

# A Causal Predictive Model for Real-World Rogue Wave Probabilities

DION HÄFNER\*

*Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark*

JOHANNES GEMMRICH

*University of Victoria, Victoria, British Columbia, Canada*

MARKUS JOCHUM

*Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark*

## ABSTRACT

Extreme waves in the ocean (“rogue waves”) are well studied in theory and lab experiments under idealized conditions, but relatively little research is based on direct observations. Therefore, it is still unclear how well common approximations hold up in the real ocean. Here, we present a predictive model that combines parameters representing both linear and nonlinear wave dynamics through an artificial neural network, trained directly on observations from wave buoys. By imposing strong architectural constraints we arrive at a model that is approximately invariant under retraining across a wide regime of sea states, suggesting causal consistency with the dominant rogue wave generation processes, and that we can analyze in detail. We find that a combination of crest-trough correlation, characteristic steepness, directionality index, and relative water depth has high causal consistency, outputs well-calibrated probabilities, and achieves good predictive scores on unseen data. This paves the way towards a higher quality rogue wave forecast.

## 1. Introduction

Oceanic rogue waves (also called freak waves) are extreme ocean waves that are suspected to have caused countless accidents, often with fatal consequences (Didenkulova 2019). They are typically defined as any wave whose crest-to-trough height  $H$  exceeds a certain threshold relative to the significant wave height  $H_s$ . The significant wave height in turn is defined as 4 times the standard deviation of the sea surface elevation. Here, we use a rogue wave criterion with a threshold of 2.0 instead of the perhaps more common 2.2 in order to increase the number of studied rogue waves (but our results can be extended to the stricter threshold):

$$H/H_s > 2.0 \quad (1)$$

A rogue wave is therefore by definition an unlikely sample from the tail of the wave height distribution (with a probability of about  $3 \times 10^{-4}$  under linear theory, Longuet-Higgins 1952). This also implies that rogue waves are an extreme event that can in principle occur by chance under any circumstance, which makes them difficult to analyze, and requires massive amounts of data. Therefore, research has mostly focused on theory and idealized experiments in wave tanks, often considering only 1-dimensional wave propagation (see Dudley et al. 2019, for a review).

In previous work we assembled a database of over 1 billion wave observations from buoys (FOWD, Häfner et al. 2021a). FOWD is a catalogue that maps individual wave observations to about 80 characteristic sea state parameters describing the circumstances under which the wave was measured. This allowed us to analyze how rogue wave occurrence probabilities depend on the sea state, where we found that crest-trough correlation is by far the best univariate predictor for rogue wave occurrence (Häfner et al. 2021b). This parameter is not included in today’s operational freak wave forecasts such as that of the European Centre for Medium-Range Weather Forecasts (ECMWF), which instead focuses on nonlinear effects governed by parameters like characteristic steepness, directionality index, and the Benjamin-Feir index (ECMWF 2021). This suggests that an improved rogue wave forecast that takes crest-trough correlation into account is within reach. However, all analysis of FOWD so far has been purely descriptive and cannot be used for forecasting.

In this study, we present a neural network-based machine learning model that predicts rogue wave probabilities from the sea state, trained solely on observations. The resulting model respects the causal structure of rogue wave generation — that means it is robust to distributional shift and can be used to infer the relative importance of rogue wave generation mechanisms. We achieve this by combining a careful a-priori analysis of causal pathways that leads to

---

\*Corresponding author: Dion Häfner, dion.haefner@nbi.ku.dk

a set of presumed causal parameters (Section 2), regularization constraints, and an a-posteriori model evaluation to identify the model architecture that shows the highest invariance to shifting environments (Section 3). Analyzing this model allows us to study how rogue wave probabilities depend on the sea state, even in higher-dimensional settings (Section 4).

## 2. The causes of rogue waves

To ensure that our machine learning model learns causal relationships instead of mere associations, it is essential to only include parameters that carry causal meaning (otherwise the model might prefer spurious associations that are easier to learn).

There are many suspected causes of rogue waves (see Adcock and Taylor 2014, for an overview). Typically, research focuses on linear bandwidth-limited seas (Tayfun and Fedele 2007), weakly nonlinear seas (Gemmrich and Garrett 2011; Fedele et al. 2016), or the highly nonlinear modulational instability (Onorato et al. 2006). Apart from these universal mechanisms, there are also countless possible interactions with localized features like topography such as (Trulsen et al. 2012) or underwater currents, interactions with currents like in the Agulhas (Mallory 1974) or in Drake passage (Didenkulova et al. 2021), or crossing sea states at high crossing angles (McAllister et al. 2019).

The go-to tool to analyze causal relationships is a causal DAG (directed acyclic graph), where nodes represent variables and edges  $A \rightarrow B$  imply that  $A$  is a cause of  $B$  (usually in the probabilistic sense in that the probability distribution  $P(B)$  depends on  $A$ ). In the frame of this analysis, we would like to relate sea state parameters  $\mathcal{P}$  to physical effects  $\Phi$ , which in turn influence wave observations  $\mathcal{O}$ . The resulting causal graph for rogue waves containing the previously discussed pathways is shown in Fig. 1.

Following this causal structure, we use the following set of sea state parameters as candidates for representing the various causal pathways (see Appendix A for more information on each parameter):

**Crest-trough correlation  $r$**  (a parameter related to spectral bandwidth) to account for linear effects. As we showed in Häfner et al. (2021b), the rogue wave probability  $p$  is conditionally independent of other bandwidth measures (such as narrowness and peakedness) when conditioning on  $r$ , but not vice-versa. This suggests that  $r$  is the dominant causal factor behind linear rogue wave formation.

**Steepness  $\varepsilon$**  governing weakly nonlinear effects (such as second-order and third-order bound waves) and wave breaking (Miche 1944; Goda 2010).

**Relative high-frequency energy  $E_h$**  (fraction of total energy contained in the spectral band 0.25 Hz to 1.5 Hz) as a proxy for the strength of local winds.

**Relative depth  $\tilde{D}$**  (based on peak wavenumber), which is central for nonlinear effects (Korteweg and De Vries 1895; Janssen 2018) and wave breaking (Miche 1944).

