Ocean Engineering 266 (2022) 113032

Contents lists available at ScienceDirect

Ocean Engineering

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

On limiting the influence of serial correlation in metocean data for prediction of extreme return levels and environmental contours

Ryan G. Coe [a,*], Lance Manuel [b], Andreas F. Haselsteiner [c]

[a] Sandia National Laboratories, Albuquerque, NM, USA
[b] The University of Texas at Austin, Austin, TX, USA
[c] University of Bremen, Bremen, Germany

ARTICLE INFO

Keywords:
Environmental contour
Metocean extremes
Joint distribution
Serial correlation
Storms

ABSTRACT

Metocean conditions change slowly, over the course of hours, sometimes even days, as storms develop and swells travel across the globe. Thus, measurements of these conditions are often serially correlated. However, many commonly employed methods for predicting metocean conditions for engineering design analyses are built upon an assumption of statistical independence of the data (e.g., hourly significant wave heights). In this brief study, we present an assessment of the serial (temporal) dependence in a selected metocean dataset. A method for processing a dataset that identifies and groups data sequences as "storm" events, and thus reduces serial dependence, is proposed and tested for estimating extreme metocean return levels. The results of this study show that the proposed procedure does indeed limit dependence and that environmental contours produced using this storm grouping procedure are reasonable based on the original dataset and when compared with associated alternative contours that ignore temporal dependence.

1. Introduction

The ocean is a dynamic system, in which currents, winds, and waves are processes with memory/inertia. Ocean storms do not last for a single hour and then instantly transition to an unrelated (statistically independent) condition, instead they build over time and persist for many hours, often even days. However, to inform engineering design analyses, metocean measurements and/or simulation outputs are often taken hourly or every three hours. Treating serial metocean observations as statistically independent observations, a common assumption in the statistical methods employed in engineering analyses, is questionable at best.

In ocean engineering, we are often interested in characterizing extreme environmental events that cause immediate structural failure (so-called "ultimate limit states"). Thus, metocean extremes are often defined as "return levels" that are exceeded at a given "return period". When ultimate limit states are considered, this return period should be defined based on storms peaks, as opposed to all 1-hour sea state observations: The return period, T_r, is the average time between two consecutive storm peak events exceeding the threshold x. The threshold is called return level. Note that this storm peak definition leads to different return values for a given period than if all metocean data were used. This is due to the serial correlation of hourly metocean data. The time series of hourly observations that makes up a storm contains many individual observations. For an ultimate limit analysis, these joint-occurring observations, however, should be treated as a single event: If the structure fails due to the conditions at 2 PM, another failure from the subsequent sample at 3 PM is likely not relevant. Consequently, if all hourly observations are used to calculate return levels too many extreme events are included and, consequently, return levels for a given return period are overestimated. Failure to account for the presence of serial correlation in metocean data will result in an over-prediction of extreme return levels (see, e.g., Beirlant et al., 2004), and thus an over-design of structures.

While many common methods and engineering workflows do treat serial metocean conditions as statistically independent, accounting for serial correlation in metocean data is by no means new (see, e.g., Ochi, 2005; Jonathan and Ewans, 2013). Perhaps the most straightforward approach to limit the influence of dependence in time series data is to downsample the data, considering, for example, only every third measurement from recorded hourly samples (see, e.g., Vanem, 2018). Other methods employ block maxima and or peak-over-threshold data subsets (Coles, 2001; Ferro and Segers, 2003; Leford and Tawn, 1996). However, as discussed by Fawcett and Walshaw (2012), such techniques inevitably reduce the quantity of data used and, thus, decrease the precision of return level predictions at target return period levels.

* Corresponding author.
E-mail addresses: rcoe@sandia.gov (R.G. Coe), lmanuel@mail.utexas.edu (L. Manuel), a.haselsteiner@uni-bremen.de (A.F. Haselsteiner).

https://doi.org/10.1016/j.oceaneng.2022.113032
Received 18 June 2021; Received in revised form 25 October 2022; Accepted 27 October 2022
Available online 24 November 2022

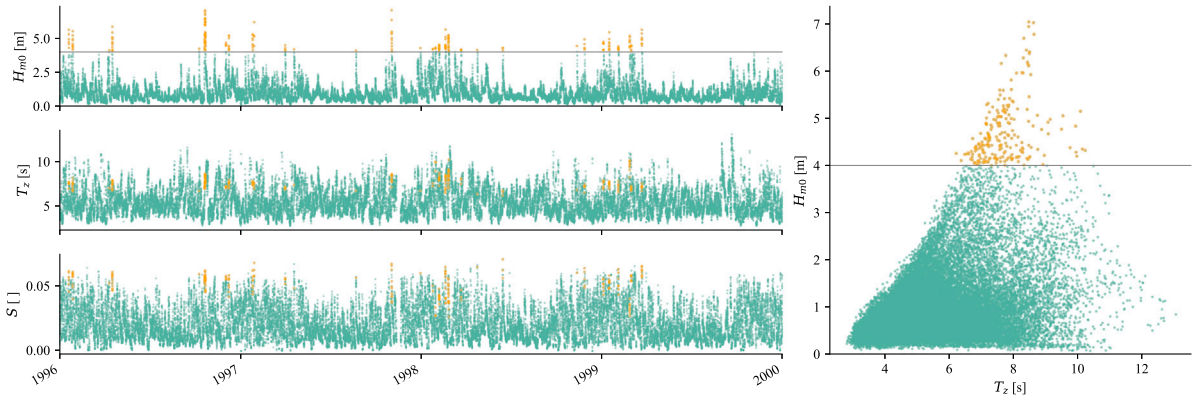


Fig. 1. Time series and scatter diagrams of data chosen from NDBC 44007 showing significant wave height (H_{m0}), zero-up-crossing period (T_z), and spectral wave steepness (S). Orange markers indicate sea states where $H_{m0} > 4$ m. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

A means of generating extreme environmental contours defined in terms of environmental variables based on storm peaks was proposed by Mackay and Jonathan (2020), who utilized block maxima for wave height and accounted for possibly non-coincident wave periods in a resampling step in order to generate contours. Derbanne and de Hauteclocque (2019) proposed a contour method in which data were first projected at lines at various angles in the 2D variable space and then peak-over-threshold was used to decluster the projected 1D data. Fogle et al. (2008) examined methods outlined in the IEC wind turbine design guidelines (IEC 61400-1, IEC, 2019), while specifically considering the effectiveness of block maxima methods in limiting the influence of serial dependence. While Fogle et al. (2008) did identify noticeably different levels of serial correlation depending on the selected block size when using the block maxima method, this effect did not necessarily translate to significant effects on extreme value prediction.

In this brief study, we examine the degree to which hourly measurements from a wave buoy exhibit serial dependence and propose a means of processing the measurements to reduce the influence of that dependence and support the prediction of extreme sea state contours. First, we assess the serial dependence of a case-study dataset. Next, based on the idea that a metocean statistical analysis might be formulated based on distinct "storm" events, we propose a relatively simple method for selecting such storms to generate a new dataset on which to operate. Finally, we compare the original dataset with several distilled storm datasets by comparing levels of serial correlation, the probability distributions of each dataset, and environmental contours produced using each dataset.

2. Analysis

2.1. Data

This study was conducted using data from NDBC 44007,[1] which is a wave measurement buoy located off the coast of Maine. This same dataset was also used in the recent "EC Benchmarking" exercise (Haselsteiner et al., 2019a, 2021). In particular, we consider hourly observations of zero-up-crossing period (T_z) and significant wave height (H_{m0}). Additionally, we present the spectral wave steepness (Myrhaug, 2018), which is given by

$$S = \frac{2\pi H_{m0}}{g T_z^2}, \tag{1}$$

where g is the acceleration due to gravity. As we are concerned with the temporal correlation in the measurements, we select a mostly contiguous four-year period[2] of collected data beginning in 1996. Time series and scatter plots of the data are shown in Fig. 1.

