Can rogue waves be predicted using characteristic wave parameters?

A.D.Cattrell[1], M.Srokosz[2], B. I. Moat[2], R.Marsh[3]

Affiliation:

1) Fluid Structure Interactions, Engineering and the Environment, Boldrewood Innovation Campus, University of Southampton, SO17 1BJ, UK

2) National Oceanography Centre, University of Southampton Waterfront Campus, European Way, Southampton, SO14 3ZH, UK.

3) Ocean and Earth Science, University of Southampton, National Oceanography Centre Southampton, SO14 3ZH, UK.

Author mailing address: A. Cattrell, 176/3001, Boldrewood Innovation Campus, University of Southampton, SO17 1BJ, United Kingdom.

Email: A.Cattrell@Southampton.ac.uk

Key points:

- Largest dataset of oceanic rogue waves is obtained from wave buoys.
- Rogue wave occurrence displays no clear link with short-term wave statistics.
- Potential predictability of rogue wave occurrence from long-term wave statistics.

Abstract:

Rogue waves are ocean surface waves larger than the surrounding sea that can pose a danger to ships and offshore structures. They are often deemed unpredictable without complex measurement of the wave field and computationally intensive calculation which is infeasible in most applications, consequently there a need for fast predictors.

Here we collate, quality control, and analyse the largest dataset of single-point field measurements from surface following wave buoys to search for predictors of rogue wave occurrence. We find that analysis of the sea state parameters in bulk yields no predictors, as the subset of seas containing rogue waves sits within the set of seas without. However, spectral bandwidth parameters of rogue seas display different probability distributions to normal seas, but these parameters are rarely provided in wave forecasts. When location is accounted for, trends can be identified in the occurrence of rogue waves as a function of the average seas state characteristics at that location. These trends follow a power law relationship with the characteristic sea state parameters: mean significant wave height and mean zero up-crossing wave period. We find that frequency of occurrence of rogue waves and their generating mechanism is not spatially uniform, and each location is likely to have its own unique sensitivities which increase in the coastal seas. We conclude that forecastable predictors of rogue wave occurrence will need to be location specific and reflective of their generation mechanism. Therefore, given location and a sufficiently long historical record of sea state characteristics, the likelihood of occurrence can be obtained for mariners and offshore operators.

1. Introduction:

Rogue waves are transient surface gravity waves of height much greater than expected for the surrounding sea, and can severely damage ships and offshore structures (Dysthe et al., 2008). The most common method of categorising a rogue wave from a normal sea is to use a wave or crest height that exceeds a threshold in relation to the significant wave height (Haver, 2000):

$$\frac{H_{max}}{H_s} > 2 \quad \text{Eq. 1}$$

$$\text{and/or } \frac{C_{max}}{H_s} > 1.25 \quad \text{Eq. 2}$$

where H_{max} is the zero-crossing wave height, C_{max} is the crest height, and H_s is the significant wave height, here estimated as four times the standard deviation of the sea surface elevation from a 20-minute observation period. Therefore, rogue waves are not always extreme waves, just larger than statistically expected.

There are several competing theories for the physical mechanism explaining the formation of oceanic rogue waves (Forristall, 2005). First, wave energy concentration through spatio-temporal wave focusing due to the dispersive nature of water waves in intermediate and deep water (Draper, 1966; Kharif et al., 2009; Slunyaev et al., 2005), which is further enhanced by nonlinearities (Longuet-Higgins, 1963; Tayfun, 1980, 2008). Second, modulational instability or Benjamin−Feir instability, the generation of spectral-sidebands and eventual breakup of the waveform into pulses through nonlinearity (Benjamin & Feir, 1967). Taking inspiration from rogue waves in aforementioned non-oceanic media, these nonlinear interactions have been suggested as a cause of oceanic rogue waves (Kharif & Pelinovsky, 2003a). Breather solitons (Akhmediev et al., 1987) and the Peregrine soliton (Peregrine, 1983), which “appears from nowhere and disappears without a trace” (Akhmediev et al., 2009), have also been suggested as causes (Kibler et al., 2010) and have been demonstrated experimentally in a one-dimensional water channel (Chabchoub et al., 2012) and in very shallow water wind waves (Costa et al., 2014). The real ocean is rarely unidirectional, and the importance of the instability is questioned with recent studies explaining rogue wave formation without the aid of modulational instability (Birkholz et al., 2016; Fedele et al., 2016). Other theories suggest the importance of local physical forcing, such as the presence of ocean currents or the bottom topography in shallow waters focusing energy (T. T. Janssen & Herbers, 2009).

Wave prediction using a deterministic approach typically uses radar images of the sea surface at given locations in space and time, combined with the physical laws, to predict the future sea surface elevation (Dannenberg et al., 2010). The process is heavily dependent on signal processing theory and is computationally expensive (Blondel-Couprie & Naaijen, 2012); it is therefore generally only used operationally to predict that the wave heights will remain below a threshold (Belmont et al., 2014).

Precursor analysis is the identification of characteristic behaviours prior to extreme events (Hallerberg et al., 2008). For rogue waves, the detection of instabilities in their infancy before they develop can act as a predictor of rogue wave occurrence, thus alleviates the need to solve the governing equations. This was demonstrated in a computational approach, unproven in the real-ocean, by *Cousins & Sapsis* (2016), who analysed the interplay between nonlinear wave mechanisms that define which wave groups will focus due to modulation instabilities, and the power spectrum which defines wave group formation due to random phase difference between harmonics. They defined a critical length scale over which, the locally concentrated energy acts as a trigger of nonlinear focussing, thus deriving short-term precursors of rare events. This method still requires accurate sensing of the wave field, whereas attributing rogue wave occurrence to sea state parameters that form part of a traditional wave forecasts could yield a computationally cheap method of predicting rogue wave likelihood, that is most useful to mariners and offshore operators.

Large datasets of oceanic rogue waves, as compiled here, can be used to assess these theories of formation, and facilitate the investigation of predictability. A Baylor wave staff mounted on the Meetpost Noordwijk platform in 18-m average water depth, recorded 5000 waves in the southern North Sea in January 1998 (Tayfun, 2008). The largest waves were attributed to the constructive focusing of spectral components enhanced by second-order bound modes. Supporting this, *Christou & Ewans*, (2014) analysed 122 million wave profiles collected from fixed offshore platforms at 22 locations in North Sea, 5 in Gulf of Mexico, 5 in South China Sea, and one on the North-West shelf of Australia. The dataset contained 3649 rogue waves, the occurrence of which was found to be not governed by sea state parameters, but rare events of the normal population caused by dispersive focusing.

