0.1245	1.6000	1.0062	-0.3960	0.1245	1.6000
12.0	1.1270	-0.3354	0.1216	1.6000	1.1270	-0.3354	0.1216	1.6000
24.0	1.2197	-0.2425	0.1177	1.6400	1.2197	-0.2425	0.1177	1.6400
36.0	1.2989	-0.2223	0.1071	1.6000	1.2989	-0.2223	0.1071	1.6000
48.0	1.3327	-0.0724	0.0804	1.6000	1.3327	-0.0724	0.0804	1.6000
60.0	1.2993	-0.0755	0.0841	1.6000	1.2993	-0.0755	0.0841	1.6000
72.0	1.2266	-0.0005	0.0739	1.6000	1.2266	-0.0005	0.0739	1.6000
84.0	1.0887	-0.0587	0.0900	1.6400	1.0887	-0.0587	0.0900	1.6400
96.0	0.9357	-0.0093	0.0810	1.6000	0.9357	-0.0093	0.0810	1.6000
108.0	0.7609	0.1624	0.0565	1.6000	0.7609	0.1624	0.0565	1.6000
120.0	0.5546	-0.0641	0.0737	1.6000	0.5546	-0.0641	0.0737	1.6000
132.0	0.3602	0.0683	0.0534	1.5600	0.3602	0.0683	0.0534	1.5600
144.0	0.1770	0.5143	0.0224	1.5600	0.1770	0.5143	0.0224	1.5600
156.0	0.0305	0.1234	0.0225	1.5600	0.0305	0.1234	0.0225	1.5600
168.0	-0.0321	-0.4437	0.0109	1.5600	-0.0321	-0.4437	0.0109	1.5600
180.0	-0.0607	-0.2366	0.0100	1.6000	-0.0607	-0.2366	0.0100	1.6000
192.0	-0.0754	-0.3358	0.0126	1.6000	-0.0754	-0.3358	0.0126	1.6000
204.0	-0.0827	-0.5789	0.0205	1.5600	-0.0827	-0.5789	0.0205	1.5600
216.0	-0.0826	-0.5009	0.0198	1.6400	-0.0826	-0.5009	0.0198	1.6400
228.0	-0.0766	-0.5970	0.0214	1.6000	-0.0766	-0.5970	0.0214	1.6000
240.0	-0.0663	-0.7225	0.0228	1.6000	-0.0663	-0.7225	0.0228	1.6000
252.0	-0.0501	-0.7629	0.0183	1.5600	-0.0501	-0.7629	0.0183	1.5600
264.0	-0.0283	-0.1749	0.0054	1.6400	-0.0283	-0.1749	0.0054	1.6400
276.0	0.0316	-0.3399	0.0114	1.5200	0.0316	-0.3399	0.0114	1.5200
288.0	0.1681	-0.3577	0.0471	1.6000	0.1681	-0.3577	0.0471	1.6000
300.0	0.3126	-0.2949	0.0826	1.6000	0.3126	-0.2949	0.0826	1.6000
312.0	0.4623	-0.2377	0.1030	1.6000	0.4623	-0.2377	0.1030	1.6000
324.0	0.6219	-0.0607	0.0868	1.6000	0.6219	-0.0607	0.0868	1.6000
336.0	0.7586	0.0424	0.0779	1.6000	0.7586	0.0424	0.0779	1.6000
348.0	0.8760	-0.1831	0.1055	1.6000	0.8760	-0.1831	0.1055	1.6000

Table 13
Contribution 8: Fitted joint distributions.