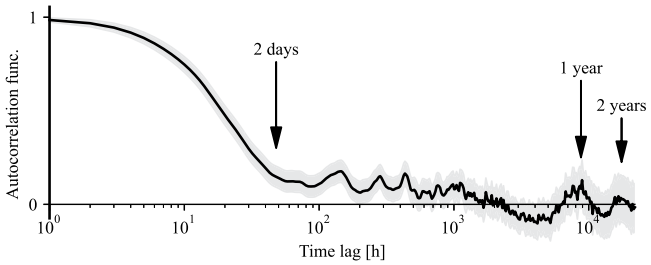


Fig. 2. Estimated autocorrelation function of hourly H_{m0} values from the selected dataset with 95% confidence interval computed via Bartlett's formula.

From Fig. 1, it is clear to see that large wave measurements are often closely spaced temporally. Choosing an arbitrary threshold of 4 m, we can see that sea states with significant wave heights larger than this level are fairly rare. In total, there are 217 observations with $H_{m0} > 4$ m (0.6% of the four-year dataset). These 217 observations occur in fairly small number of tightly-grouped clusters. This strongly grouped nature of the extreme sea states supports the idea that metocean conditions are indeed serially dependent to some degree.

2.2. Testing for serial correlation

The estimated autocorrelation function of hourly H_{m0} values from the selected dataset is presented in Fig. 2 with a 95% confidence interval computed via Bartlett's formula (see,e.g., Brockwell and Davis, 2009). We can note the strong serial correlation for time lags up to roughly two days. It is also clear to see an annual harmonic, reflecting the seasonal nature of ocean waves (e.g., with largest waves generally occurring each winter).

Elaborating on our qualitative observation about the temporal cluster of observations with $H_{m0} > 4$ m (recall Fig. 1), if we select the n largest H_{m0} measurements (for different values of n), we might be interested in the subset of measurements that occur in a given time interval (e.g., 24 h) of another large sea state. For any selected value of n (the largest n measurements), let us refer to the number of measurements that occur near one another within a designated time window, $\pm t_c$, as m. The fraction of these large wave measurements that

[1] NDBC 44007 is located off the coast of Portland, ME: https://www.ndbc.noaa.gov/station_page.php?station=44007.

[2] With measurements recorded every hour and recalling that 1996 was a leap year, there should be $(4 \cdot 365 + 1) \cdot 24 = 35{,}064$ measurements. Due to some short gaps in the dataset, there are actually 34,296 measurements (i.e., there are 768 missing data points—roughly 2%).

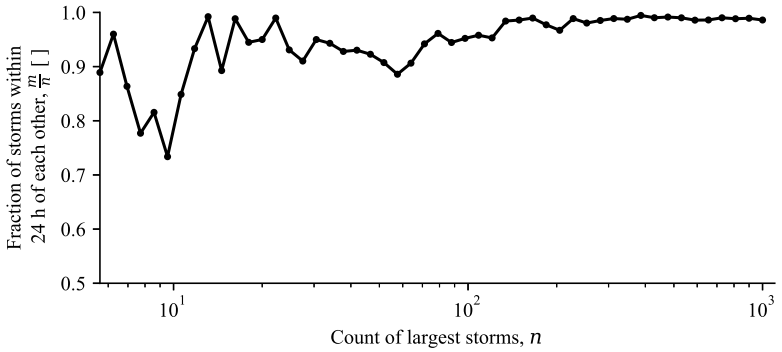


Fig. 3. Fraction of the n largest samples within a cutoff period of $t_c = 24$ h of each other (m/n).

occur close each other is then (m/n). The results of such tallying of data, for various values of n with $t_c = 24$ h, are presented in Fig. 3. From Fig. 3, we note that generally greater than 90% of the large wave measurements occur within ± 24 h of another large wave measurement.

Clearly, hourly sea states exhibit strong serial correlation. This is also true of the most extreme sea states (as shown in Figs. 1 and 3). One common means of dealing with serial correlation is to use a block maxima approach. Mackay and Jonathan (2020) used a block maxima approach to model the distribution of independent storm peaks. As a second step, they resampled time series from the original dataset to capture the complete range of environmental data and calculated environmental contours. The resampled time series were serially correlated such that the approach did not control for serial correlation in the bi-variate distribution used for contour construction. Here, we propose a somewhat similar concept that is also related to the block maxima method, but is different in the way joint block maxima are defined and that we do not resample storm time series here. Note also that our proposed method operates solely on H_{m0} (without considering T_z) – more discussion on this aspect of the proposed approach and why it appears viable for the dataset considered here is provided in Section 2.4.

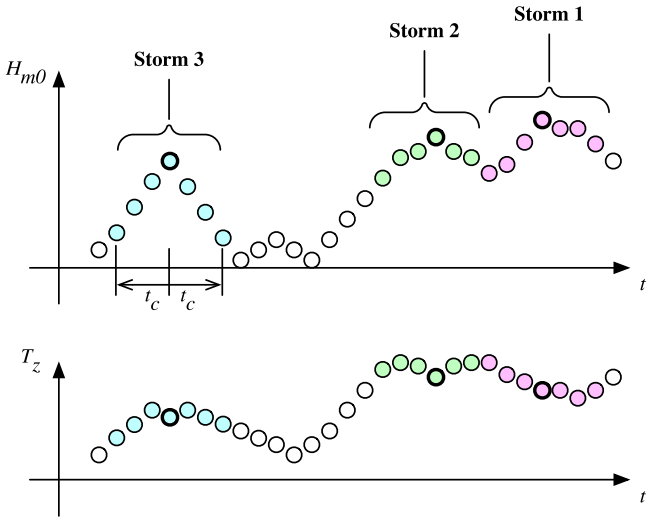


Fig. 4. Illustration of storm identification procedure using the time series of significant wave height (H_{m0}) and zero-upcrossing period (T_z). The members of each storm are shown with filled markers; maxima for each storm are shown with a heavier marker.

2.3. Grouping by storm

With the recognition now that many large wave measurements occur near in time to one another, it is useful to group a series of consecutive measurements into "storms". Accordingly, we refer herein to a storm as a grouping of sea state measurements that take place in some designated short time window. These storms should be chosen in such a way that they are serially independent, while minimizing their size to ensure the largest possible dataset on which to perform subsequent analyses. We shall also, in turn, consider the maxima of such defined storms.

The following procedure is proposed to group the metocean data into storms:

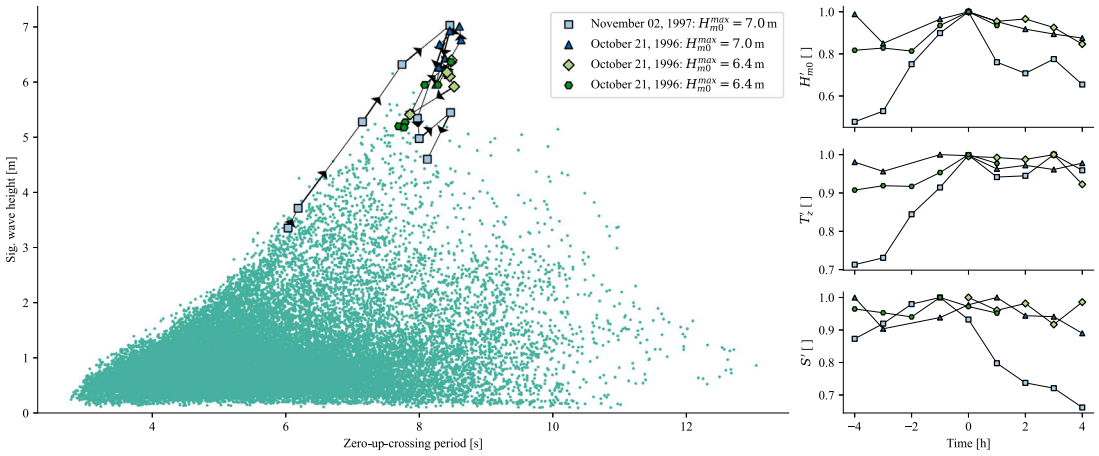
1. **Sort:** Sort the dataset by H_{m0} in descending order.
2. **Identify storms:** Starting with the largest H_{m0} observed, find all other observations that occur within the stated cut-off time ($\pm t_c$) and group all such measurements as (lesser) members of a distinct storm.
3. **Repeat:** Remove the lesser members of each storm from the base dataset. Repeat Step 2 until a desired number of storms (n_s) is obtained.