**Benjamin-Feir index BFI** which controls third-order nonlinear free waves (Janssen 2018) and the modulational instability (Janssen 2003).

**Ursell number  $Ur$**  which quantifies nonlinear effects in shallow water (Ursell 1953).

**Dominant directional spread  $\sigma_\theta$**  which has an influence on third-order nonlinear waves (Janssen 2018) and wave breaking (McAllister et al. 2019).

**Spectral bandwidth  $\nu$**  appearing in the expression for the influence of third-order nonlinear waves (Janssen 2018).

**Directionality index  $R$**  (the ratio of directional spread and spectral bandwidth) controlling third-order nonlinear free waves (often used in conjunction with the BFI, Janssen 2018).

There are some notable omissions from this list. Firstly, mean period and significant wave height, perhaps the most studied sea state parameters of all. While these parameters *do* play an important role in the rogue wave generation process, they appear higher up in the causal graph and are therefore not direct causes of rogue waves (instead, they *generate the conditions* that are causing rogue waves). Secondly, surface elevation skewness and kurtosis, which are also studied extensively in connection with rogue waves (Stansell 2004; Mori and Janssen 2006; Fedele and Tayfun 2009), but which we have shown to be too noisy to be of use for rogue wave prediction (Häfner et al. 2021b).

Unfortunately, there are still entirely unobserved causal paths, since we do not have access to data on local winds, topography, or currents. Additionally, all measurements are potentially biased estimates of the true sea state parameters. Therefore we cannot rely on the resulting model to be causally consistent by itself, and we will have to perform a-posteriori verification on the learned model to check for causal consistency across different environments (see Section 3c).

## 3. A causally consistent predictive model

### a. Input data

The main data source for this analysis is the Free Ocean Wave Dataset (FOWD, Häfner et al. 2021a). The pre-filtered version of FOWD consists of 1.4 billion wave measurements, originally recorded by 158 CDIP wave buoys (Behrens et al. 2019) along the Pacific and Atlantic coasts of the US, Hawaii, and overseas US territories at water depths between 10 m to 4000 m and a minimum significant wave height of 1 m. Each buoy records the sea surface elevation at

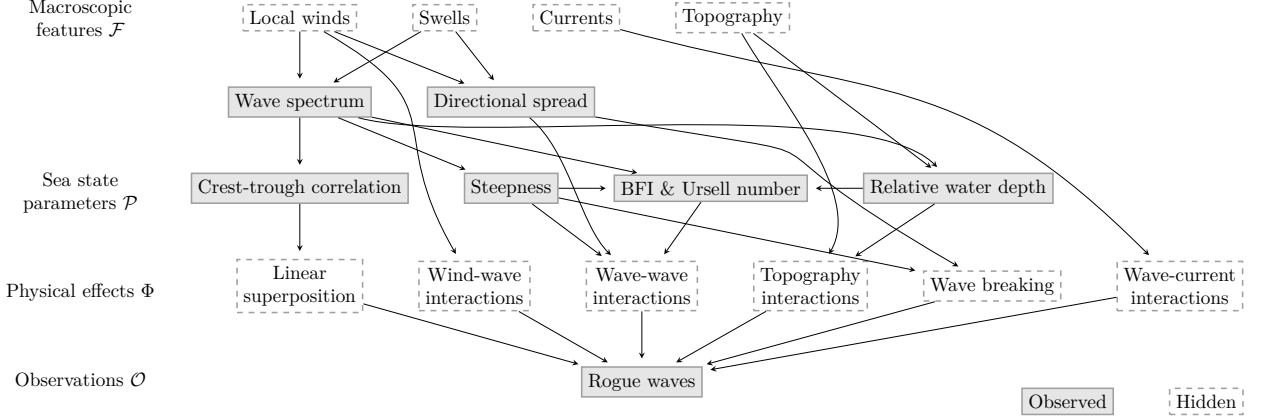


FIG. 1. The causes of rogue waves as a causal DAG (directed acyclic graph). Arrows  $A \rightarrow B$  imply that  $A$  causes  $B$ .

a sampling frequency of 1.28 Hz, which amounts to over 700 yr of time series in total. FOWD then extracts every zero-crossing wave from the CDIP data and computes about 20 characteristic sea state parameters from the history of the wave within a 10 min, 30 min, and dynamically sized window (as suggested in Boccotti 2000). Here, we only use parameters based on the dynamic window size.

Due to the massive data volume of the full FOWD catalogue ( $\sim 1$  TB), we use an aggregated version that maps each sea state to the maximum wave height of the following 100 waves (similar to what is used in Häfner et al. 2021b). This reduces the data volume by a factor of 100 and inflates all rogue wave probabilities, which can be corrected for via  $p = 1 - (1 - \hat{p})^{1/100}$  where  $\hat{p}$  is the inflated probability, assuming that rogue waves occur independently from each other.

This process leaves us with 12.9M data points containing just over 100,000 rogue waves exceeding  $2H_s$ . This input dataset is freely available for download at <https://erda.ku.dk/archives/7072a6eb3d181149deb56b7d8739805e/published-archive.html>.

### b. Model architecture

Our core assumption is that the rogue wave probability can be modelled as:

$$\text{logit } P(y = 1 \mid \mathbf{x}) \sim \sum_i f_i(\mathbf{x}^{(S_i)}) + b \quad (2)$$

where  $y$  is a binary label indicating whether the current wave is a rogue wave,  $\mathbf{x}^{(S_i)}$  denotes the  $i$ -th subset of all causal sea state parameters  $\mathbf{x}$  (see Section 2),  $\text{logit}(p) = \log(p) - \log(1 - p)$  is the logit function,  $f_i$  are arbitrary nonlinear functions to be learned, and  $b$  is a constant bias term. By including only a subset  $\mathbf{x}^{(S_i)}$  of all parameters  $\mathbf{x}$  as input for each term we can limit which parameters

may interact with each other as an additional regularizing constraint.