Dataset	**Significant wave height**		**Zero-up-crossing period, 2-p. Weibull distribution**						
	2-p. Weibull distribution		$\alpha_{tz}(h_s) = c_1 h_s^2 + c_2 h_s + c_3$			$\beta_{tz}(h_s) = c_4 h_s^2 + c_5 h_s + c_6$			
	α (scale)	β (shape)	c_1	c_2	c_3	c_4	c_5	c_6	
A	1.0651	1.6399	-0.0245	0.4387	-0.4702	0.07056	-1.132	5.787	
B	1.3654	1.8893	-0.02063	0.4798	0.5028	0.04746	-0.7643	4.963	
C	5.0475	5.5633	0.135	-0.824	1.964	0.06753	-0.6624	3.637	

	Significant wave height		**Wind speed, 2-p. Weibull distribution**								
	2-p. Weibull distribution		$\alpha_u(h_s) = c_7 h_s^3 + c_8 h_s^2 + c_9 h_s + c_{10}$				$\beta_u(h_s) = c_{11} h_s^3 + c_{12} h_s^2 + c_{13} h_s + c_{14}$			
	α (scale)	β (shape)	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}	c_{13}	c_{14}
D	1.7148	1.5292	0.01211	-0.3048	3.972	3.496	0.01414	-0.177	1.848	2.35
E	2.018	1.6249	0	-0.1399	3.412	3.621	0	-0.008746	0.9706	1.718
F	2.849	1.716	0	-0.09185	2.735	2.825	0	-0.002556	0.5414	1.327

Table 14
Contribution 9: Fitted joint distributions.

Dataset	Significant wave height			Zero-up-crossing period, lognormal distribution					
	Translated Weibull distribution			$\mu_{tz}(h_s) = c_1 + c_2 h_s^{c_3}$			$\sigma_{tz}(h_s) = c_4 + c_5 \exp(c_6 h_s)$		
	α (scale)	β (shape)	γ (location)	c_1	c_2	c_3	c_4	c_5	c_6
A	0.4983	0.8573	0.4187	1.4306	0.2561	0.5556	0.0150	0.3004	-0.2884
B	0.6539	0.9710	0.5658	1.3805	0.2686	0.5254	0.0150	0.2311	-0.2339
C	0.7291	1.0134	0.3910	0.7702	0.8061	0.2624	0.0150	0.1452	-0.2069
	Significant wave height			**Wind speed, 2-parameter Weibull distribution**					
	Translated Weibull distribution			$\alpha_u(h_s) = c_7 + c_8 h_s^{c_9}$			$\beta_u(h_s) = c_{10} + c_{11} h_s^{c_{12}}$		
	α (scale)	β (shape)	γ (location)	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}
D	1.2528	1.1186	0.3389	2.3134	1.2987	1.0594	-1.3969	8.5546	0.5129
E	1.4836	1.1963	0.4101	2.1616	1.0011	1.1721	-0.3624	6.7912	0.6232
F	2.0550	1.2284	0.6104	1.9762	0.2397	1.3660	-1.3854	6.0949	0.5972

Table 15
The results presented in this study can be reproduced by running the listed open-source Python files.