As the cut-off time is decreased ($t_c \to 0$) and the number of storms is increased ($n_s \to 365.25 \cdot 24$), the storm dataset will approach the original dataset. Note that herein, for better intuitive clarity, we use n_s to refer to the average number of storms per year.

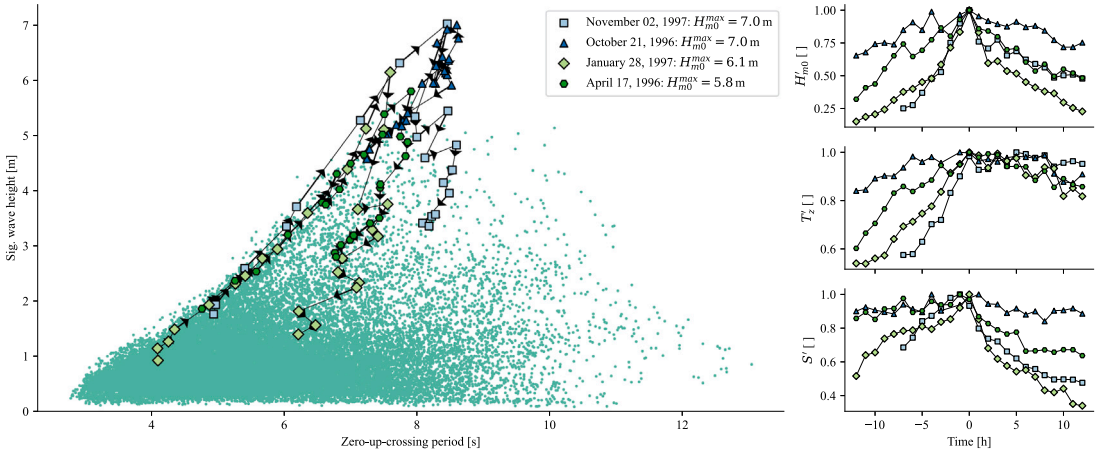
An illustration of the process outlined above is presented in Fig. 4. Here, the procedure illustrates identification of three distinct storms. Note that the maxima of these storms are not necessarily the three largest data points in the overall record but, per the outlined procedure above, they are local maxima that are separated by a period at least as long as t_c. Additionally, the identified storms are not necessarily evenly spaced, as in most block maxima approaches, and can in fact be asymmetric about the storm maxima (see Storm 2 in Fig. 4) due to the fact that storm membership is mutually exclusive. Note that this storm selection procedure operates entirely on the significant wave height and does not consider the zero-up-crossing period. Thus, when considering the two-dimensional $H_{m0} - T_z$ space, the procedure does not necessarily filter for high quantile observations, but instead uses the significant wave height and cut-off time separate the observations into non-overlapping blocks.

The results of this procedure are illustrated with our selected dataset in Fig. 5(a). In this case, we look for the four largest storms where members of a storm must occur within four hours of the local maxima ($t_c = 4$ h). The left-hand portion of Fig. 5(a) shows a scatter diagram of significant wave height and zero-up-crossing period. Arrows connecting storm data points indicate the progression of time. The three plots on the right-hand side of the figure show time series of normalized significant wave height, zero-up-crossing period, and spectral wave steepness.

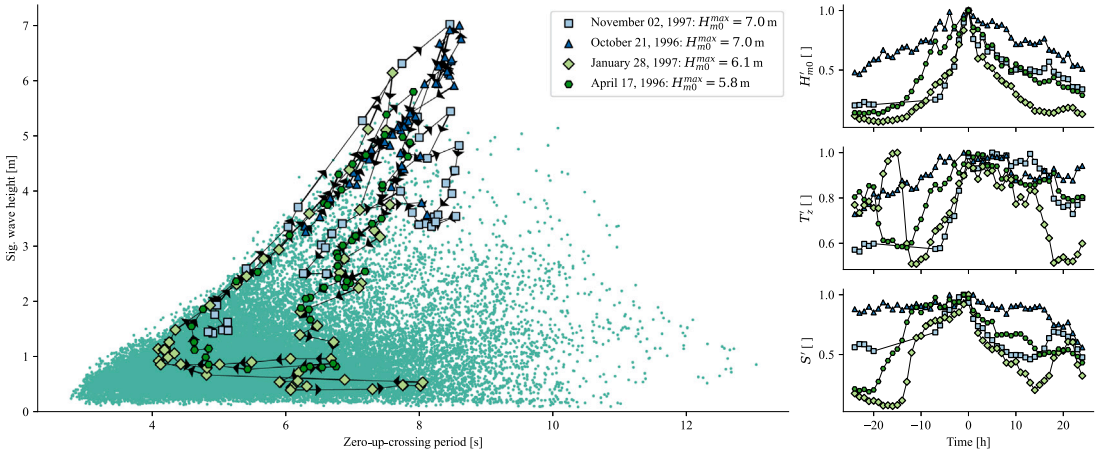
Recognizing that the selection of $t_c = 4$ h is somewhat arbitrary, it is instructive to consider other values for this threshold. Accordingly, Figs. 5(b) and 5(c) show the resulting plots for $t_c = [12, 24]$ h as well. We can see from Fig. 5(a), when $t_c = 4$ h, that the 2nd, 3rd,



(a) $t_c = 4$ h



(b) $t_c = 12$ h



(c) $t_c = 24$ h

Fig. 5. Scatter plots showing the four largest storms within the four-year dataset using different cutoff times ($t_c = [4, 12, 24]\,h$). Smaller plots to right show normalized time histories of significant wave height (H'_{m0}), zero-up-crossing period (T'_z), and spectral wave steepness (S'), centered at the storm maxima.

and 4th largest storms all occurred on the same day (October 21, 1996);[3] we should probably not consider these each to be distinct storms but, instead, one long sustained storm. When t_c is increased to 12 h (Fig. 5(b)), all of the four largest storms are seen to have occurred

[3] When $t_c = 4$ h, the 2nd, 3rd, and 4th largest storms are actually directly adjacent to each other. This is evident from closer examination of time series plots on the right hand side of Fig. 5(a), where we can see that the 2nd, 3rd, and 4th largest storms are asymmetric about their maxima. This is similar to the scenario shown for "Storm 2" in Fig. 4.

on distinct days. Setting $t_c = 24$ h (Fig. 5(c)) results in the same storms as for $t_c = 12$ h, but with many observations of relatively low significant wave heights included as well, perhaps indicating that $t_c = 24$ h is too large, as these four very large storms may begin to incorporate smaller (but independent) storm events.

It is also interesting to note from Fig. 5 how the storms build over time. As is well understood (see, e.g., Ochi, 2005), developing storms can exhibit very steep waves. Following the paths shown (by arrows) between points on the scatter plots in Fig. 5, we can confirm that the storms often grow by ascending along the upper left extents of the data (i.e., close to the wave breaking limit). This behavior appears fairly consistent for this site. The time series on the right side of Fig. 5 also show this same evolutionary progression, with the storm maxima preceded by increasingly steep waves. With $t_c = 24$ h (Fig. 5(c)), we note two cases where the wave period decreases dramatically, likely indicating the arrival of strong local winds followed by growth in wave height.

Beyond the qualitative assessments that can be applied based on Fig. 5, we can also apply more objective statistical tests to assess serial correlation. The Durbin–Watson test uses a statistic defined as (Durbin and Watson, 1950)

$$C = \frac{\sum_{t=2}^{T} \left(x_t - x_{t-1}\right)^2}{\sum_{t=1}^{T} x_t^2}. \tag{2}$$

Values of C close to 2 indicate no evidence of autocorrelation, whereas values close to 0 indicate positive autocorrelation and values close to 4 indicate negative autocorrelation.[4]

Additionally, Blum's statistic (Blum et al., 1961) can be used to test for serial correlation in time series, as was done by Fogle et al. (2008). For any two independent variables, X and Y, each of length N, we can compute the following test statistic:

$$B = \frac{\pi^4}{2N^4} \sum_{j=1}^{N} \left(N_1(j)N_4(j) - N_2(j)N_3(j)\right)^2$$

where

$$\begin{aligned} &N_1(j) \text{ is the number of } (x, y) \text{ pairs where } x \le x_j \text{ and } y \le y_j \\ &N_2(j) \text{ is the number of } (x, y) \text{ pairs where } x > x_j \text{ and } y \le y_j \\ &N_3(j) \text{ is the number of } (x, y) \text{ pairs where } x \le x_j \text{ and } y > y_j \\ &N_4(j) \text{ is the number of } (x, y) \text{ pairs where } x > x_j \text{ and } y > y_j \end{aligned} \tag{3}$$

In considering a single time series, (3) can be used where Y is simply a "lag-1" version of X (i.e., $y_j = x_{j+1}$). At a 1% significance level, the null hypothesis of independence is rejected for values of $B > 4.23$.