For example, to include the effects of linear superposition and nonlinear corrections for free and bound waves similar to ECMWF (2021) we can use:

$$\text{logit } P(y = 1 \mid \mathbf{x}) \sim \underbrace{f_1(r)}_{\text{linear}} + \underbrace{f_2(\text{BFI}, R)}_{\text{free waves}} + \underbrace{f_3(\varepsilon, \tilde{D})}_{\text{bound waves}} \quad (3)$$

with Benjamin-Feir index BFI, directionality index  $R$ , steepness  $\varepsilon$ , and relative depth  $\tilde{D}$ .

We parametrize the functions  $f_i$  via fully connected neural networks (FCNs), which have been shown to be universal function approximators (Hornik 1991), and that can be trained efficiently for large amounts of data. The set of functions  $f_i$  can be represented as a single multi-head FCN (one head for each input subset  $\mathbf{x}^{(S_i)}$ ) with a linear output layer (Fig. 2). We use a simple feedforward architecture with 3 hidden layers and ReLU activation functions (rectified linear unit, Nair and Hinton 2010).

The neural network outputs a scalar  $\tilde{p} = \text{logit } P(y = 1 \mid \mathbf{x}) \in (-\infty, \infty)$ , the log-odds of a rogue wave occurrence for the given sea state. For training, we use the Adam optimizer and backpropagation to minimize a typical cross-entropy loss for binary classification with added  $\ell_1$  and  $\ell_2$  regularization terms for kernel parameters:

$$L(p, y, \theta) = y \cdot \log(p) + (1 - y) \cdot \log(1 - p) + \lambda_1 \|\theta\|_1 + \lambda_2 \|\theta\|_2 \quad (4)$$

with predicted probability  $p = \text{logit}^{-1}(\tilde{p})$ , observed labels  $y \in \{0, 1\}$  (rogue wave or not), and neural network kernel parameters  $\theta$ .

To estimate uncertainties in the neural network parameters and resulting predictions, we use Gaussian stochastic weight averaging (SWAG, Maddox et al. 2019). For this,

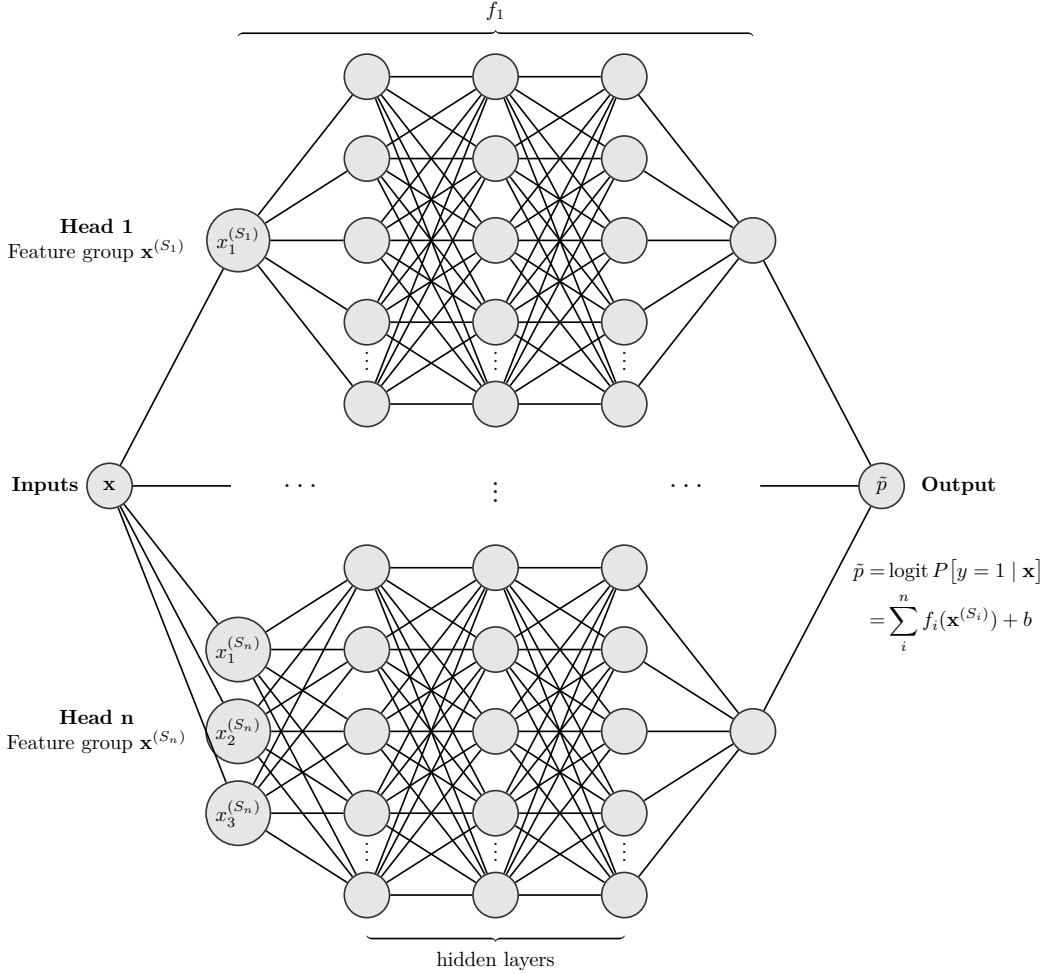


FIG. 2. Neural network architecture (multi-head FCN) used to predict rogue wave probabilities. Each input head receives a different subset of the full parameter set  $\mathbf{x}$  to limit the amount of non-causal interactions between parameters.

we train the network for 50 epochs, then start recording the optimizer trajectory after each epoch for another 50 epochs. The observed covariance structure of the sampled parameters is then used to construct a multivariate Gaussian approximation of the loss surface close to the minimum that we can sample from. This results in slightly better predictions, but more importantly, also gives us a way to quantify how confident the neural network is in its predictions.

Appendix B lists the full set of model hyperparameters.

### c. Causal consistency

Even though we only include input parameters that we assume to have a direct causal connection with rogue wave generation, there is no guarantee that the neural network will learn the correct causal connections. In fact, the presence of measurement bias and unobserved causal paths makes it unlikely that the model will converge to the true causal structure unless the right amount of regularization

and architectural constraints are applied. To search for a maximally causally consistent model we will have to quantify its causal consistency.