Offshore of California and Oregon, wave profiles from 16 Datawell Directional Waverider buoys form a dataset with approximately 1 million waves (Baschek & Imai, 2011). Of these, 2843 exceeded H > 2.0 Hs and 258 exceeded H > 2.2 Hs. The buoy locations were categorised, into deep water, representative of the open ocean; shallow water; and coastal ocean, of variable depth sheltered by islands. There are spatial differences across the region, showing that rogue wave occurrence per annum is less frequent in the shallow and the sheltered locations than in the open ocean. To estimate the likelihood of encounter on a global scale, the probability of encountering a freak wave at the five open ocean buoys was applied to global wave heights, empirically derived from 25-km resolution QuikSCAT wind speed data, yielding a world map of the extrapolated likelihood of encountering rogue waves in the open ocean within a 24 hour period (Baschek & Imai, 2011). We include and extend the data from these buoys in our study, to compile the largest dataset to date for the study of rogue waves.

Analysis of vertical displacement time series data from surface following wave buoys allows the study of waves away from the influences of offshore structures. The dataset used in this study is an order of magnitude larger than previous studies, which is important when analysing rare events. The dataset offers a unique spatial insight into the cause of formation of rogue waves in a range of wave environments covering multiple ocean basins. Analysis of the time series data allows for the assessment of sea state characteristics as a predictor of rogue waves and to study the shape of rogue waves.

This paper is organised as follows. First, we detail the measurement and the quality control of the dataset of observed rogue waves. Second, the potential causal links between rogue waves and sea state parameters is investigated. Third, we examine the average shape of rogue waves for a range of size criteria. Fourth, the spatial distribution of rogue waves is mapped. We conclude by discussing the implications of our analysis in the context of previous rogue wave studies.

2. Dataset:

The data analysed here consists of vertical displacement recorded by 80 Datawell waverider buoys around the coast of North America and Pacific Ocean islands, and covers diverse wave environments, from fetch-limited coastal bays to the deep ocean away from coastal processes. The earliest record began in August 1993, and the most recent data from active buoys cut-off at February 2017, with buoy record lengths varying. In contrast to many previous wave buoy studies, the buoys are continuously measuring, not just switched on during storms.

The wave buoys are managed by, and the data freely available from, the Coastal Data Information Program (CDIP), operated by Scripps Institution of Oceanography. Datawell waverider buoys use accelerometers to measure waves with periods of 1.6–30 s and wave heights up to 40 m with a vertical resolution of 0.01 m. The vertical displacement of the buoy is sampled at a rate of 3.84 Hz; however, data are transmitted and logged on-board with a sampling frequency of 1.28 Hz. Here we use data from the buoy's memory card data to avoid transmission losses.

Wave buoys can underestimate the wave peaks by avoiding the 3-D peak of the wave (Allender et al., 1989) or by being dragged through the crest, avoiding short-crested extreme waves (Seymour & Castel, 1998). In addition, the fluid structure interactions of a wave buoy can linearise the wave time series (James, 1986; Magnusson et al., 1999). Wave buoys are also subject to biofouling (Thomson et al., 2015), vandalism (Beets et al., 2015), and affected by tidal currents. These drawbacks in sampling using wave buoys are mitigated by the unparalleled spatial distribution, length of record, and consistency of continuous surface elevation measurement by the Datawell Waverider buoys (Casas-Prat & Holthuijsen, 2010).

3. Quality control (QC) and initial processing of the dataset:

Field measurements of waves are subject to errors that must be removed to obtain a high quality and reliable dataset. Therefore, a strict QC procedure is required. Furthermore, since this study is looking at extreme individual wave events, not just sea state statistics where the occasional spike would be smoothed in the large sample, a stringent QC procedure for data failing flags was applied.

Each displacement time series was split into non-overlapping 20-minute seas, the typical observational period. The buoy automatically flags questionable, bad, or missing data points in the same time domain as the vertical displacement, and CDIP also runs a shore-side QC process. Any 20-minute sea with an error flag was removed, as sufficient quantity of data allowed this rather than attempting to fix observations by removing single erroneous data points (Makri et al., 2016). For each sea, the vertical displacement time series was linearly interpolated to increase the time resolution by a factor of 10, and the zero up-crossing wave period, wave height, and crest height were calculated.

Screening of erroneous values not identified by the buoy or CDIP's QC took place using a series of filters. The entire 20-minute sea was removed if it had values in excess of the buoy's displacement limits or failed any of the following flags based on the QC process undertaken by *Christou and Ewans,* (2014):

Flag a) Individual waves with a zero-crossing wave period >25 seconds.

Flag b) The rate of change of surface elevation, *Sy*, exceeded by a factor of two:

$$Sy = \left(\frac{2\pi\sigma}{Tz}\right)\sqrt{(2\ln N_z)} \qquad \text{Eq. 3}$$

where σ is the is the standard deviation of the surface elevation η, N_Z is the number of zero up-crossing periods (Tz).

Flag c) Flag b, running from time maxima to minima.

Flag d) 10 consecutive data points of the same value.

Flag e) Absolute crest or trough elevation is greater than 5 times the standard deviation of the 20-minute water surface elevation.

Flag f) A single zero-crossing containing >1499 data points.

Seas where then categorised as normal or rogue using Eq. 1 and Eq. 2. Seas not containing rogue waves are hereafter referred to as normal seas. Rogue waves were then subject to a visual QC as performed by Christou & Ewans, (2014) and Makri et al., (2016) to ensure an erroneous wave was not included in the analysis. Although subjective, experience gained reviewing rogue waves and previous literature allowed sound identification of instrument error.