Results from each of these independence tests are shown in Fig. 6 for storm maxima with varying values of cutoff period (t_c) and number of storms per year (n_s). (Tabulated results are also presented in Table 1.) The results for the storm maxima datasets are shown along with the original dataset ($t_c = 0$, $n_s = 8579$) and a block maxima method. The block maxima method uses all of the data with block sizes equivalent to those used by the storm grouping (e.g., for $t_c = 12$ h, the storm grouping procedure will include the maxima along with observations within ± 12 h, so the equivalent block maxima approach uses a block of size 25 h).

The results shown in Fig. 6 clearly indicate that the storm maxima datasets exhibit weaker dependence than the original dataset. Fig. 6 also shows that dependence generally decreases as the cut-off time (t_c) increases. This result is consistent with our *ad hoc* conclusions drawn from Fig. 5. Additionally, as expected, dependence also decreases when fewer storms are used ($n_s \to 0$). For the same level of data retention, the storm maxima datasets exhibit lower levels of serial correlations than the block maxima datasets.

Table 1
Data dependence statistics for the original dataset ($t_c = 0$, $n_s = 8,579$) and storm maxima datasets with different values of cutoff time (t_c) and number of sea states per year (n_s). The Durbin–Watson test statistic is C, where values closer to 2 indicate less dependence; the Blum statistic is B, where the null hypothesis of independence is rejected for $B > 4.23$ (at 1% significance level). The fraction of data used is given by f_{max} and f_{tot}, which are the fraction of the original data used in terms of the storm maxima and total storm grouping, respectively.

t_c	n_s	C	B	f_{max}	f_{tot}
0	8579	0.03	4.1e4	–	1
4	60	1.1	17	0.007	0.028
	120	0.91	51	0.014	0.056
	240	0.7	1.6e2	0.028	0.11
12	60	1.8	1.1	0.007	0.084
	120	1.6	14	0.014	0.17
	240	1.3	64	0.028	0.34
24	60	2	1.7	0.007	0.17
	120	2	1.5	0.014	0.34
	240	1.6	32	0.028	0.67

2.4. Environmental contours from storm maxima

These storm maxima datasets, which exhibit lower levels of serial dependence than the original dataset, can be used within a larger engineering analysis workflow. Environmental contours are widely used in ocean engineering to efficiently estimate extreme return levels (Ross et al., 2020). As a example case, we shall employ the storm maxima datasets in producing environmental contours.

Prior to generating contours, we must consider the appropriate exceedance probability level for our storm model data. If we assume that there are n "independent" observations per year and we are interested in extremes associated with a return period of T_r years, then the target exceedance probability level would be

$$\alpha = \frac{1}{n} \cdot \frac{1}{T_r}. \tag{4}$$

If we consider only the five largest storms in the four-year dataset, we have $n = 5/4$, and for a 25-year return period ($T_r = 25$), we would be interested in H_{m0} values associated with a target exceedance probability level of $\alpha = (5/4 \cdot 25)^{-1} = 0.032$ based on a distribution constructed using maxima from only the selected storms. In a similar manner, if we considered a more typical scenario where measurements are taken once each hour (when we assume that "all" data are considered independent), we would have $\alpha = (24 \cdot 365.25 \cdot 25)^{-1} = 4.56 \times 10^{-6}$.

For a cumulative distribution, $F(x)$, the complementary cumulative distribution, which describes the exceedance probability, is $1 - F(x)$. Since we are dealing with two different types of distributions (i.e., the original dataset has measurements every hour whereas the storm maxima datasets have n_s measurements per year), to perform a comparison we can perform a normalization. The daily exceedance probability $(1 - F_D(x))$ can be obtained from the daily cumulative distribution function. For example, if 4-hour storm maxima are considered,[5] the daily distribution function is obtained as:

$$F_D(x) = \left[F_{4\text{h}}(x)\right]^{24/4}. \tag{5}$$

With a clear means of comparing the original dataset (with hourly observations) and the storm maxima datasets (with n_s observations per year), we may begin to examine the distributions of these different

[4] Critical values of the Durbin–Watson statistic at a given significance level are dependent on the number of samples considered (Turner, 2020). In the case of this study, where the number of samples ranges from 34,296 for the original dataset to 240 for the most restrictive of the storm grouping datasets, the upper critical values of the test statistic range from 1.974 to 1.707, respectively.

[5] Note that (5) assumes serial independence; as one means of accounting for serial correlation, (5) can be altered to include sub-asymptotic extremal index (Mackay et al., 2021; Ledford and Tawn, 2003).

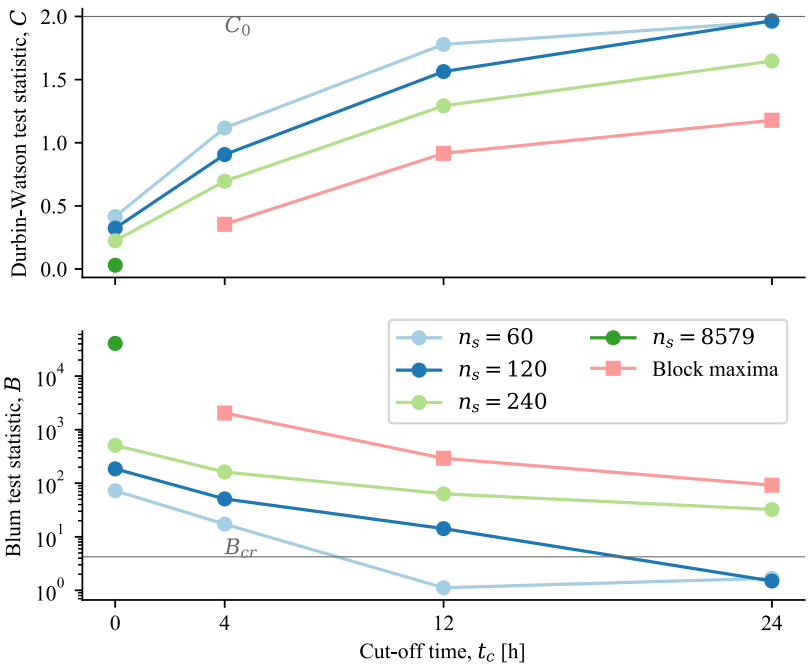


Fig. 6. Durbin–Watson and Blum independence test statistics for varying values of cutoff time (t_c) and number of storms per year (n_s). Original dataset is marked as $t_c = 0$, $n_s = 8579$.

datasets. Empirical (marginal) exceedance probability plots for H_{m0} are shown in Fig. 7. The distribution for the original dataset is shown with those of three storm maxima datasets ($t_c = [4, 12, 24]$ h and $n_s = 240$). Bootstrapping has been performed on the original dataset so as to include a 95% confidence interval in Fig. 7. From this, we can see that the 1-year return level varies between 6.2 and 7 m, with a median of 6.7 m when hourly observations are considered to be independent. It is also interesting to note that many of the nine largest observations in the original dataset (shown with the red bracket in Fig. 7) all occurred within ±12 h of each other, as these observations have been reduced to the single largest observation ($H_{m0} = 7$ m) in the storm maxima datasets for $t_c \geq 12$ h.