We can achieve this through the concept of invariant causal prediction (ICP, Peters et al. 2016, 2017). The key insight behind ICP is that only the true causal model will be invariant under distributional shift, in the sense that re-training the model on data with different correlations *between* features should still lead to the same dependency of the target *on* the features. Ideally, the chosen environments lead to a significant distributional shift by changing the relative importance of different causal mechanisms, but not the causal structure itself.

To exploit ICP we split the full dataset randomly into separate training and validation sets, in chunks of 1M waves (to ensure that the validation data is truly unseen by the model, but covers roughly the same range of sea states). We train the model on the training dataset (66% of all

TABLE 1. The subsets of the validation data set used to evaluate model invariance.

Subset name	Condition	# waves
southern-california	Longitude $\in (-123.5, -117)^\circ$ , Latitude $\in (32, 38)^\circ$	233M
deep-stations	Water depth $> 1000\text{m}$	33M
shallow-stations	Water depth $< 100\text{m}$	138M
summer	Day of year $\in (160, 220)$	44M
winter	Day of year $\in (0, 60)$	88M
$H_s > 3\text{m}$	$H_s > 3\text{m}$	55M
high-frequency	Relative swell energy $< 0.15$	40M
low-frequency	Relative swell energy $> 0.7$	42M
long-period	Mean zero-crossing period $> 9\text{s}$	40M
short-period	Mean zero-crossing period $< 6\text{s}$	90M
cnoidal	Ursell number $> 8$	34M
weakly-nonlinear	Steepness $> 0.04$	80M
spectral-narrow	Directionality index $< 0.3$	68M
spectral-wide	Directionality index $> 1$	37M
full	(all validation data)	438M

data) and perform ICP on the validation dataset (34 % of all data), which we partition into subsets representing different conditions in space, time, depth, mean period, and degrees of non-linearity (Table 1). Then, we re-train the model separately on each subset and compare the performance on the  $k$ -th data subset  $\mathbf{x}_{(k)}$  between the re-trained model  $P_k$  and the full model  $P_{\text{tot}}$  by computing the root-mean-square error of predictions:

$$\mathcal{E}_k^2 = \frac{1}{n_k} \sum_i^{n_k} \left( \text{logit } P_k(\mathbf{x}_i^{(k)}) - \text{logit } P_{\text{tot}}(\mathbf{x}_i^{(k)}) \right)^2 \quad (5)$$

where  $n_k$  is the number of data points in the subset  $\mathbf{x}_{(k)}$ . As the total consistency score we use the root-mean-square across all environments:

$$\mathcal{E} = \sqrt{\frac{1}{n_E} \sum_k^{n_E} \mathcal{E}_k^2} \quad (6)$$

Under a noise-free, infinite dataset and an unbiased training process that identifies the true causal model we would find  $\mathcal{E} = 0$ , i.e., re-training the model on the unseen data subset would not contribute any new information and leave the model perfectly invariant. Since all 3 of these assumptions are violated here, we merely search for the model that minimizes  $\mathcal{E}$ , while also having a competitive predictive score and well-calibrated probabilities (since for example a trivial model predicting  $P(\mathbf{x}) = c$  with constant  $c$  has perfect invariance of predictions).

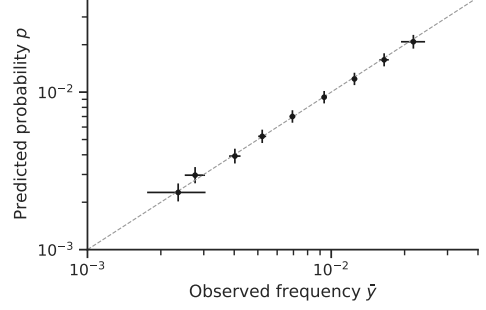


FIG. 3. Our model outputs well-calibrated probabilities. Shown is the binned predicted probability  $p$  vs. the observed rogue wave incidence  $\bar{y}$ . Error bars for  $p$  indicate 3 standard deviations estimated via SWAG sampling. Error bars for  $\bar{y}$  indicate 99 % credible interval assuming  $\bar{y}_i \sim \text{Beta}(n_i^+, n_i^-)$  with  $n_i^+$  rogue and  $n_i^-$  non-rogue measurements in the  $i$ -th bin. Dashed line indicates perfect calibration.

#### d. Other performance metrics

We also need performance metrics that describe the predictive capabilities of the model and measure overfitting, specifically whether the model tends to produce overconfident or underconfident predictions.

To evaluate predictive performance we use the mean of a base rate-adjusted log-likelihood score across all environments (Table 1):

$$\mathcal{L}(p, \bar{y}) = \frac{1}{n_E} \sum_k^{n_E} (I(p_k) - I(\bar{y}_k)) \quad (7)$$

$$I(x) = x \cdot \log(x) + (1-x) \cdot \log(1-x) \quad (8)$$

with predicted probabilities  $p$  and base rate  $\bar{y} = \frac{1}{n} \sum_i^n y$  for each environment. This gives the log-likelihood of observing the data given the model, relative to a model that just predicts its base rate  $\bar{y}$ . Using the mean over environments ensures that the model performs well in different physical regimes instead of focusing on average conditions. Since we do not have access to ground truth probabilities we cannot quantify predictive quality in an absolute sense, and the above score can only be used to compare different models on the same data.

To evaluate model calibration we compute a calibration curve by binning the predicted probabilities and comparing each bin to the observed rogue wave frequency (Fig. 3). As the total calibration score we use the root-mean-square error between measured and predicted log-odds:

$$C = \sqrt{\frac{1}{n_b} \sum_{k=1}^{n_b} (\text{logit}(p_i) - \text{logit}(\bar{y}_i))^2} \quad (9)$$

with number of bins  $n_b$  and corresponding mean prediction  $p_i$  and observed base rate  $\bar{y}_i$ .

#### 4. Results

We train a total of 24 different candidate models using the procedure described in Section 3 and evaluate their performance in terms of calibration, data likelihood, and causal consistency (Table 2).

We observe a clear anti-correlation between model complexity and predictive score on the one hand and causal consistency on the other hand, as long as only causal parameters are included (as identified in Section 2). When including additional non-causal parameters (experiment 24), the predictive score decreases slightly while the causal consistency error increases drastically (as expected).