4. Results:

From an initial dataset size equivalent to 13.2 million 20-minute seas, 11.4 million seas (86%) passed QC. These seas contain 1.1 billion individual wave profiles; of these, 74,262 were rogue waves with Abnormality Index (h/Hs; AI) of 2<AI<3, 120 with 3<AI<4, 30 with 4<AI<5, and 19 with AI>5 (Figure 2a). 21,682 had a C_{max}/Hs ratio exceeding 1.25, 324 exceeding 1.75, 137 exceeding 2.25, and 67 exceeding 2.75 (Figure 2b). The dataset covers extensive range of significant wave heights up to 14 m, peak wave heights exceeding 20 m, and crest elevations up to 14 m.

4a. Sea state parameters:

Assessing the occurrence of rogue waves as a function of the statistics of the sea state in which they occur could indicate the method of their generation. Furthermore, a link between forecastable wave parameters and rogue wave occurrence could facilitate a low computational-cost predictor of rogue wave events.

Wave steepness has been cited as an explanation for rogue wave formation because, under certain conditions, nonlinear interactions beyond second order can provide significant increases in wave elevation and steepness (Gibson & Swan, 2007). Plotting the common wave parameters significant wave height (Hs) and peak wave period (T*p*; Figure 3), with each point representing a 20-minute sea that passed the QC procedure, gives an indication of steepness. The seas containing rogue waves primarily lie within the distribution of normal seas, and normal seas are as steep as or steeper than rogue seas therefore, steepness cannot be the exclusive causal factor in rogue event formation. The marginal PDF of Hs indicates that the majority of rogue waves occur in seas with low significant wave height, and that there is no discernible link between Hs and rogue wave occurrence when bulk analysing the dataset as a many independent seas. The marginal PDF of T*p* shows a bimodal distribution for both rogue sea and normal seas, with peaks at 8 s and 14 s. Rogue seas display increased probability, relative to normal seas, in seas with T*p* <6 s. We discuss the distribution of period further below.

Another assessment of the role of steepness is the analysis of maximum crest height in the 20-minute sea as a function of the mean sea state steepness S_1 (Figure 4b):

$$S_1 = \frac{2\pi}{g}\frac{H_s}{T_1^2} \quad \text{Eq. 4}$$

where g is gravitational acceleration, and mean wave period $T_1 = m_0/m_1$ calculated from the first two moments of the wave spectrum:

$$\mathrm{m_n} = \int_0^\infty \mathrm{f^n\,S(f)}\ \partial\mathrm{f} \quad \text{Eq. 5}$$

where $S(f)$ is the non-directional energy density spectrum, with $H_s = 4\sqrt{m_0}$.

As previously seen, the rogue seas mostly sit within the normal seas, and there are normal seas with greater steepness than rogue seas and the marginal PDF of S_1 shows little deviation between the rogue and normal seas (Figure 4e). Furthermore, the distributions of maximum values for rogue seas and normal seas do not form separate distributions (Figure 2).

The relative importance of nonlinearities can be measured by looking at the maximum crest height as a function of wave skewness λ_3 (Figure 4c) and the excess kurtosis λ_{40}(Figure 4d):

$$\lambda_3 = \frac{\overline{\eta^3}}{\sigma^3} \qquad \text{Eq. 6}$$

$$\lambda_{40} = \frac{\overline{\eta^4}}{\sigma^4} - 3 \qquad \text{Eq. 7}$$

where overbars denote statistical averages, and σ the standard deviation of the surface elevation η (n.b. $\sigma^2 = m_0$). For a Gaussian sea $\lambda_3 = 0,\ \lambda_{40} = 0$. The skewness describes the effects of nonlinearities on the geometry and statistics of the sea surface, with increased skewness implying more pointed crests and shallower, more rounded, troughs (Fedele & Tayfun, 2009; Tayfun, 1980; Tayfun & Fedele, 2007). The rogue seas sit within the bounds of the normal seas (Figure 4c) and the marginal PDF of skewness shows that rogue seas are not particularly skewed (Figure 4f). Therefore, skewness cannot distinguish rogue-containing seas from normal seas.

Rogue seas have increased excess kurtosis compared to normal seas (Figure 4d, g); however, by definition a sea with a rogue wave will have a wave much larger than the surrounding sea, hence an increased kurtosis, and removing the rogue wave from the 20-minute sea reduces the kurtosis (Stansell, 2004).

Spectral bandwidth can be an indicator of the strength of nonlinear focusing (P. Janssen, 2003). The spectral width parameters ε and ν are calculated by:

$$\varepsilon = \sqrt{1 - \frac{m_2^2}{m_0 m_4}} \qquad \text{Eq. 8}$$

$$\nu = \sqrt{\frac{m_2 m_0}{m_1 m_1}} - 1 \qquad \text{Eq. 9}$$

where m_0, m_1, m_2, and m_4 are the zeroth-, first-, second-, and fourth-order spectral moments, respectively, calculated from Eq. 5. For narrow bandwidths ε and ν approach zero, and the wave energy is concentrated near the peak frequency, as individual waves have similar frequency with differing amplitudes modulated by the wave envelope. Values of ε and ν approaching 1 are due to a wide spectrum, with wave energy distributed over widespread frequencies.

Typical values for wave conditions during a storm are $\nu \approx 0.3$-0.5 (Massel, 2013), and normal seas form a distribution about this with a peak at 0.45. The distribution of ν indicates that although the most likely spectral bandwidth is similar for rogue and normal seas (Figure 5a), the probability of getting rogues increases in seas with a higher bandwidth. The distribution of ε (Figure 5b) supports this by indicating rogue waves with an AI>2 are more likely to occur at higher spectral widths, and this would suggest that these rogues are unlikely to be generated by modulational instability. The distribution for the crest height criterion differs from this however, showing higher probability in seas with narrow spectral bandwidth.

The spectral width parameter ν is preferred to ε because ε depends on the fourth order moment of the spectrum (Eq. 8) and tends to infinity logarithmically with the high-frequency cut-off (Tucker & Pitt, 2001). Although ν also depends on a high frequency cut-off, f_c, the variation is less than 10% for $f_c \times T_p > 5$ (Rye, 1977). The wave buoys apply a low-pass filter of 1.5 Hz due to geometric attenuation, when the wave wavelength becomes comparable to the buoy dimensions, and the buoy can no longer follow them. Therefore, for T_p > 3.33 s the variation in ν is less than 10%.