Object	Content	URL
Figure 7	E1 sea state contour overlay	https://github.com/ec-benchmark-organizers/ec-benchmark/blob/master/results/exercise-1/plot_benchmark_contours_dataset_abc.py
Figure 8	E1 maxima along the sea state contours	https://github.com/ec-benchmark-organizers/ec-benchmark/blob/master/results/exercise-1/plot_e1_maxima.py
Table 5	E1 sea state points outside	https://github.com/ec-benchmark-organizers/ec-benchmark/blob/master/results/exercise-1/create_points_outside_table_abc.py
Figure 9	E1 wind-wave contour overlay	https://github.com/ec-benchmark-organizers/ec-benchmark/blob/master/results/exercise-1/plot_benchmark_contours_dataset_def.py
Figure 10	E1 maxima along the wind-wave contours	https://github.com/ec-benchmark-organizers/ec-benchmark/blob/master/results/exercise-1/plot_e1_maxima.py
Table 6	E1 wind-wave points outside	https://github.com/ec-benchmark-organizers/ec-benchmark/blob/master/results/exercise-1/create_points_outside_table_def.py
Figure 12	E2 confidence intervals	https://github.com/ec-benchmark-organizers/ec-benchmark/blob/master/results/exercise-2/plot_benchmark_contours_e2.py
Figure 15	Autocorrelation function of H_s and U_{10}	https://github.com/ec-benchmark-organizers/ec-benchmark/blob/master/results/discussion/autocorrelation_of_datasets.py
Table 7	Upper tail dependence	https://github.com/ec-benchmark-organizers/ec-benchmark/blob/master/results/discussion/extremal_dependence.py
Figure 16	Effect of autocorrelation on the results of E2	https://github.com/ec-benchmark-organizers/ec-benchmark/blob/master/results/discussion/e2_autocorrelation.py
Figure 17	Interpreting the submitted contours as model uncertainty.	https://github.com/ec-benchmark-organizers/ec-benchmark/blob/master/results/discussion/e1_confidence_bound.py
Provided datasets		https://github.com/ec-benchmark-organizers/ec-benchmark/tree/master/datasets
Retained datasets		https://github.com/ec-benchmark-organizers/ec-benchmark/tree/master/datasets-retained

B. Reproducing the results of this study

The figures and tables that presented the major results of this study can be reproduced by running open-source Python scripts that are available at a dedicated Github repository. The repository is available at `https://github.com/ec-benchmark-organizers/ec-benchmark` . Table 15 lists the individual scripts that were used to produce the presented figures and tables. Additionally, a procedure for applying these benchmark exercises to a new contour are provided in the Github repository. The steps to do so are detailed in a `README.md` file in the exercise-1 folder: `https://github.com/ec-benchmark-organizers/ec-benchmark/blob/master/results/exercise-1/readme.md`.

References

Baarholm, G.S., Haver, S., Økland, O.D., 2010. Combining contours of significant wave height and peak period with platform response distributions for predicting design response. Marine Structures 23, 147–163. doi:10.1016/j.marstruc.2010.03.001.

Bao, Y., Song, Z., Qiao, F., 2020. FIO-ESM version 2.0: Model description and evaluation. Journal of Geophysical Research: Oceans 125, 1–21. doi:10.1029/2019JC016036.

Bitner-Gregersen, E.M., 2015. Joint met-ocean description for design and operations of marine structures. Applied Ocean Research 51, 279 – 292. doi:10.1016/j.apor.2015.01.007.

Bitner-Gregersen, E.M., Haver, S., 1991. Joint environmental model for reliability calculations, in: Proc. 1st International Offshore and Polar Engineering Conference (ISOPE 1991), Edinburgh, United Kingdom. pp. 246–253.

Chai, W., Leira, B.J., 2018. Environmental contours based on inverse SORM. Marine Structures 60, 34 – 51. doi:10.1016/j.marstruc.2018.03.007.

Cheng, Z., Svangstu, E., Moan, T., Gao, Z., 2019. Long-term joint distribution of environmental conditions in a Norwegian fjord for design of floating bridges. Ocean Engineering 191, 106472. doi:10.1016/j.oceaneng.2019.106472.

Coles, S., 2001. An introduction to statistical modeling of extreme values. Springer, London; New York.

Coles, S., Heffernan, J., Tawn, J., 1999. Dependence measures for multivariate extremes. Extremes 2, 339–365. doi:10.1023/A:1009963131610.

Derbanne, Q., de Hauteclocque, G., 2019. A new approach for environmental contour and multivariate de-clustering. doi:10.1115/OMAE2019-95993.

Dimitrov, N., 2020. Inverse directional simulation: An environmental contour method providing an exact return period. Journal of Physics: Conference Series doi:10.1088/1742-6596/1618/6/062048.

DNV GL, 2017. Recommended practice DNVGL-RP-C205: Environmental conditions and environmental loads. Technical Report.