Still, all experiments show a significant non-zero consistency error  $\mathcal{E}$  of around 0.1 (mean SWAG uncertainties in predictions are about 0.03). Selecting among the remaining models is therefore a trade-off between bias (prediction score) and variance (consistency score). There are 2 models that represent a compromise with good performance in all metrics:

1. Model 16 with two parameter groups  $S_1 = \{r, R\}$ ,  $S_2 = \{\varepsilon, \bar{D}, R\}$ .
2. Model 17 with a single parameter group  $S_1 = \{r, \varepsilon, \bar{D}, R\}$ .

We apply Occam’s razor and choose model 16 as the reference model for further analysis — even though it has slightly lower predictive and calibration scores — due to the lower number of possible parameter interactions (at most 3-way interactions instead of 4-way).

Our analysis of this reference model and selected other experiments leads to the following 3 main results of this study.

*a. Rogue wave models should account for crest-trough correlation, steepness, relative depth, and directionality*

Only this parameter combination achieves good causal consistency and predictive scores at the same time, and experiments that exclude any of these parameters perform unconditionally worse (see e.g. experiments 1–6). Especially the exclusion of crest-trough correlation leads to catastrophic results, even when including other bandwidth measures in its place (experiments 3, 13).

Models that use the directionality index  $R$  over raw directional spread  $\sigma_\theta$  are generally more causally consistent (experiment 17 vs. 19). Also, an interaction between crest-trough correlation and directionality index seems essential to achieve optimal performance (12 vs. 16). Higher predictive performance can be achieved by including spread  $\sigma_\theta$  and spectral narrowness  $\nu$  directly (instead of via  $R = \sigma_\theta^2/2\nu^2$ ), at the expense of increased overfitting (experiment 21).

This suggests that this set of parameters represents the dominant rogue wave generation processes in the form of

linear bandwidth-limited superposition with a directional correction ( $r, R$ ) and weakly nonlinear corrections ( $\varepsilon, \bar{D}, R$ ). This is consistent with other empirical studies such as Fedele et al. (2019), which consider the same parameters in conjunction with rogue crests during storms (except crest-trough correlation, which is not meaningful for crest heights).

This set of parameters is also similar to the ingredients to ECMWF’s freak wave forecast (ECMWF 2021), which is based on second and third-order bound and free waves and uses steepness, relative depth, directional spread, and spectral bandwidth. However, in our model these parameters are combined differently; a model enforcing the same interactions (steepness and relative depth for bound wave contribution, BFI and directionality index for free wave contribution) performs poorly, even in the deep-water regime where the BFI is most applicable (experiment 14).

Numerous previous studies have found the BFI to be a poor predictor of rogue wave risk in realistic sea states (Fedele et al. 2016, 2019; Gramstad and Trulsen 2007; Xiao et al. 2013; Häfner et al. 2021b) due to its strong underlying assumptions (such as unidirectionality). This study extends this to the fully nonparametric and nonlinear case, in which the predictive qualities of the BFI remain low, even when including interactions with directionality. This suggests that the contribution of the modulational instability and third-order free waves to rogue wave risk is negligible.

We study how our model uses different parameters by visualizing their impact on the prediction of the respective head of the neural network. For this, we make use of a functional decomposition called accumulated local effects (ALE, Apley and Zhu 2019), which measures the influence of infinitesimal changes in each parameter on the prediction outcome (see also Molnar 2020). This removes correlations between parameters, so we can plot the contribution of each parameter and 2-parameter interaction in isolation, where the total effect is the sum of all individual contributions. For example, a total value of ALE = 1 implies a predicted rogue wave probability that is about  $\exp(1) = 2.71$  times higher than baseline.

From the ALE plot (Fig. 4), we find that crest-trough correlation has by far the biggest influence and explains about 1 order of magnitude in rogue wave risk variation (as already observed in Häfner et al. 2021b). To first order, higher crest-trough correlation, lower directionality index, higher relative depth, and higher steepness lead to higher rogue wave risk, but parameter interactions can lead to more complicated, non-monotonic relationships (for example in very shallow water, see Section 4c).

Perhaps the most surprising finding is the strong interaction between crest-trough correlation and directionality index. This could imply that directional corrections are necessary for the accurate modelling of linear interactions, and could suggest a promising line for further theoretical research.

TABLE 2. Full list of experiments.  $\mathcal{L}$ : Prediction score (higher is better).  $\mathcal{E}$ : Invariance error (lower is better).  $\mathcal{C}$ : Calibration error (lower is better). Color coding ranges between (median – IQR, median + IQR) with inter-quartile range IQR.

ID	Feature groups			Scores		
	1	2	3	$\mathcal{L} \times 10^4$	$\mathcal{E} \times 10^2$	$\mathcal{C} \times 10^2$
1	$\{r\}$			4.62	8.52	6.90
2	$\{r, R\}$			5.05	8.58	3.86
3	$\{\varepsilon, \tilde{D}, R\}$			0.03	22.59	6.21
4	$\{r, \tilde{D}, R\}$			5.56	7.95	4.34
5	$\{r, \varepsilon, R\}$			5.49	8.83	3.83
6	$\{r, \varepsilon, \tilde{D}\}$			5.35	8.89	7.05
7	$\{r, R\}$	$\{\varepsilon, \tilde{D}\}$		5.77	9.19	4.46
8	$\{r, R, \text{Ur}\}$			5.70	7.99	3.94
9	$\{r, R\}$	$\{\text{Ur}, R\}$		5.64	7.49	4.31
10	$\{r, R, \text{BFI}\}$			5.60	7.75	4.51
11	$\{r, R\}$	$\{\text{BFI}, R\}$		5.46	8.20	4.44
12	$\{r\}$	$\{\varepsilon, \tilde{D}, R\}$		5.67	9.24	4.67
13	$\{\sigma_f\}$	$\{\varepsilon, \tilde{D}, R\}$		4.11	12.16	6.30
14	$\{r\}$	$\{\varepsilon, \tilde{D}\}$	$\{\text{BFI}, R\}$	5.64	9.77	6.02
15	$\{r, R\}$	$\{\varepsilon, \tilde{D}, \sigma_\theta\}$		6.22	10.63	5.20
16	$\{r, R\}$	$\{\varepsilon, \tilde{D}, R\}$		5.87	8.63	3.62
17	$\{r, \varepsilon, \tilde{D}, R\}$			5.98	8.60	2.96
18	$\{r\}$	$\{\varepsilon, \tilde{D}\}$	$\{\text{BFI}, \sigma_f, \sigma_\theta\}$	6.01	11.10	8.43
19	$\{r, \varepsilon, \tilde{D}, \sigma_\theta\}$			5.97	9.71	6.45
20	$\{r, \varepsilon, \tilde{D}, R, E_h\}$			6.10	9.14	5.33
21	$\{r, \varepsilon, \tilde{D}, \sigma_\theta, \nu\}$			6.31	10.04	4.00
22	$\{r, \varepsilon, \tilde{D}, R, \text{BFI}\}$			6.05	8.84	6.81
23	$\{r, \varepsilon, \tilde{D}, \sigma_\theta, \sigma_f, E_h, \text{BFI}, R\}$			6.91	12.69	3.68
24	$\{r, \varepsilon, \tilde{D}, \sigma_\theta, \sigma_f, E_h, H_s, \bar{T}, \kappa, \mu, \lambda_p\}$			6.70	56.44	7.27