4b. Average wave shape:

Mariners describe the shape of rogue waves as "walls of water" or "holes in the ocean" (Gibbs & Taylor, 2005), fitting the crest height (Eq. 2) and wave height criteria (Eq. 1), respectively. A rogue crest would appear as a "wall of water" above the mean surface level, and for a height criteria rogue, the ship would fall into a deep preceding trough, far below the mean surface level, appearing as a "hole in the ocean". The buoys store surface elevation continuously, allowing an analysis of the shape of rogue waves (Figure 6).

When averaged, the waves that exceed the crest elevation criterion (Eq. 2) have an average crest elevation of 1.48, exceeding the 1.25 threshold. This average rogue wave shape has a larger crest and shallower preceding trough than the average shape of the largest 1% of normal waves, as described by *Walker et al.*, (2004). This differs from the shape seen by *Christou & Ewans*, (2014), which had deeper troughs and a peak of equal height.
However, waves that exceed the wave height criterion (Eq. 1) do not exceed their individual threshold when averaged. This thought to be a consequence of the normalising and averaging, which smooths out the troughs, making them shallower.

We examine this more closely in Figure 7, and try to improve the normalisation by normalising by T_{wave} rather than T_p where:

$$T_{wave} = T_{following\ trough} - T_{preceding\ trough} \qquad \text{Eq. 10}$$

Furthermore, we now average the waves by using the median, a more stable average than the mean, as it is less sensitive to outliers, allowing an improved representation of the average

shape. With an input AI of >2 (Figure 7a), the AI of the average wave is 1.9. This is due to troughs not perfectly aligning and becoming smoothed in the median averaging.
The trough preceding the peak is deeper than that following. To get an average AI of 2, then $AI \geq 2.136$. Increasing the input to $AI \geq 3$ the average AI exceeds the input, with $AI=3.336$. In this case, the trough following the peak is deeper than that preceding. This trend continues with input $AI \geq 4$ and $AI \geq 5$, with the following trough getting deeper, relative to the preceding trough, and displays increased noise, likely due to the reduction in the number of samples with high AI. A deeper trough following a high crest could result in an experience like falling into a “hole” in the ocean that mariners report.

As expected, the crests are peaky and the troughs more rounded, this evidencing the non-linearity despite the wave buoys linearising the sea (Longuet-Higgins, 1963; Tayfun, 1980). The average rogue wave by (crest height criterion only) shape from the *Christou and Ewans*, (2014) database revealed equal minimum elevation of troughs preceding and following the peak, and the shape of six rogue waves, including the Draupner wave, revealed no relationship (Benetazzo et al., 2017).

4c. Spatial variations:

The frequency of occurrence of rogue waves is not the same everywhere (Baschek & Imai, 2011). The spatially diverse dataset compiled here allows for the novel analysis of rogue wave occurrence as a function of averaged sea state parameters (Figure 8).

Rare hazardous events occur at a range of intensities, with the occurrence rate being a decreasing function of their intensity, and often follow a power law rate-intensity relationship With increasing rogue wave prevalence, the height of freak waves (Figure 8a), the significant wave height (Figure 8b), and the zero-crossing period (Figure 8c) of the seas in which they occur, decrease. Zero crossing wave period bifurcates (Figure 8c), with buoys in the Atlantic showing a stronger dependence on wave period compared to those in the Pacific, with Pacific wave period greater than Atlantic locations. This is likely the explanation of the bimodal distribution in the marginal PDF of Tp (Figure 3).

In the Pacific Ocean, rogue wave occurrence shows a relationship with spectral bandwidth parameters and could be indicative of the generation mechanism at specific sites (Figure 9). The distribution of percentage rogue wave occurrence shows that rogue waves are more prevalent in the Southern Californian Bight (SCB; Figure 9). The wave climate in the region is complex (Adams et al., 2008; O'Reilly et al., 2016). Aleutian low sourced waves, approach the SCB from the northwest during La Niña, and more from the west during El Niño (Adams et al., 2008; Graham & Diaz, 2001). There is Northwest swell generated along the California coast, tropical storms formed off Mexico (Inman et al., 1996; Inman & Jenkins, 1997), Southern Hemisphere swell during summer months with small wave height and long period, sea-breeze waves, and Santa Ana wind waves (Adams et al., 2008; Guzman-Morales et al., 2016). The complexity is further compounded by wave refraction, diffraction, and sheltering by Point Conception, at the northern end of the SCB, which blocks waves from >315°, the complex bathymetry of the California Borderlands, and the Channel Islands (Adams et al., 2008; Pawka, 1983; Pawka et al., 1984). It is therefore logical to have high average ν in the region (Figure 9), confirming that the role of instability in forming the rogues in the SCB is likely minimal. Additionally, Kaumalapau, Lanai, Hawaii (CDIP buoy 146) shows high rogue wave occurrence and a large ν.

In contrast, there is high rogue wave occurrence in the Cook Inlet, Alaska (CDIP buoys 175 and 204) but low average ν. The Cook Inlet, has a tidal range of 8-9 m, forcing currents about 1–2 ms^{-1} during full tidal flow, and currents are also generated by wind and baroclinic forcing (Singhal et al., 2013). Wave height and steepness could increase due to a strong opposing current (Kharif & Pelinovsky, 2003b; Onorato et al., 2011; Toffoli et al., 2003). Currents can also alter the dispersion relation and spatially focus wave energy, forming rogue waves (Heller et al., 2008; Lavrenov, 1998; Peregrine, 1976).

In the Southern Gulf of Alaska, Ocean Station Papa (50°N, 145°W) is situated on the southern edge of the cyclonic northeast Pacific subpolar gyre (Pelland et al., 2016). The currents are weak in the low energy Gulf of Alaska (Freeland, 2007) and hence the site is representative of the open Pacific Ocean. The site has low average spectral bandwidth and low freak wave prevalence, further indicating that coastal processes enhance rogue wave occurrence likelihood.

The buoys on the Eastern seaboard of North America are located on the continental shelf and have prevailing offshore winds, explaining a lower average significant wave height compared to the West coast. The prevalence of rogue waves here is greater but their cause of formation difficult to define with the available data. Spectral bandwidth is average in the southern sites and narrows with increasing latitude (Figure 9).