Eckert, A., Martin, N., Coe, R.G., Seng, B., Stuart, Z., Morrell, Z., 2021. Development of a comparison framework for evaluating environmental contours of extreme sea states. Journal of Marine Science and Engineering 9. doi:10.3390/jmse9010016.

Eckert-Gallup, A., Martin, N., 2016. Kernel density estimation (KDE) with adaptive bandwidth selection for environmental contours of extreme sea states, in: OCEANS 2016 MTS/IEEE Monterey, IEEE, Monterey, CA, USA. pp. 1–5. doi:10.1109/OCEANS.2016.7761150.

Eckert-Gallup, A.C., Sallaberry, C.J., Dallman, A.R., Neary, V.S., 2014. Modified Inverse First Order Reliability Method (I-FORM) for Predicting Extreme Sea States. Technical Report SAND2014-17550. Sandia National Laboratories. Albuquerque, NM (United States).

Fazeres-Ferradosa, T., Taveira-Pinto, F., Vanem, E., Reis, M.T., das Neves, L., 2018. Asymmetric copula-based distribution models for met-ocean data in offshore wind engineering applications. Wind Engineering 42, 304–334.

Ferreira, J.A., Guedes Soares, C., 2002. Modelling bivariate distributions of significant wave height and mean wave period. Applied Ocean Research 24, 31–45. doi:10.1016/S0141-1187(02)00006-8.

Gramstad, O., Vanem, E., Bitner-Gregersen, E.M., 2018. Uncertainty of environmental contours due to sampling variability, in: Proc. 37th International Conference on Ocean, Offshore and Arctic Engineering (OMAE 2018), American Society of Mechanical Engineers (ASME), Madrid, Spain. doi:10.1115/OMAE2018-77810.

Groll, N., Weisse, R., 2016. coastDat-2 North Sea wave hindcast for the period 1949-2014 performed with the wave model WAM. doi:10.1594/WDCC/coastDat-2_WAM-North_Sea.

Groll, N., Weisse, R., 2017. A multi-decadal wind-wave hindcast for the North Sea 1949 -– 2014: coastDat2. Earth System Science Data 9, 955–968. doi:10.1594/WDCC/coastDat-2.

Guedes Soares, C., 1993. Long term distribution of non-linear wave induced vertical bending moments. Marine Structures 6, 475–483. doi:10.1016/0951-8339(93)90033-Y.

Guedes Soares, C., Henriques, A.C., 1996. Statistical uncertainty in long-term distributions of significant wave height. Journal of Offshore Mechanics and Arctic Engineering 118, 284–291. doi:10.1115/1.2833917.

Guedes Soares, C., Lopes, L., Costa, M., 1988. Wave climate modelling for engineering purposes, in: Schreffler, B.A., Zienkiewicz, O.C. (Eds.), Computer Modelling in Ocean Engineering. Rotterdam: A.A. Balkema Pub., pp. 169–175.

Hannesdóttir, Á., Kelly, M., Dimitrov, N., 2019. Extreme fluctuations of wind speed for a coastal/offshore climate: statistics and impact on wind turbine loads. Wind Energy Science 4, 325–342. doi:10.5194/wes-4-325-2019.

Haselsteiner, A., Ohlendorf, J.H., Thoben, K.D., 2017a. Environmental contours based on kernel density estimation, in: Proceedings of the 13th German Wind Energy Conference DEWEK 2017, Bremen, Germany.

Haselsteiner, A.F., Coe, R.G., Manuel, L., Nguyen, P.T.T., Martin, N., Eckert-Gallup, A., 2019. A benchmarking exercise on estimating extreme environmental conditions: Methodology & baseline results, in: Proc. 38th International Conference on Ocean, Offshore and Arctic Engineering (OMAE 2019), American Society of Mechanical Engineers (ASME), Glasgow, UK. doi:10.1115/OMAE2019-96523.