When storm maxima are considered, the uncorrected empirical 1-year return value decreases to between 5.6 m ($t_c = 4$ h) and 6.2 m ($t_c = 24$ h). This effect is to be expected and consistent with theory: If serially correlated data such as our original time series are directly used to estimate return values, these return values overestimate the true return value for a given return period (Beirlant et al., 2004, p. 381). The storm maxima are less serially dependent and therefore lead to less overestimation. The effect of serial correlation on extreme values is often characterized with the extremal index $\theta \in [0, 1]$, where independent extremes lead to $\theta = 1$ and, at asymptotic levels, serially dependent extremes lead to $\theta < 1$ (Beirlant et al. 2004, pp. 376). The extremal index characterizes extremal dependence at asymptotic levels, but not at levels that might be of interest in any particular application. At finite return periods, the effect of serial correlation can be quantified using a sub-asymptotic extremal index θ_T which is less or equal than the asymptotic extremal index θ. Mackay et al. (2021) showed that θ_T directly relates to the bias associated with the return period of a value x: $\theta_T = \tilde{T}(x)/T(x)$, where $T(x)$ is the true return period of value x, and $\tilde{T}(x)$ is the return period of the equivalent independent sequence. For our example, that means that the 1-year return value of 6.7 m that we computed based on the original dataset does not have a true return period of 1 year if we are interested in distinct storm events; instead, its average recurrence period of exceedance is longer.

While many approaches are available to construct environmental contours, accounting for different dependence structures among the variables using parametric and non-parametric approaches (Haselsteiner et al., 2021; Manuel et al., 2018), we apply the joint model

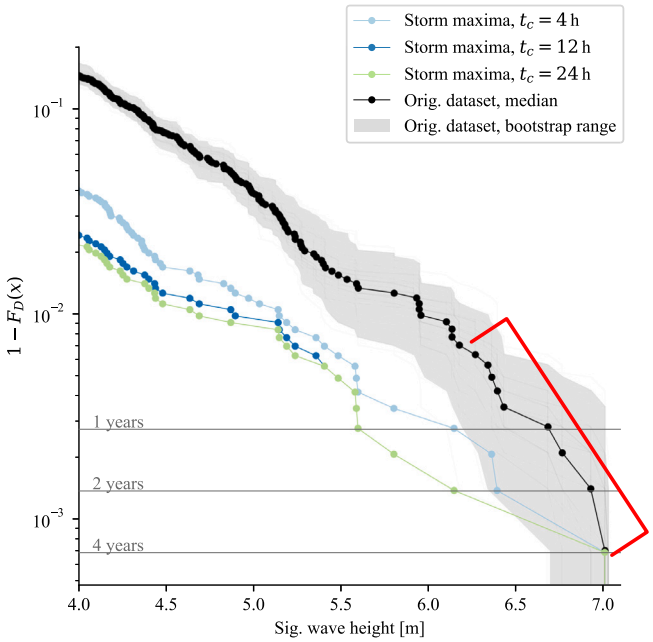


Fig. 7. Daily empirical exceedance probability for significant wave height for original dataset and three different storm maxima datasets with $t_c = [4, 12, 24]$ h and $n_s = 240$. Red bracket shows observations in the original dataset that occurred within ±12 h of each other. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and fitting procedure suggested by Haselsteiner et al. (2020). From the joint distributions, an environmental contour based on the inverse first-order reliability method (IFORM) is constructed (Winterstein et al., 1993). This procedure is implemented in the ViroCon package (Haselsteiner et al., 2019b). The joint model used was originally developed to describe the distribution for all sea states; however, with different parameters, we have found that it provides reasonable fits to the

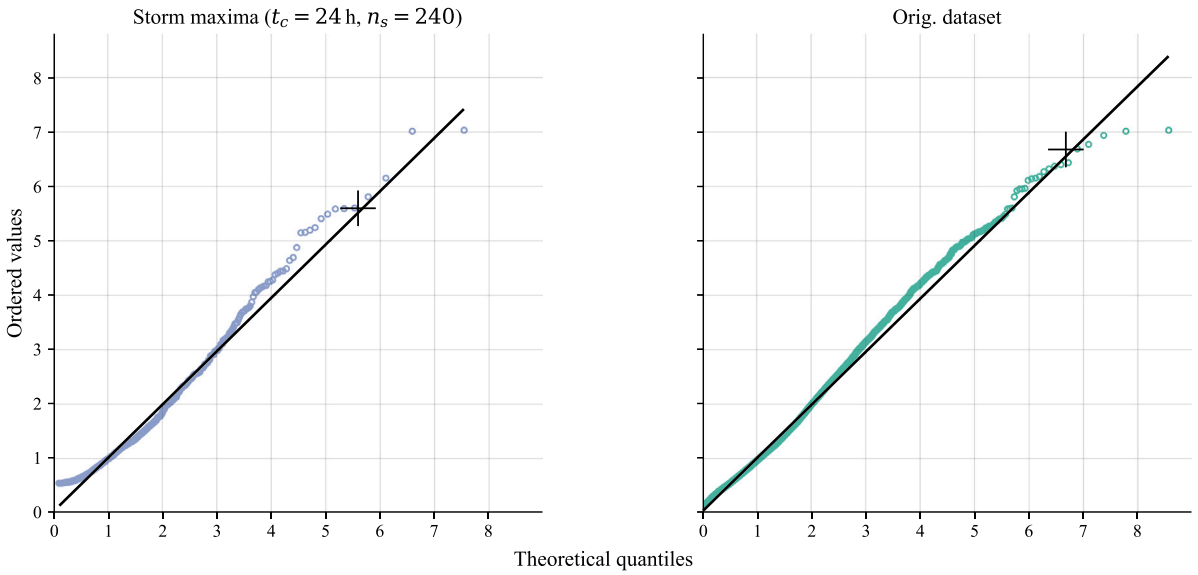


Fig. 8. Significant wave height QQ plots for original dataset and storm maxima dataset with $t_c = 24$ h and $n_s = 240$. Empirical 1-year return level indicated with '+' marker.

distribution for storm maxima as well. See Appendix for a complete description of the model, goodness of fit metrics, and additional details.

The resulting fits for both the storm maxima ($t_c = 12$ h, $n_s = 240$) and original dataset are illustrated in Fig. 8. The empirical 1-year return levels are shown with '+ ' markers. The models for both storm maxima and original datasets show good agreement with the data overall, but the model for storm maxima datasets tracks the high quantile observations more closely than the model for the original dataset.

The goodness of fit for joint two-dimensional model with $t_c = 12$ h and $n_s = 240$ is illustrated in Fig. 9. The model's isodensity curves show good agreement with those from a kernel density estimate performed on the dataset.[6] The agreement is generally better for the extremes than the high probability data, which is expected given the fitting procedures (see Appendix for further details and tabulated goodness of fit metrics).

Based on the fitted distributions, we may construct extreme sea state contours. Nine 1-year contours constructed based on storm maxima are shown in Fig. 10. In each case, the contour constructed based on the storm maxima is shown alongside a contour based on the original dataset. Each row of plots in Fig. 10 uses a certain cut-off time ($t_c = [4, 12, 24]$ h working from top to bottom), while each column corresponds to a certain number of storms used ($n_s = [60, 120, 240]$ working from left to right).

The contours from the storm maxima dataset give smaller significant wave heights for a given zero-up-crossing period than the contour produced using the original dataset. As discussed for the marginal distribution, this is expected and consistent with theory on serial dependence of extremes: Environmental contours, which represent bivariate return values, are overestimated if the underlying distribution is based on serially correlated data.

In most cases, the shape of the storm maxima contours are similar to that of the original dataset. Note that the lower boundaries of the storm maxima contours do not capture some low H_{m0} values in the original dataset. This is expected because the contours describe the extremes of storm maxima instead of the extremes of all sea states. However, this is not a major issue, given that these contours are generally employed in design studies where large waves are the primary concern.