5. Discussion

Wave forecasts provide the characteristic sea state parameters (Hs, Tp, Tz, etc.), and a relationship between them and rogue wave occurrence would provide mariners a computationally cheap tool to assess the likelihood of rogue waves; however, when analysed as a dataset of 1.1 million individual 20-minute seas, no clear link can be found, supporting *Christou & Ewans*, (2014) finding that "rogue waves are not governed by sea state parameters". When the data is examined as 80 spatially differing time series, the rogue wave occurrence likelihood at the location can be examined as a function of the average sea state characteristics. This yields power law relationships between occurrence and mean H_s and mean T_z (Figure 8). This would allow the likelihood of rogue wave occurrence to be predicted at a location given the long-term average sea state characteristics. Furthermore, the application of machine learning tools on the dataset may find novel links based on these parameters by building predictive models that extract patterns from large datasets. To the author's knowledge, this has not been undertaken on an ocean wave dataset and will be performed in a follow-up study.

The spectral width parameter ν could provide a novel indicator of rogue wave occurrence: seas with a high spectral bandwidth may have increased rogue wave likelihood. This finding is in contrast to that of *Christou and Ewans*, (2014) who showed that freak waves were more narrow-banded. Wave groups in seas with narrow spectral bandwidth stay coherent for a longer period than a broadband spectrum; thus, nonlinear instabilities, such as the Benjamin-Feir instability or modulational instability, are more effective. Rogue waves occurring in seas with a broad spectral bandwidth indicates that Benjamin-Feir instability may not be the cause of rogue wave occurrence.

Spatial analysis is complex as wave characteristics at a local scale cannot fully be understood by solely looking at the local conditions as both the locally generated waves, the wind sea, and swell waves from distant storms need to be understood, but this is beyond the scope of our present analysis. In addition, the buoys provide some directionality information through their north and west displacement, which has not been incorporated into this study due to computational constraints. This information could allow the investigation of crossing seas and spreading angle as a rogue wave generation mechanism with the statistical power that this large dataset provides. Again, this is left to a future study.

The cause of formation of rogue waves differs with location: In the Southern Californian Bight, rogues occur with high spectral bandwidth, and therefore mariners may be able to use this as a statistical predictor. In the Cook Inlet however, this would not yield a suitable warning, as entirely different processes may generate the rogue waves. Therefore, it is unlikely that a predictor can be based on one parameter, and any predictors will need to be region specific.

Rogue wave occurrence is low at Ocean Station Papa, the most open-ocean like buoy in the dataset. This suggests that coastal processes amplify the number of rogue waves. However, deep open ocean areas are under-sampled, and hence under-represented in this, and all previous studies, due to the complications of offshore mooring systems for buoys in deep waters and the cost of maintenance.

Wave buoys provide a single point time series and therefore only capture rogue waves occurring at that point, but whether or not a wave is breaking cannot be determined from the time series. It is possible that rogue waves could occur nearby but not directly at the buoy’s locations and hence the likelihood of rogue waves is under represented by buoys (Benetazzo et al., 2015; Fedele et al., 2013). This can be investigated numerically with simulations of high-order spectral calculations of the Euler equations for water waves (Dommermuth &

Yue, 1987; Fedele et al., 2016), and experimentally using stereo imagery to form spatio-temporal records of 3D wave fields (Benetazzo et al., 2012; Gallego et al., 2011). A recent study by *Benetazzo et al.*, (2017), used this method to show that the probability of encountering rogue waves in space and time is at least an order of magnitude larger than when restricting the analysis to a point time series. Additionally, the spatial element is important when considering the rogue wave encounter likelihood for ships and offshore structures which have a spatial footprint rather than simply being at a point (Benetazzo et al., 2017).

The scientific definition of a rogue wave (Eq.1 and Eq.2) form somewhat arbitrary thresholds that do not account for the sudden and severe characteristics of a real rogue wave as reported by mariners. Further work is required to formulate an improved definition that better encompasses the severity and unexpected nature of rogue waves as reported by mariners. It would then be valuable to assess the likelihood of exceeding this improved definition using extreme value analysis.

6. Summary and conclusions

We collated and quality controlled the largest dataset of individual wave profiles for the investigation of rogue waves. The large size still did not yield a discernible link between rogue wave occurrence and the statistics of the 20-minute seas in which they occurred. When the data was assessed as 80 separate locations with a long record of seas, power law relationships of rogue wave occurrence and the average rogue wave height, max wave height, significant wave height and zero crossing wave period were found. With increasing rogue wave prevalence, the height of freak and highest waves, and the significant wave height and zero-crossing period of the seas in which they occur, decrease.

Looking spatially at percentage rogue wave occurrence and the average statistics for each buoy showed that the generation mechanisms for rogue waves is not the same everywhere, and rarely seem to be due to modulational instabilities. The high rogue wave occurrence in the southern California Bight are likely generated by a complex crossing wave fields, whereas in the semi-enclosed seas in Alaska, tidal currents are likely the main mechanism. Therefore, predictors of rogue wave occurrence will need to be region specific.

Future work will use machine learning algorithms to search for novel links between sea state characteristics that have not been sought using the traditional analysis of this paper. Furthermore, the directionality data from the buoys will also be analysed to better understand the influence of crossing seas.

7. Acknowledgements

The wave data were provided by, and freely available from, the Coastal Data Information Program (CDIP), Integrative Oceanography Division, operated by the Scripps Institution of Oceanography, under the sponsorship of the US Army Corps of Engineers and the California Department of Boating and Waterways. Data accessible at: http://thredds.cdip.ucsd.edu/thredds/catalog/cdip/archive/catalog.html. We thank James Behrens for discussions regarding the CDIP buoy data.

This project is partially supported by the Lloyd's Register Foundation, further funding was supplied through the Southampton Marine and Maritime Institute at the University of Southampton. Ben Moat was supported by the NERC funded Project ACSIS.