Haselsteiner, A.F., Ohlendorf, J.H., Wosniok, W., Thoben, K.D., 2017b. Deriving environmental contours from highest density regions. Coastal Engineering 123, 42–51. doi:10.1016/j.coastaleng.2017.03.002.

Haselsteiner, A.F., Sander, A., Ohlendorf, J.H., Thoben, K.D., 2020. Global hierarchical models for wind and wave contours: Physical interpretations of the dependence functions, in: Proc. 39th International Conference on Ocean, Offshore and Arctic Engineering (OMAE2020).

Haselsteiner, A.F., Thoben, K.D., 2020. Predicting wave heights for marine design by prioritizing extreme events in a global model. Renewable Energy 156, 1146–1157. doi:10.1016/j.renene.2020.04.112.

de Hauteclocque, G., Mackay, E., Vanem, E., in preparation. Quantitative assessment of environmental contour approaches .

Haver, S., 1985. Wave climate off northern Norway. Applied Ocean Research 7, 85–92. doi:10.1016/0141-1187(85)90038-0.

Haver, S., 1987. On the joint distribution of heights and periods of sea waves. Ocean Engineering 14, 359–376. doi:10.1016/0029-8018(87)90050-3.

Haver, S., Winterstein, S.R., 2009. 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 116, 116–127.

Heffernan, J.E., Tawn, J.A., 2004. A conditional approach for multivariate extreme values. Journal of the Royal Statistical Society. Series B (Statistical Methodology) 66, 497–546. doi:10.1111/j.1467-9868.2004.02050.x.

Heredia-Zavoni, E., Montes-Iturrizaga, R., 2019. Modeling directional environmental contours using three dimensional vine copulas. Ocean Engineering 187. doi:10.1016/j.oceaneng.2019.06.007.

Horn, J.T., Bitner-Gregersen, E., Krokstad, J., Leira, B.J., Amdahl, J., 2018. A new combination of conditional environmental distributions. Applied Ocean Research 73, 17–26.

Huseby, A.B., Vanem, E., Eskeland, K., 2017. Evaluating properties of environmental contours, in: Chepin, M., Bris, R. (Eds.), Safety and Reliability, Theory and Applications. Proceedings of the European Safety and Reliability Conference (ESREL 2017). doi:10.1201/9781315210469-265.

Huseby, A.B., Vanem, E., Natvig, B., 2013. A new approach to environmental contours for ocean engineering applications based on direct Monte Carlo simulations. Ocean Engineering 60, 124–135. doi:10.1016/j.oceaneng.2012.12.034.

Huseby, A.B., Vanem, E., Natvig, B., 2014. A new Monte Carlo method for environmental contour estimation, in: Proc. ESREL 2014, European Safety and Reliability Association (ESRA).

Huseby, A.B., Vanem, E., Natvig, B., 2015. Alternative environmental contours for structural reliability analysis. Structural Safety 54, 32 – 45. doi:https://doi.org/10.1016/j.strusafe.2014.12.003.

International Electrotechnical Commission, 2019. Wind energy generation systems - Part 3-1: Design requirements for fixed offshore wind turbines. Technical Report IEC 61400-3-1.

Jonathan, P., Ewans, K., Flynn, J., 2014. On the estimation of ocean engineering design contours. Journal of Offshore Mechanics and Arctic Engineering 136, 41101–1 to 041101–8. doi:10.1115/1.4027645.

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

Jones, O.P., Gibson, R., Shaffrey, L., 2018. Estimating low probability events using long climate simulations, in: Proc. 2018 Offshore Structural Reliability Conference, API, Houston, TX, USA.

Karmakar, D., Bagbanci, H., Guedes Soares, C., 2016. Long-term extreme load prediction of spar and semisubmersible floating wind turbines using the environmental contour method. Journal of Offshore Mechanics and Arctic Engineering 138, 021601–1 to 021601–9. doi:10.1115/1.4032099.