The largest sea state from each contour is reported in Table 2 as H_{m0}^{max} and the corresponding zero-up-crossing period is $T_z(H_{m0}^{max})$. As also indicated by the exceedance plot shown in Fig. 7, as t_c is decreased

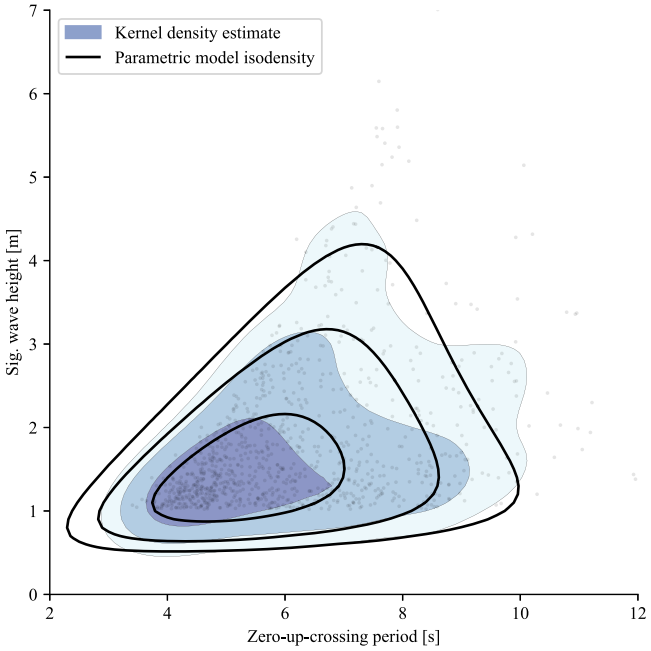


Fig. 9. Joint model goodness of fit for storm maxima dataset with $t_c = 12$ h and $n_s = 240$ illustrated by isodensity curves (contour lines) and kernel density estimate (filled contour) for probability density levels of $\alpha = [0.01, 0.032, 0.1]\,\mathrm{m^{-1}s^{-1}}$. See Appendix for additional joint goodness of fit plots.

and as n_s is increased, the contours from the storm maxima datasets tend to approach the contour from the original dataset (see Fig. 10 and the H_{m0}^{max} and $T_z(H_{m0}^{max})$ columns in Table 2).

From Fig. 10 and Table 2, we can observe that the contours generated the storm maxima datasets are fairly similar. However, subtle differences do exist, particularly when T_z is large, and it is well-known that sea states on the contour other than that with the largest significant wave height can cause the largest loading for certain types of responses (see, e.g., Muliawan et al., 2013; Edwards and Coe, 2018; de Hauteclocque et al., 2022). Based on the results presented in Section 2.2, adding more sea states to the storm maxima dataset by using higher values for n_s will tend to increase serial dependence. Additionally, looking at the contours in Fig. 10 from left to right

[6] The bandwidth for Gaussian kernel was chosen based on method proposed by Scott (2015).

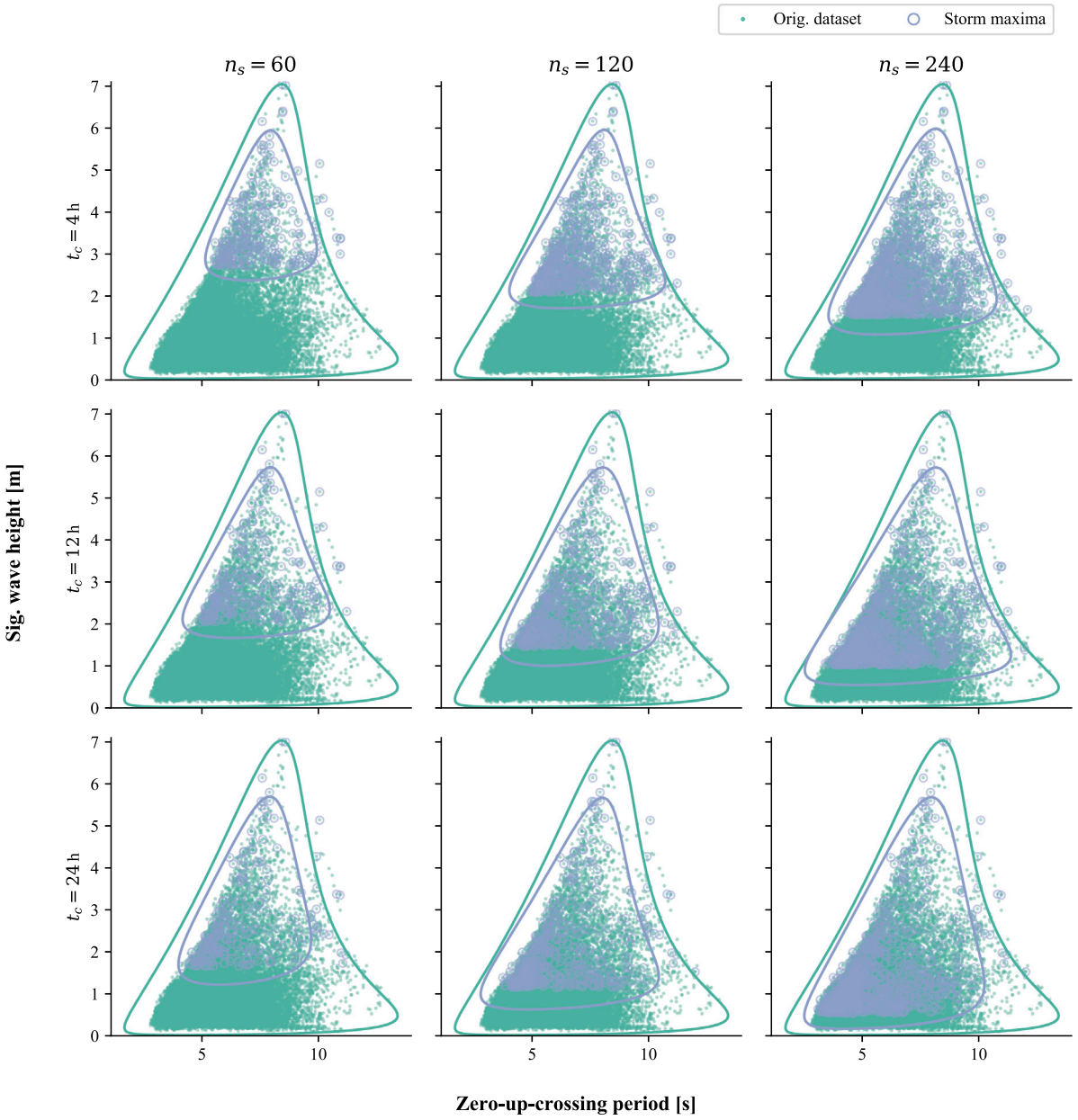


Fig. 10. One-year return contours based on original dataset and storm maxima with n_s = [60, 120, 240] and t_c = [4, 12, 24] h. As n_s increases and t_c decreases, the storm maxima environmental contours' peaks H_{m0} increase, approaching that of the original dataset.

(increasing n_s), we see that higher values of n_s mostly tend to increase the number of low wave height observations that are included in the storm maxima dataset. Interestingly, the storm peak datasets also hold events with extreme zero-up-crossing period, although, as discussed in Section 2.3 and illustrated in Fig. 4, the storm grouping procedure used here operates only based significant wave height. The presence of these extreme zero-up-crossing period observations suggest that more complex schemes based on storm resampling or based on block maxima where variables are taken from different times within the block may not necessarily be required to obtain a dataset that can be used for a sea state environmental contour. In practice, one might use the results shown in Fig. 6 to balance the opposing demands of limiting serial dependence and using a sufficiently large fraction of the dataset to enable robust analyses.

3. Discussion and conclusion

This paper presents a relatively simple method for reducing serial correlation in metocean datasets for further use in extreme condition predictions. The method is based on grouping the dataset into distinct storms, and using the maxima from each storm to form a new dataset with less serial correlation. The initial results presented in this paper show that the method does reduce dependence and gives extreme value results that are generally in-line with those when using a dataset with hourly measurements, but with lower significant wave heights predicted at most wave periods. We also showed that joint models that were developed for all sea states data can be applied to storm maxima data.