#### Symbols

$r$	Crest-trough correlation	$\nu$	Spectral bandwidth (narrowness)
$\sigma_f$	Spectral bandwidth (peakedness)	$\sigma_\theta$	Directional spread
$\varepsilon$	Peak steepness $H_s k_p$	$R$	Directionality index $\sigma_\theta^2/(2\nu^2)$
BFI	Benjamin-Feir index	$\tilde{D}$	Relative peak water depth $Dk_p/(2\pi)$
$E_h$	Relative high-frequency energy	Ur	Ursell number
$\bar{T}$	Mean period	$\kappa$	Kurtosis
$\mu$	Skewness	$H_s$	Significant wave height

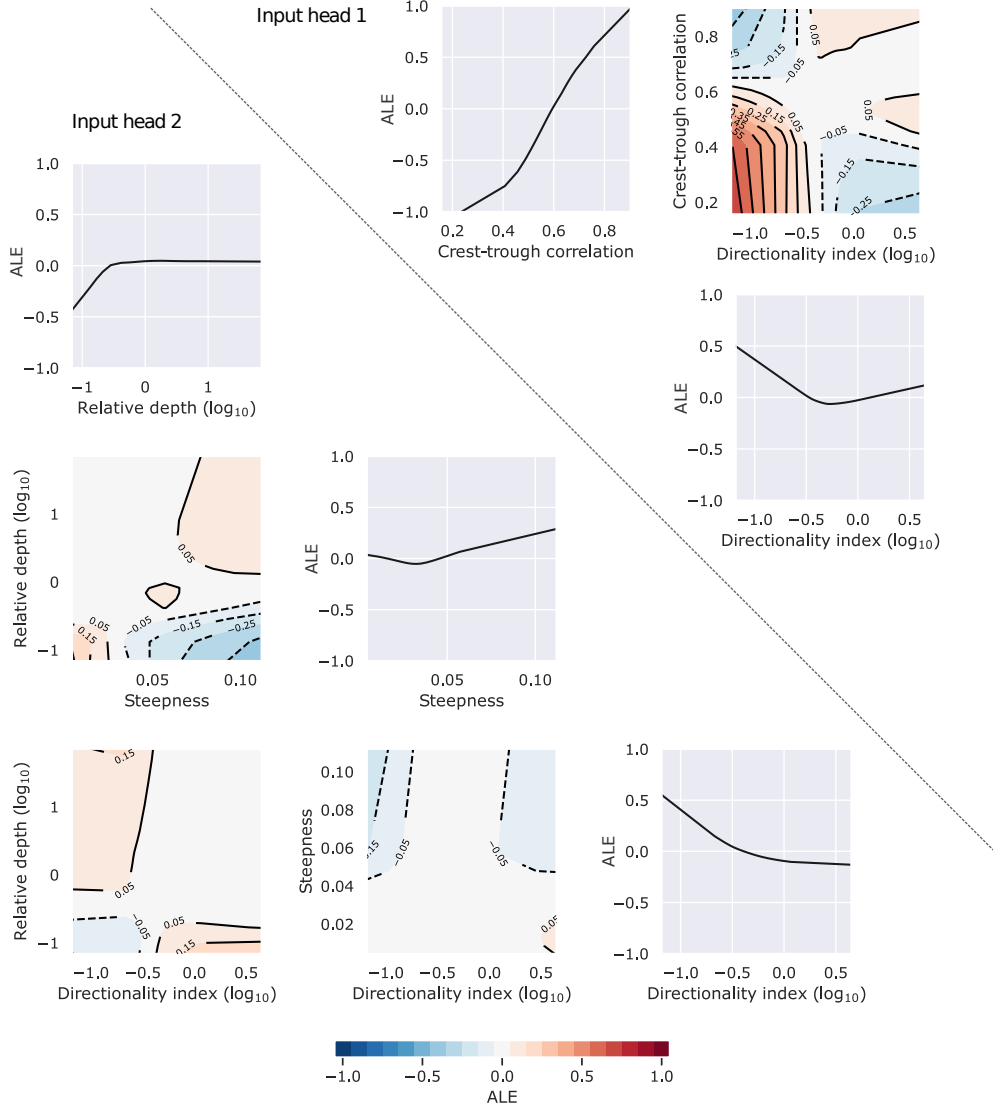


FIG. 4. ALE (accumulated local effects) plot matrix for experiment 16. Shown is the change in rogue wave risk (in logits) from the average as each parameter is varied. The total effect is the sum of all 1D, 2D, and higher-order contributions (not shown).