8. References

Adams, P. N., Inman, D. L., & Graham, N. E. (2008). Southern California Deep-Water Wave Climate: Characterization and Application to Coastal Processes. *Journal of Coastal Research*, *244*, 1022–1035. https://doi.org/10.2112/07-0831.1

Akhmediev, N., Eleonskii, V. M., & Kulagin, N. E. (1987). Exact first-order solutions of the nonlinear Schrodinger equation. *Theoretical and Mathematical Physics*, *72*(2), 809–818. https://doi.org/10.1007/BF01017105

Akhmediev, N., Ankiewicz, A., & Taki, M. (2009). Waves that appear from nowhere and disappear without a trace. *Physics Letters, Section A: General, Atomic and Solid State Physics*, *373*(6), 675–678. https://doi.org/10.1016/j.physleta.2008.12.036

Allender, J., Audunson, T., Barstow, S. F., Bjerken, S., Krogstad, H. E., Steinbakke, P., … Graham, C. (1989). The wadic project: A comprehensive field evaluation of directional wave instrumentation. *Ocean Engineering*, *16*(5–6), 505–536. https://doi.org/10.1016/0029-8018(89)90050-4

Baschek, B., & Imai, J. (2011). Rogue Wave Observations Off the US West Coast. *Oceangraphy*, *24*(2), 158–165. https://doi.org/10.5670/oceanog.2011.35

Beets, R., Hill, C., Coniglione, R., & Portmann, H. (2015). Counter-vandalism at NDBC. In *2014 Oceans - St. John's, OCEANS 2014*. https://doi.org/10.1109/OCEANS.2014.7003182

Belmont, M. R., Christmas, J., Dannenberg, J., Hilmer, T., Duncan, J., Duncan, J. M., & Ferrier, B. (2014). An Examination of the Feasibility of Linear Deterministic Sea Wave Prediction in Multidirectional Seas Using Wave Profiling Radar. *American Meteorologial Society*, *31*, 1601–1614.

Benetazzo, A., Ardhuin, F., Bergamasco, F., Cavaleri, L., Guimarães, P. V, Schwendeman, M., … Torsello, A. (2017). On the shape and likelihood of oceanic rogue waves. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-07704-9

Benetazzo, A., Fedele, F., Gallego, G., Shih, P. C., & Yezzi, A. (2012). Offshore stereo measurements of gravity waves. *Coastal Engineering*, *64*, 127–138. https://doi.org/10.1016/j.coastaleng.2012.01.007

Benetazzo, A., Fedele, F., Gallego, G., Shih, P. C., & Yezzi, A. (2015). Observation of extreme sea waves in a space-time ensemble. *Journal of Physical Oceanography*, *45*(In press), 2261–2275. https://doi.org/10.1175/JPO-D-15-0017.1

Benjamin, T. B., & Feir, J. E. (1967). The disintegration of wave trains on deep water Part 1. Theory. *Journal of Fluid Mechanics*, *27*(03), 417.

https://doi.org/10.1017/S002211206700045X

Birkholz, S., Breé, C., VeseliÄ ‡, I., Demircan, A., & Steinmeyer, G. (2016). Ocean rogue waves and their phase space dynamics in the limit of a linear interference model. *Scientific Reports*, *6*. https://doi.org/10.1038/srep35207

Blondel-Couprie, E., & Naaijen, P. (2012). Deterministic Prediction of Ocean Waves Based on X-Band Radar Measurements. *13emes Journees de l'Hyrdodynamique*, 1–16.

Casas-Prat, M., & Holthuijsen, L. H. (2010). Short-term statistics of waves observed in deep water. *Journal of Geophysical Research: Oceans*, *115*(9). https://doi.org/10.1029/2009JC005742

Chabchoub, A., Hoffmann, N., Onorato, M., & Akhmediev, N. (2012). Super rogue waves: observation of a higher-order breather in water waves. *Physical Review X*, *2*(1), 011015. https://doi.org/10.1103/PhysRevX.2.011015

Christou, M., & Ewans, K. (2014). Field Measurements of Rogue Water Waves. *Journal of Physical Oceanography*, *44*(9), 2317–2335. https://doi.org/10.1175/JPO-D-13-0199.1

Costa, A., Osborne, A. R., Resio, D. T., Alessio, S., Chrivì, E., Saggese, E., … Long, C. E. (2014). Soliton turbulence in shallow water ocean surface waves. *Physical Review Letters*, *113*(10). https://doi.org/10.1103/PhysRevLett.113.108501

Cousins, W., & Sapsis, T. P. (2016). Reduced order prediction of rare events in unidirectional nonlinear water waves. *J. Fluid Mech.*, *790*, 368–388. https://doi.org/10.1017/jfm.2016.13

Dannenberg, J., Naaijen, P., Hessner, K., van den Boom, H., & Reichert, K. (2010). The On board Wave Motion Estimator OWME. In *Proceedings of the Twentieth (2010) International Offshore and Polar Engineering Conference*. Beijing, China: ISOPE.

Dommermuth, D. G., & Yue, D. K. P. (1987). A high-order spectral method for the study of nonlinear gravity waves. *Journal of Fluid Mechanics*, *184*(1), 267. https://doi.org/10.1017/S002211208700288X

Draper, L. (1966). Freak Ocean Waves. *Weather*, *21*(1), 2–4. https://doi.org/10.1002/j.1477-8696.1966.tb05176.x

Dysthe, K., Krogstad, H. E., & Müller, P. (2008). Oceanic Rogue Waves. *Annual Review of Fluid Mechanics*. https://doi.org/10.1146/annurev.fluid.40.111406.102203

Fedele, F., Benetazzo, A., Gallego, G., Shih, P. C., Yezzi, A., Barbariol, F., & Ardhuin, F. (2013). Space-time measurements of oceanic sea states. *Ocean Modelling*, *70*, 103–115. https://doi.org/10.1016/j.ocemod.2013.01.001

Fedele, F., Brennan, J., Ponce De León, S., Dudley, J., & Dias, F. (2016). Real world ocean rogue waves explained without the modulational instability. *Scientific Reports*, *6*. https://doi.org/10.1038/srep27715

Fedele, F., & Tayfun, M. A. (2009). On nonlinear wave groups and crest statistics. *Journal of Fluid Mechanics*, *620*, 221. https://doi.org/10.1017/S0022112008004424

Forristall, G. Z. (2005). Understanding rogue waves: Are new physics really necessary? In *Rogue Waves, Proc. 14th Aha Huliko: A Hawaiian Winter Workshop, University of Hawaii, Honolulu, HI.* (pp. 29–35).