As mentioned in Section 2.1, we chose to perform the analysis using four years of nearly contiguous data. Given the nature of the analysis

Table 2
Distribution fit parameters and mean absolute error (MAE) for $H_{m0} > 4$ m. The largest sea state on the 1-year contour has a significant wave height of H_{m0}^{max} and a zero-up-crossing period $T_z(H_{m0}^{max})$.

t_c	n_s	α	β	δ	c_0	c_1	c_2	c_3	MAE	H_{m0}^{max}	$T_z(H_{m0}^{max})$
0	8579	0.34	0.78	4.5	3.8	5.5	1.8e−20	0.34	0.18	7	8.4
4	60	0.07	0.58	9.7e+03	5.2	3.5	1.6e−12	2.2	0.099	5.9	7.9
	120	0.11	0.61	1.2e+03	4.2	5	1.9e−14	85	0.13	5.9	8.1
	240	0.16	0.65	1.8e+02	3.9	5.4	2.2e−21	0.47	0.17	6	8.1
12	60	0.24	0.71	2.1e+02	4.4	4.6	5.3e−15	1.6e+02	0.18	5.7	8
	120	0.23	0.69	71	4	5.3	5.6e−19	0.43	0.24	5.7	8
	240	0.18	0.64	39	3.3	6.4	2.3e−21	0.53	0.3	5.7	8.1
24	60	0.53	0.85	29	3.9	5.3	1.8e−23	0.57	0.2	5.7	7.9
	120	0.38	0.76	18	3.2	6.3	2.4e−16	0.63	0.29	5.7	8
	240	0.44	0.77	5.9	3.4	5.9	1.3e−21	0.33	0.32	5.7	7.9

and the concern with serial correlation, this dataset was considered preferable over others with larger proportions of missing data. To allow for more direct benchmarking with the raw dataset, we similarly considered analyses with return periods of one to four years. While these time periods are generally shorter than those considered for offshore engineering projects, the underlying theory of the proposed method is agnostic to the total period of record or the return period. With the short dataset employed here, the 1-year return period contours are more appropriate to study the role of serial correlation; had we considered longer return period contours, the added complexity arising from extrapolation leading to greater uncertainty in the tails would mask the issue of including or excluding serial correlation in storms. Nevertheless, future studies should apply this storm grouping procedure to more scenarios more typical in practice (e.g., using a dataset of 15 years to predict extremes for a return period of 100 years).

As only a single location and dataset were considered in this study, future studies should consider the application and performance of this storm grouping approach with a wider range of locations. Additionally, while we have considered a two-dimensional contour for the significant wave height and zero-up-crossing period, it would also be useful and interesting to consider other design variables, such as wave direction, current, and wind. Furthermore, there are many methods for producing contours from a dataset, which produce dramatically different results—some exploration of how these different contours pair with the storm grouping approach should be considered.

CRediT authorship contribution statement

Ryan G. Coe: Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing. **Lance Manuel:** Conceptualization, Writing – review & editing. **Andreas F. Haselsteiner:** Conceptualization, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

RGC would like to acknowledge funding support from the US Department of Energy's Water Power Technologies Office. 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.

Appendix. Probabilistic models & fitting

As suggested by Haselsteiner et al. (2020), we employ the following probabilistic model. An exponentiated Weibull distribution is fit to H_{m0} and a log-normal distribution is fit to T_z conditional on H_{m0}.

$$F(h_{m0}) = \left(1 - \exp\left(-\frac{h_{m0}}{\alpha}\right)^{\beta}\right)^{\delta} \tag{6a}$$

$$F(t_z|h_{m0}) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{\ln(t_z) - \mu_{t_z}}{\sqrt{2}\sigma_{t_z}}\right)\right) \tag{6b}$$

Here, α is the scale parameter while β and δ are shape parameters. The log-normal distribution parameter, μ_{t_z}, is modeled by

$$\mu_{t_z} = \ln\left(c_1 + c_2\sqrt{\frac{h_{m0}}{g}}\right), \tag{7}$$

where g is the acceleration due to gravity. The parameters c_1 and c_2 determined based on fitting. The parameter σ_{t_z} is modeled by

$$\sigma_{t_z} = c_3 + \frac{c_4}{1 + c_5 h_{m0}}. \tag{8}$$

The parameters of the joint models were estimated in multiple steps. First the marginal distribution of H_s was estimated by minimizing the sum of the squared errors, where errors were weighted by H_{m0}^2 so as to give more influence to the larger waves (Haselsteiner and Thoben, 2020). Next the T_z data was sorted into intervals based on H_{m0} bins of size 1 m. Using maximum likelihood estimation, distributions were fit to each of these subsets of T_z containing more than ten observations. Finally, the dependence functions (7)–(8) were fit by minimizing the least squares error. All estimated parameters and the model's errors for $H_{m0} > 4$ m are shown in Table 2.

The goodness of fit for the storm maxima joint models is also illustrated by Fig. 11, which shows isodensity curves for the joint models over top of the storm maxima datasets and isodensity curves based on a kernel density estimate, where the bandwidth was chosen based on method proposed by Scott (2015). The fits look reasonable (i.e., the parametric models isodensity curves follow those from the kernel density estimate), especially considering the largest and longest waves. The fits do not capture the wave breaking limit particularly well, nor do they capture the distribution for the smallest waves in the storm maxima datasets, but this second issue is of little concern given the application of these contours in predicting large extreme sea states.

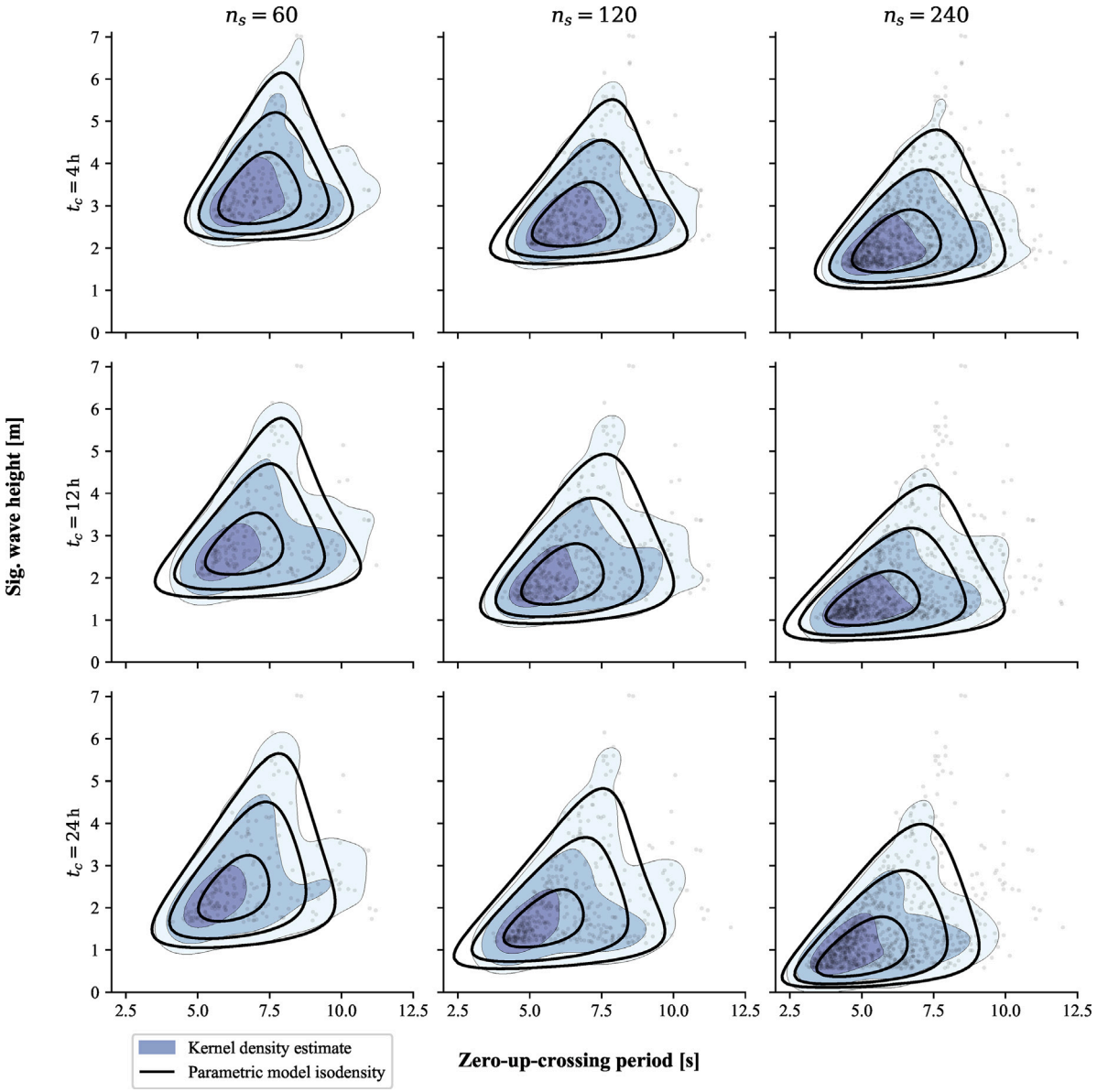


Fig. 11. Joint model goodness of fit for storm maxima datasets illustrated by isodensity curves (contour lines) and kernel density estimate (filled contour) for probability density levels of $\alpha = [0.01, 0.032, 0.1]\,\mathrm{m^{-1}s^{-1}}$.