*b. The Rayleigh distribution is an upper bound for real-world rogue wave risk*

Despite the clear enhancing properties of weakly nonlinear corrections, the Rayleigh wave height distribution remains an upper bound for real-world (crest-to-trough) rogue waves. The Rayleigh distribution is the theoretical wave height distribution for linear narrowband waves (Longuet-Higgins 1952), i.e., the limit  $r \rightarrow 1$ ,  $\varepsilon \rightarrow 0$ ,  $\tilde{D} \rightarrow \infty$ , and  $R \rightarrow 0$ , and reads:

$$P(H/H_s > k) = \exp(-2k^2) \quad (10)$$

Only in the most extreme conditions does our model predict a similarly high probability, for example for  $R = 0.05$ ,  $\varepsilon = 0.01$ ,  $r = 0.85$ , and  $\tilde{D} = 0.2$ , which gives  $p = 3.2 \times 10^{-4}$  (compared to  $3.3 \times 10^{-4}$  for Rayleigh distributed waves).

In the opposite extreme, rogue wave probabilities can fall to as little as  $10^{-5}$  for low values of  $r$  and high values of  $R$  (as e.g. in a sea with a strong high-frequency component and high directional spread). This suggests that bandwidth effects can create sea states that efficiently suppress extremes — a fact that could lead to safer shipping routes should a forecast based on these parameters become available.



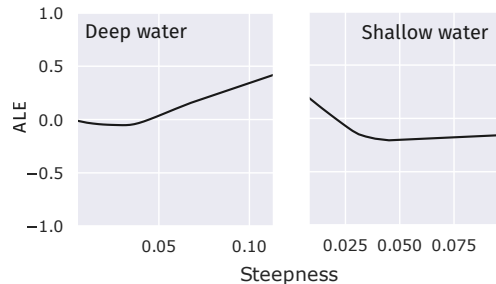


FIG. 5. Our model predicts a positive association between steepness and rogue waves in deep water, and a negative association in shallow water. Shown is the 1-dimensional ALE (accumulated local effects) plot in both cases. Here, deep water are sea states with  $\tilde{D} > 3$  and shallow water with  $\tilde{D} < 0.1$ .

*c. There is a clear separation between deep water and shallow water regimes*

The inclusion of an interaction between steepness and relative water depth is essential for predictive invariance across several environments. Looking at this more closely, we find that a stratification on deep and shallow water sea states reveals 2 distinct regimes (Fig. 5).

In deep water, rogue wave risk is strongly positively associated with steepness, as expected from the contribution of second and third-order nonlinear bound waves (Janssen 2018).

The opposite is true in shallow water, where we find a clear *negative* association with steepness. This is less expected, since nonlinear effects are typically considered to *increase* rogue wave occurrence. But the theoretical expression for the contribution of wave-induced currents does contain a term proportional to  $-\varepsilon \tilde{D}^{-1}$  (Janssen 2018) and could therefore be responsible for the observed negative scaling with  $\varepsilon$  in shallow waters where  $\tilde{D}$  is small. Another possible explanation is wave breaking, which also depends inversely on relative depth (e.g. Goda 2010). In very shallow waters, more sea states have a steepness close to the breaking threshold, which removes taller waves that tend to have a higher steepness than average.

## 5. Limitations

This analysis has some notable limitations due to the fact that we are only using wave buoy observations:

1. We do not have observations for all causal pathways. This includes wind-wave interactions, currents, and local topography (see Fig. 1). Fortunately, all 3 of these unobserved causal pathways relate to localized phenomena that are unlikely to play a major roll in bulk analysis. Nevertheless, this implies that local rogue wave probabilities could be dramatically different in regimes where these causal pathways play a major role,

e.g. over sloping topography (Trulsen et al. 2012) or in strong currents (Ying et al. 2011).

2. We only have 1-dimensional (time series) data. Because of that we cannot capture conditions where parameters are "imported" from elsewhere, e.g., a soliton generated in different conditions travelling into the observation area. While we expect most sea state parameters to be sufficiently stable in space for this to play a minor role, this is one mechanism through which we might underestimate the importance of nonlinear free waves.
3. Sensor bias. Systematic sensor bias is common in buoys, e.g. in the form of linearization of the sea state (McAllister and van den Bremer 2019) and can also lead to spurious causal relationships. For example, the fact that the best performing model uses both crest-trough correlation and the directionality index could also be because the directionality index can be used to correct noisy measurements of crest-trough correlation for this particular sensor.

If the only goal is maximized predictive performance this adaptation to sensor characteristics is actually a good thing, since several noisy quantities can be synthesized into more robust ones. On the other hand, this may obscure the true causal structure and generalize poorly to other sensors.

All of these factors can potentially reduce the capabilities of our model to detect relevant causal pathways, and may lead to an underestimation of the true rogue wave risk, even though we are confident that we accurately capture the leading-order dynamics of rogue wave generation.

Since our analysis is agnostic to the data source at hand (all that is needed is a mapping from parameters to observed wave heights), it can easily be repeated on different data sources as they become available to validate our findings.

## 6. Next steps

### *a. Comparisons to theory*

This study emphasizes the importance of bandwidth effects to predict rogue waves, while also including higher-order nonlinear effects that are governed by steepness and depend critically on relative water depth and directionality index. These parameters are frequently suggested by theory, but combined in a non-standard way to achieve greatly improved predictive performance and causal consistency. A logical next step would be to compare our predictions to existing theory, such as Fedele (2015); Janssen (2018), to study which parts of the prediction surface can be understood through it (and in turn quantify their predictive quality) and which parts are truly novel. We hope that this may ultimately lead to an improved theoretical understanding of real-world rogue waves.

### b. An improved rogue wave forecast

Another, perhaps even more obvious next step is the comparison of ECMWF’s operational freak wave forecast (ECMWF 2021). This operational forecast focuses on envelope wave heights and does not include crest-trough correlation or any other bandwidth parameter for linear wave interactions. Therefore we are confident that large improvements are within reach in terms of predicting crest-to-trough rogue waves.

However, the fact that our model is trained on observations may be problematic, since the model also learns to correct for systematic sensor bias (see Section 5). Also, having access to in-situ data on the sea state might lead to improved performance by itself. To correct for these effects, our model should be re-trained on *forecast* sea state parameters  $\mathbf{x}$  and observed labels  $\mathbf{y}$ . This way, both the current operational forecast and the empirical model would use the same input data, which allows for an apples-to-apples comparison.