Freeland, H. (2007). A short history of Ocean Station Papa and Line P. *Progress in Oceanography*, *75*(2), 120–125. https://doi.org/10.1016/j.pocean.2007.08.005

Gallego, G., Yezzi, A., Fedele, F., & Benetazzo, A. (2011). A variational stereo method for the three-dimensional reconstruction of ocean waves. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(11 PART 2), 4445–4457. https://doi.org/10.1109/TGRS.2011.2150230

Gibbs, R. H., & Taylor, P. H. (2005). Formation of walls of water in “fully” nonlinear simulations. *Applied Ocean Research*, *27*(3), 142–157. https://doi.org/10.1016/j.apor.2005.11.009

Gibson, R. S., & Swan, C. (2007). The evolution of large ocean waves: the role of local and rapid spectral changes. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *463*(2077), 21–48. https://doi.org/10.1098/rspa.2006.1729

Graham, N. E., & Diaz, H. F. (2001). Evidence for intensification of North Pacific winter cyclones since 1948. *Bulletin of the American Meteorological Society*, *82*(9), 1869–1893. https://doi.org/10.1175/1520-0477(2001)082<1869:EFIONP>2.3.CO;2

Guzman-Morales, J., Gershunov, A., Theiss, J., Li, H., & Cayan, D. (2016). Santa Ana Winds of Southern California: Their climatology, extremes, and behavior spanning six and a half decades. *Geophysical Research Letters*, *43*(6), 2827–2834. https://doi.org/10.1002/2016GL067887

Hallerberg, S., Bröcker, J., & Kantz, H. (2008). Prediction of Extreme Events. In R. V Donner & S. M. Barbosa (Eds.), *Nonlinear Time Series Analysis in the Geosciences: Applications in Climatology, Geodynamics and Solar-Terrestrial Physics* (pp. 35–59). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-78938-3_3

Haver, S. (2000). Evidences of the Existence of Freak Waves. In *Rogues Waves 2000*.

Heller, E. J., Kaplan, L., & Dahlen, A. (2008). Refraction of a Gaussian seaway, *113*(May), 1–14. https://doi.org/10.1029/2008JC004748

Inman, D. L., Jenkins, S. A., & Elwany, M. H. S. (1996). Wave climate cycles and coastal engineering practice. In *25th International Conference Coastal Engineering Research Council ASCE* (pp. 314–327). Orlando, Florida.

Inman, D. L., & Jenkins, S. A. (1997). Changing wave climate and littoral drift along the California coast. In *California and the World Ocean ‘97* (pp. 538–549). San Diego, ASCE.

James, I. D. (1986). A note on the theoretical comparison of wave staffs and wave rider buoys in steep gravity waves. *Ocean Engineering*, *13*(2), 209–214. https://doi.org/10.1016/0029-8018(86)90028-4

Janssen, P. (2003). Nonlinear Four-Wave Interactions and Freak Waves. *Journal of Physical Oceanography*, *33*(4), 863–884. https://doi.org/10.1175/1520-0485(2003)33<863:NFIAFW>2.0.CO;2

Janssen, T. T., & Herbers, T. H. C. (2009). Nonlinear Wave Statistics in a Focal Zone. *Journal of Physical Oceanography*, *39*(8), 1948–1964. https://doi.org/10.1175/2009JPO4124.1

Kharif, C., & Pelinovsky, E. (2003a). Physical mechanisms of the rogue wave phenomenon. *European Journal of Mechanics - B/Fluids*, *22*(6), 603–634. https://doi.org/10.1016/j.euromechflu.2003.09.002

Kharif, C., & Pelinovsky, E. (2003b). Physical mechanisms of the rogue wave phenomenon.

European Journal of Mechanics, B/Fluids, *22*(6), 603–634. https://doi.org/10.1016/j.euromechflu.2003.09.002

Kharif, C., Pelinovsky, E., & Slunyaev, A. (2009). *Rogue Waves in the Ocean*. *Advances in Geophysical and Environmental Mechanics and Mathematics*. Berlin: Springer-Verlag. https://doi.org/10.1017/CBO9781107415324.004

Kibler, B., Fatome, J., Finot, C., Millot, G., Dias, F., Genty, G., … Dudley, J. M. (2010). The Peregrine soliton in nonlinear fibre optics. *Nature Physics*, *6*(10), 790–795. https://doi.org/10.1038/nphys1740

Lavrenov, I. V. (1998). The wave energy concentration at the Agulhas Current off South Africa. *Natural Hazards*, *17*(2), 117–127. https://doi.org/10.1023/A:1007978326982

Longuet-Higgins, M. S. (1963). The effect of non-linearities on statistical distributions in the theory of sea waves. *Journal of Fluid Mechanics*, *17*(3), 459–480. https://doi.org/10.1017/S0022112063001452

Magnusson, A. K., Donelan, M. A., & Drennan, W. M. (1999). On estimating extremes in an evolving wave field. *Coastal Engineering*, *36*(2), 147–163. https://doi.org/10.1016/S0378-3839(99)00004-6

Makri, I., Rose, S., Christou, M., Gibson, R., & Feld, G. (2016). Examining Field Measurements of Deep-Water Crest Statistics. In *ASME 2016 35th International Conference on Ocean, Offshore and Arctic Engineering* (p. 10). Busan, South Korea. https://doi.org/10.1115/OMAE2016-54363

Massel, S. R. (2013). *Ocean Surface Waves: Their Physics and Prediction* (2nd ed.). Advanced Series on Ocean Engineering. Retrieved from http://www.worldscientific.com/worldscibooks/10.1142/8682

O’Reilly, W. C., Olfe, C. B., Thomas, J., Seymour, R. J., & Guza, R. T. (2016). The California coastal wave monitoring and prediction system. *Coastal Engineering*, *116*, 118–132. https://doi.org/10.1016/j.coastaleng.2016.06.005

Onorato, M., Proment, D., & Toffoli, A. (2011). Triggering rogue waves in opposing currents. *Physical Review Letters*, *107*(18). https://doi.org/10.1103/PhysRevLett.107.184502

Pawka, S. S. (1983). Island shadows in wave directional spectra. *Journal of Geophysical Research–Oceans and Atmospheres*, *88*, 2579–2591.