Leira, B.J., 2008. A comparison of stochastic process models for definition of design contours. Structural Safety 30, 493–505.

Li, L., Gao, Z., Moan, T., 2015. Joint environmental data at five European offshore sites for design of combined wind and wave energy devices. Journal of Offshore Mechanics and Arctic Engineering 137, 031901–1 to 031901–16. doi:10.1115/1.4029842.

Li, L., Yuan, Z.M., Gao, Y., Zhang, X., Tezdogan, T., 2019. Investigation on long-term extreme response of an integrated offshore renewable energy device with a modified environmental contour method. Renewable Energy 132, 33–42. doi:10.1016/j.renene.2018.07.138.

Li, Q., Gao, Z., Moan, T., 2016. Modified environmental contour method for predicting long-term extreme responses of bottom-fixed offshore wind turbines. Marine Structures 48, 15–32. doi:10.1016/j.marstruc.2016.03.003.

Lin, Y., Dong, S., Tao, S., 2020. Modelling long-term joint distribution of significant wave height and mean zero-crossing wave period using a copula mixture. Ocean Engineering 197, 106856. doi:10.1016/j.oceaneng.2019.106856.

Liu, J., Thomas, E., Goyal, A., Manuel, L., 2019. Design loads for a large wind turbine supported by a semi-submersible floating platform. Renewable Energy 138, 923 – 936. doi:https://doi.org/10.1016/j.renene.2019.02.011.

Lucas, C., Guedes Soares, C., 2015. Bivariate distributions of significant wave height and mean wave period of combined sea states. Ocean Engineering 106, 341–353. doi:10.1016/j.oceaneng.2015.07.010.

Mackay, E., Haselsteiner, A.F., 2021. Marginal and total exceedance probabilities of environmental contours. Marine Structures 75. doi:10.1016/j.marstruc.2020.102863.

Mackay, E.B., Jonathan, P., 2020. 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).

Manuel, L., Nguyen, P.T., Canning, J., Coe, R.G., Eckert-Gallup, A.C., Martin, N., 2018. Alternative approaches to develop environmental contours from metocean data. Journal of Ocean Engineering and Marine Energy 4, 293–310. doi:10.1007/s40722-018-0123-0.

Mathisen, J., Bitner-Gregersen, E., 1990. Joint distributions for significant wave height and wave zero-up-crossing period. Applied Ocean Research 12, 93–103. doi:10.1016/S0141-1187(05)80033-1.

Montes-Iturrizaga, R., Heredia-Zavoni, E., 2017. Assessment of uncertainty in environmental contours due to parametric uncertainty in models of the dependence structure between metocean variables. Applied Ocean Research 64. doi:10.1016/j.apor.2017.02.006.

Muliawan, M.J., Gao, Z., Moan, T., 2013. Application of the contour line method for estimating extreme responses in the mooring lines of a two-body floating wave energy converter. Journal of Offshore Mechanics and Arctic Engineering 135, 031301–1 to 031301–10. doi:10.1115/1.4024267.

Naess, A., Moan, T., 2013. Random environmental process, in: Stochastic dynamics of marine structures. Cambridge University Press, Cambridge, United Kingdom, pp. 191–208. doi:10.1017/CBO9781139021364.

NORSOK, 2007. NORSOK standard N-003: Actions and action effects. Technical Report.

Rendon, E.A., Manuel, L., 2014. Long-term loads for a monopile-supported offshore wind turbine. Wind Energy 17, 209–223. doi:10.1002/we.1569.

Ross, E., Astrup, O.C., Bitner-Gregersen, E., Bunn, N., Feld, G., Gouldby, B., Huseby, A., Liu, Y., Randell, D., Vanem, E., Jonathan, P., 2019. On environmental contours for marine and coastal design. Ocean Engineering doi:10.1016/j.oceaneng.2019.106194.