References

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., 2004. Statistics of extremes: Theory and applications.

Blum, J.R., Kiefer, J., Rosenblatt, M., 1961. Distribution free tests of independence based on the sample distribution function. Ann. Math. Stat. 32 (2), 485–498, URL http://www.jstor.org/stable/2237758.

Brockwell, P.J., Davis, R.A., 2009. Time Series: Theory and Methods. Springer science & business media.

Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values. Springer, London; New York.

de Hauteclocque, G., Mackay, E., Vanem, E., 2022. Quantitative comparison of environmental contour approaches. Ocean Eng. 245, 110374. http://dx.doi.org/10.1016/j.oceaneng.2021.110374, URL https://www.sciencedirect.com/science/article/pii/S0029801821016693.

Derbanne, Q., de Hauteclocque, G., 2019. A new approach for environmental contour and multivariate de-clustering. In: Proceedings of the ASME 2019 38th International Conference on Ocean, Offshore and Arctic Engineering, Vol. 3: Structures, Safety, and Reliability, no. V003T02A030. OMAE2019, American Society of Mechanical Engineers (ASME), Glasgow, Scotland, UK, http://dx.doi.org/10.1115/OMAE2019-95993.

Durbin, J., Watson, G.S., 1950. Testing for serial correlation in least squares regression: I. Biometrika 37 (3/4), 409–428, URL http://www.jstor.org/stable/2332391.

Edwards, S.J., Coe, R.G., 2018. The Effect of Environmental Contour Selection on Expected Wave Energy Converter Response. Journal of Offshore Mechanics and Arctic Engineering 141 (1), http://dx.doi.org/10.1115/1.4040834.

Fawcett, L., Walshaw, D., 2012. Estimating return levels from serially dependent extremes. Environmetrics 23 (3), 272–283. http://dx.doi.org/10.1002/env.2133, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/env.2133.

Ferro, C.A.T., Segers, J., 2003. Inference for clusters of extreme values. J. R. Stat. Soc. Ser. B Stat. Methodol. 65 (2), 545–556. http://dx.doi.org/10.1111/1467-9868.00401, URL https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9868.00401.

Fogle, J., Agarwal, P., Manuel, L., 2008. Towards an improved understanding of statistical extrapolation for wind turbine extreme loads. Wind Energy 11 (6), 613–635. http://dx.doi.org/10.1002/we.303, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/we.303.

Haselsteiner, A.F., Coe, R.G., Manuel, L., Chai, W., Leira, B., Clarindo, G., Guedes Soares, C., Hannesdóttir, Á., Dimitrov, N., Sander, A., Ohlendorf, J.-H., Thoben, K.-D., de Hauteclocque, G., Mackay, E., Jonathan, P., Qiao, C., Myers, A., Rode, A., Hildebrandt, A., Schmidt, B., Vanem, E., Huseby, A.B., 2021. A benchmarking exercise for environmental contours. Ocean Eng. 236, 109504. http://dx.doi.org/10.1016/j.oceaneng.2021.109504, URL https://www.sciencedirect.com/science/article/pii/S0029801821009033.

Haselsteiner, A.F., Coe, R.G., Manuel, L., Nguyen, P.T.T., Martin, N., Eckert-Gallup, A., 2019a. A benchmarking exercise on estimating extreme environmental conditions:

Methodology & baseline results. In: Proceedings of the ASME 2019 38th International Conference on Ocean, Offshore and Arctic Engineering, Vol. 3: Structures, Safety, and Reliability, no. V003T02A049. OMAE 2019, American Society of Mechanical Engineers (ASME), Glasgow, Scotland, UK, http://dx.doi.org/10.1115/OMAE2019-96523.

Haselsteiner, A.F., Lehmkuhl, J., Pape, T., Windmeier, K.-L., Thoben, K.-D., 2019b. ViroCon: A software to compute multivariate extremes using the environmental contour method. SoftwareX 9, 95–101. http://dx.doi.org/10.1016/j.softx.2019.01.003.

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: Proceedings of the ASME 2020 39th International Conference on Ocean, Offshore and Arctic Engineering, Vol. 2A: Structures, Safety, and Reliability, no. V02AT02A047. OMAE2020, American Society of Mechanical Engineers (ASME), Virtual, Online, http://dx.doi.org/10.1115/OMAE2020-18668.

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

IEC, 2019. Wind Turbines—Part 1: Design Requirements. Technical Report, (IEC 61400-1:2019), International Electrotechnical Commission, URL https://webstore.iec.ch/publication/26423.

Jonathan, P., Ewans, K., 2013. Statistical modelling of extreme ocean environments for marine design: A review. Ocean Eng. 62, 91–109. http://dx.doi.org/10.1016/j.oceaneng.2013.01.004, URL https://www.sciencedirect.com/science/article/pii/S002980181300019X.

Ledford, A.W., Tawn, J.A., 2003. Diagnostics for dependence within time series extremes. J. R. Stat. Soc. Ser. B Stat. Methodol. 65 (2), 521–543. http://dx.doi.org/10.1111/1467-9868.00400, arXiv:https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/1467-9868.00400.

Leford, A.W., Tawn, J.A., 1996. Statistics for near independence in multivariate extreme values. Biometrika 83 (1), 169–187. http://dx.doi.org/10.1093/biomet/83.1.169.

Mackay, E., de Hauteclocque, G., Vanem, E., Jonathan, P., 2021. The effect of serial correlation in environmental conditions on estimates of extreme events. Ocean Eng. 242, 110092. http://dx.doi.org/10.1016/j.oceaneng.2021.110092, URL https://www.sciencedirect.com/science/article/pii/S0029801821014189.

Mackay, E.B., Jonathan, P., 2020. Estimation of environmental contours using a block resampling method. In: Proceedings of the ASME 2020 39th International Conference on Ocean, Offshore and Arctic Engineering, Vol. 2A: Structures, Safety, and Reliability, Vol. V02AT02A045. OMAE 2020, American Society of Mechanical Engineers (ASME), Virtual, Online, http://dx.doi.org/10.1115/OMAE2020-18308.

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. J. Ocean Eng. Mar. Energy 4 (4), 293–310. http://dx.doi.org/10.1007/s40722-018-0123-0.

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. J. Offshore Mech. Arct. Eng. 135, 031301–1 to 031301–10. http://dx.doi.org/10.1115/1.4024267.

Myrhaug, D., 2018. Some probabilistic properties of deep water wave steepness. Oceanologia 60 (2), 187–192. http://dx.doi.org/10.1016/j.oceano.2017.10.003, URL https://www.sciencedirect.com/science/article/pii/S0078323417300970.

Ochi, M.K., 2005. Ocean Waves: The Stochastic Approach, Vol. 6. Cambridge University Press.

Ross, E., Astrup, O.C., Bitner-Gregersen, E., Bunn, N., Feld, G., Gouldby, B., Huseby, A., Liu, Y., Randell, D., Vanem, E., Jonathan, P., 2020. On environmental contours for marine and coastal design. Ocean Eng. 195, 106194. http://dx.doi.org/10.1016/j.oceaneng.2019.106194, URL https://www.sciencedirect.com/science/article/pii/S0029801819303798.

Scott, D., 2015. Multivariate Density Estimation: Theory, Practice, and Visualization. In: Wiley Series in Probability and Statistics, Wiley, URL https://books.google.com/books?id=TZg3BwAAQBAJ.

Turner, P., 2020. Critical values for the Durbin-Watson test in large samples. Appl. Econ. Lett. 27 (18), 1495–1499. http://dx.doi.org/10.1080/13504851.2019.1691711.

Vanem, E., 2018. A simple approach to account for seasonality in the description of extreme ocean environments. Mar. Syst. Ocean Technol. 13 (2), 63–73. http://dx.doi.org/10.1007/s40868-018-0046-6.

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 6^{Th} International Conference on Structural Safety & Reliability. ICOSSAR, Innsbruck, Austria.