Pawka, S. S., Inman, D. L., & Guza, R. T. (1984). Island sheltering of surface gravity waves: model and experiment. *Continental Shelf Research*, *3*(1), 35–53. https://doi.org/10.1016/0278-4343(84)90042-6

Pelland, N. A., Eriksen, C. C., & Cronin, M. F. (2016). Seaglider surveys at Ocean Station Papa: Circulation and water mass properties in a meander of the North Pacific Current. *Journal of Geophysical Research: Oceans*, *121*(9), 6816–6846. https://doi.org/10.1002/2016JC011920

Peregrine, D. H. (1976). Interaction of Waves and Currents. In *Advances in Applied Mechanics* (Vol. 16, pp. 9–117). https://doi.org/10.1016/S0065-2156(08)70087-5

Peregrine, D. H. (1983). Water waves, nonlinear Schrödinger equations and their solutions. *The Journal of the Australian Mathematical Society. Series B. Applied Mathematics*, *25*(01), 16–43.

Rye, H. (1977). The stability of some currently used wave parameters. *Coastal Engineering*, *1*, 17–30. Retrieved from https://doi.org/10.1016/0378-3839(77)90004-7

Seymour, R. J., & Castel, D. (1998). Systematic Underestimation of Maximum Crest Heights in Deep Water Using Surface-following Buoys. In ASME (Ed.), *17th International Conference on Offshore Mechanics and Arctic Engineering*.

Singhal, G., Panchang, V. G., & Nelson, J. A. (2013). Sensitivity assessment of wave heights to surface forcing in Cook Inlet, Alaska. *Continental Shelf Research*, *63*. https://doi.org/10.1016/j.csr.2012.02.007

Slunyaev, a., Pelinovsky, E., & Guedes Soares, C. (2005). Modeling freak waves from the North Sea. *Applied Ocean Research*, *27*(1), 12–22. https://doi.org/10.1016/j.apor.2005.04.002

Stansell, P. (2004). Distributions of freak wave heights measured in the North Sea. *Applied Ocean Research*, *26*(1–2), 35–48. https://doi.org/10.1016/j.apor.2004.01.004

Tayfun, M. A. (1980). Narrow-band nonlinear sea waves. *Journal of Geophysical Research: Oceans*, *85*(9), 1548–1552. https://doi.org/10.1029/JC085iC03p01548

Tayfun, M. A. (2008). Distributions of Envelope and Phase in Wind Waves. *Journal of Physical Oceanography*, *38*(2), 2784–2800. https://doi.org/10.1175/2008JPO4008.1

Tayfun, M. A., & Fedele, F. (2007). Wave-height distributions and nonlinear effects. *Ocean Engineering*, *34*(11–12), 1631–1649. https://doi.org/10.1016/j.oceaneng.2006.11.006

Thomson, J., Talbert, J., de Klerk, A., Brown, A., Schwendeman, M., Goldsmith, J., … Meinig, C. (2015). Biofouling effects on the response of a wave measurement buoy in deep water. *Journal of Atmospheric and Oceanic Technology*, *32*(6), 1281–1286. https://doi.org/10.1175/JTECH-D-15-0029.1

Toffoli, A., Lefevre, J., Monbaliu, J., Savina, H., & Gregersen, E. (2003). Freak Waves: Clues for Prediction in Ship Accidents? In *Proceedings of The Thirteenth (2003) International Offshore and Polar Engineering Conference* (pp. 23–29). Honolulu, Hawaii: The International Society of Offshore and Polar Engineers.

Tucker, M. J., & Pitt, E. G. (2001). *Waves in Ocean Engineering*. *Elsevier Ocean Engineering Book Series*. Retrieved from http://trid.trb.org/view.aspx?id=699480

Walker, D. A. G., Taylor, P. H., & Taylor, R. E. (2004). The shape of large surface waves on the open sea and the Draupner New Year wave. *Applied Ocean Research*, *26*(3–4), 73–83. https://doi.org/10.1016/j.apor.2005.02.001

Figure Captions:

Figure 1: Map showing the location and name of the 80 Datawell waverider buoys used in the study. The point colour indicates the water depth at the buoys location.

Figure 2: Grey points represent normal seas and black rogue seas and display: a) the maximum wave height of each 20-minute sea that passed QC as a function of the significant

wave height, with the degrees of abnormality index (h/Hs) marked; and b) the maximum crest elevation in each 30-minute sea as a function of significant wave height, with degrees of abnormality (C_{max}/Hs) displayed.

Figure 3: a) Probability density function of significant wave height for seas containing a rogue wave (black dashed line) and normal seas (grey fill). b) Significant wave height with peak period, indicating wave steepness, for 20-minute samples of rogue seas (black points) and normal seas (grey points). c) Probability density function of peak period height for rogue seas (black dashed line) and normal seas (grey fill).

Figure 4: a) The probability density function of the maximum crest height of the 20-miniute sea for rogue seas (black dashed line) and normal seas (grey fill). Maximum crest height as a function of b) sea state steepness S_1, c) skewness, and d) excess kurtosis. Probability density functions of e) sea state steepness S_1, f) skewness, and g) excess kurtosis for rogue seas (black dashed line) and normal seas (grey fill).

Figure 5: Probability density functions of spectral bandwidth parameters a) ν and b) ε for normal seas (grey fill), rogue seas – crest criteria (black dot), and rogue seas height criteria (black dash).

Figure 6: The average height and period normalised wave shape of rogue waves with a crest height greater than 1.25 Hs (red), rogue waves with a wave height greater than 2 Hs (blue), and the highest 1% of normal waves (green).

Figure 7: The average shape of the peak aligned and normalised (with Hs and Twave) sea surface elevation for a range of input AI: a) 2, b) 2.136, displaying an average AI of 2, c) 3, d) 4, and e) 5. One standard deviation about the median is shown in grey shade.

Figure 8: Logged statistical average (denoted by overbar) of a) freak wave height, b) significant wave height, and c) zero up-crossing wave period, as a function of logged percentage rogue seas for each of the 80 wave buoys. Water depth at the buoy location is denoted with point colour and ocean by shape: squares for Pacific Ocean and circles for Atlantic Ocean. Linear regressions and associated parameters are displayed.

Figure 9: Map of the percentage rogue seas (marker size) and the average spectral bandwidth parameter ν (marker colour).