Saranyasoontorn, K., Manuel, L., 2006. Design loads for wind turbines using the environmental contour method. Journal of Solar Energy Engineering 128, 554–561. doi:10.1115/1.2346700.

Silva-González, F., Vázquez-Hernández, A., Sagrilo, L., Cuamatzi, R., 2015. 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 53, 190–199. doi:10.1016/j.apor.2015.09.005.

Song, Z., Bao, Y., Zhang, D., Shu, Q., Song, Y., Qiao, F., 2020. Centuries of monthly and 3-hourly global ocean wave data for past, present, and future climate research. Scientific Data 7, 1–11. doi:10.1038/s41597-020-0566-8.

Vanem, E., 2015. Uncertainties in extreme value modelling of wave data in a climate change perspective. Journal of Ocean Engineering and Marine Energy 1, 339–359. doi:10.1007/s40722-015-0025-3.

Vanem, E., 2016. Joint statistical models for significant wave height and wave period in a changing climate. Marine Structures 49, 180–205.

doi:10.1016/j.marstruc.2016.06.001.

Vanem, E., 2017. A comparison study on the estimation of extreme structural response from different environmental contour methods. Marine Structures 56, 137–162. doi:10.1016/j.marstruc.2017.07.002.

Vanem, E., 2018. A simple approach to account for seasonality in the description of extreme ocean environments. Marine Systems & Ocean Technology 13, 63–73. doi:10.1007/s40868-018-0046-6.

Vanem, E., 2019. 3-dimensional environmental contours based on a direct sampling method for structural reliability analysis of ships and offshore structures. Ships and Offshore Structures 14. doi:10.1080/17445302.2018.1478377.

Vanem, E., Bitner-Gregersen, E.M., 2015. Alternative environmental contours for marine structural design- a comparison study. Journal of Offshore Mechanics and Arctic Engineering 137, 51601–1 to 51601–8.

Vanem, E., Gramstad, O., Bitner-Gregersen, E.M., 2019. A simulation study on the uncertainty of environmental contours due to sampling variability for different estimation methods. Applied Ocean Research 91, 1–15. doi:10.1016/j.apor.2019.101870.

Vanem, E., Guo, B., Ross, E., Jonathan, P., 2020. Comparing different contour methods with response-based methods for extreme ship response analysis. Marine Structures 69. doi:10.1016/j.marstruc.2019.102680.

Vanem, E., Huseby, A.B., 2018. Combined long-term and short-term description of extreme ocean wave conditions by 3-dimensional environmental contours, in: Proc. 28th International Ocean and Polar Engineering conference (ISOPE 2018), The International Society of Offshore and Polar Engineering (ISOPE).

Vanem, E., Huseby, A.B., 2020. 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).

Velarde, J., Vanem, E., Kramhøft, C., Dalsgaard, J., 2019. Probabilistic analysis of offshore wind turbines under extreme resonant response: Application of environmental contour method. Applied Ocean Research 93, 101947–1 to 101947–16. doi:10.1016/j.apor.2019.101947.

Wang, S., Wang, X., Woo, W.L., 2018. A comparison of response-based analysis and environmental contour methods for FPSO green water assessment, in: Proc. 37th International Conference on Ocean, Offshore and Arctic Engineering (OMAE 2018), American Society of Mechanical Engineers (ASME), Madrid, Spain. doi:10.1115/OMAE2018-77841.

Winterstein, S.R., Ude, T.C., Cornell, C.A., Bjerager, P., Haver, S., 1993. Environmental parameters for extreme response: Inverse FORM with omission factors, in: Proceedings of the 6th International Conference on Structural Safety & Reliability (ICOSSAR), Innsbruck, Austria.

Zhang, Y., Kim, C.W., Beer, M., Dai, H., Guedes Soares, C., 2018. Modeling multivariate ocean data using asymmetric copulas. Coastal Engineering 135, 91–111. doi:10.1016/j.coastaleng.2018.01.008.