### c. Predicting super-rogue waves

Observed wave height distributions often show a flattening of the wave height distribution towards the extreme tail (e.g. Gemmrich and Garrett 2011; Casas-Prat and Holthuijsen 2010), which Adcock and Taylor (2014) call a Type 3 distribution. Therefore, we expect rogue wave probabilities to be more pronounced for even more extreme waves (e.g. with  $H/H_s > 2.4$ ).

The lack of sufficient direct observations in these regimes calls for a different strategy. One approach could be to transform this classification problem (rogue wave or not) into a regression, where the predicted variables are the parameters of a candidate wave height probability distribution (e.g. shape and scale parameters of a Weibull distribution). Then, a similar analysis as in this study could be conducted for these parameters, which may reveal the main mechanisms influencing the risk for truly exceptional waves, and whether this flattening can be confirmed in our dataset.

**Acknowledgments.** Dion Häfner received funding from the Danish Hydrocarbon Research and Technology Centre (DHRTC).

Raw data were furnished by the Coastal Data Information Program (CDIP), Integrative Oceanography Division, operated by the Scripps Institution of Oceanography, under the sponsorship of the U.S. Army Corps of Engineers and the California Department of Parks and Recreation.

Computational resources were provided by DC<sup>3</sup>, the Danish Center for Climate Computing.

The authors thank Jonas Peters for his input on the initial design of this study.

This publication was made possible by the following open-source software stack: JAX (Bradbury et al. 2018),

flax (Heek et al. 2020), optax (Hessel et al. 2020), scikit-learn (Pedregosa et al. 2011), PyALE (Jomar 2020), NumPy (Harris et al. 2020), SciPy (Virtanen et al. 2020), matplotlib (Hunter 2007), Seaborn (Waskom 2021), pandas (Wes McKinney 2010), Jupyter (Kluyver et al. 2016).

**Data availability statement.** The preprocessed and aggregated version of the FOWD CDIP data used in this study is available for download at <https://erda.ku.dk/archives/7072a6eb3d181149deb56b7d8739805e/published-archive.html>.

## APPENDIX A

### Sea state parameters

Here, we give the definition of the sea state parameters central to this study. For a more thorough description of how parameters are computed from buoy displacement time series see Häfner et al. (2021a).

All parameters can be derived from the non-directional wave spectrum  $\mathcal{S}(f)$  (wave spectral density depending on frequency  $f$ ), with the exception of directional spread  $\sigma_\theta$ , which is estimated from the horizontal motion of the buoy and taken from the raw CDIP data.

Many parameters are computed from moments of the wave spectrum, where the  $n$ -th moment  $m_n$  is defined as

$$m_n = \int_0^\infty f^n \mathcal{S}(f) df \quad (\text{A1})$$

The expressions used for the relevant sea state parameters are:

#### Significant wave height:

$$H_s = 4\sqrt{m_0} \quad (\text{A2})$$

#### Spectral bandwidth (narrowness):

$$\nu_f = \sqrt{m_2 m_0 / m_1^2} - 1 \quad (\text{A3})$$

**Peak wavenumber**  $k_p$ , computed via the peak period (as in Young 1995):

$$\bar{T}_p = \frac{\int \mathcal{S}(f)^4 df}{\int f \mathcal{S}(f)^4 df} \quad (\text{A4})$$

The peak frequency  $f_p = 1/\bar{T}_p$  then leads to the peak wavenumber through the dispersion relation for linear waves in intermediate water:

$$f^2 = \frac{gk}{(2\pi)^2} \tanh(kD) \quad (\text{A5})$$

with gravitational acceleration  $g$  and water depth  $D$ . An approximate inverse is given in Fenton (1988).

**Relative depth:**

$$\tilde{D} = \frac{D}{\lambda} = \frac{1}{2\pi} k_p D \quad (\text{A6})$$

with wave length  $\lambda$ .

**Peak steepness:**

$$\varepsilon = H_s k_p \quad (\text{A7})$$

**Benjamin-Feir index:**

$$\text{BFI} = \frac{\varepsilon \nu}{\sigma_f} \sqrt{\max\{\beta/\alpha, 0\}} \quad (\text{A8})$$

where  $\sigma_f$  is spectral bandwidth estimated through peakedness and  $\nu, \alpha, \beta$  are coefficients depending only on  $\tilde{D}$  (full expression given in Serio et al. 2005).

**Directionality index:**

$$R = \frac{\sigma_\theta^2}{2\nu_f^2} \quad (\text{A9})$$

**Crest-trough correlation:**

$$r = \frac{1}{m_0} \sqrt{\rho^2 + \lambda^2} \quad (\text{A10})$$

$$\rho = \int_0^\infty S(\omega) \cos\left(\omega \frac{\bar{T}}{2}\right) d\omega \quad (\text{A11})$$

$$\lambda = \int_0^\infty S(\omega) \sin\left(\omega \frac{\bar{T}}{2}\right) d\omega \quad (\text{A12})$$

where  $\omega$  is the angular frequency and  $\bar{T} = m_0/m_1$  the spectral mean period (Tayfun and Fedele 2007).

## APPENDIX B

### Model implementation and hyperparameters

All performance critical model code is implemented in JAX (Bradbury et al. 2018), using neural network modules from flax (Heek et al. 2020) and optimizers from optax (Hessel et al. 2020). We run each experiment on a single Tesla P100 GPU in about 40 minutes, including SWAG sampling and re-training on every validation subset.

The hyperparameters for all experiments are shown in Table B3.

TABLE B3. Hyperparameters used in experiments.

### Hyperparameters

Optimizer	Adam
Learning rate	$10^{-4}$
Number of hidden layers	3
Neurons in hidden layers	(32, 16, 8)
$\ell_1$ penalty $\lambda_1$	0
$\ell_2$ penalty $\lambda_2$	$1 \times 10^{-5}$
Number of training epochs	50
Number of SWAG epochs	50
Number of SWAG posterior samples	100
Train-test split	66% – 34%

### References

- Adcock, T. A. A., and P. H. Taylor, 2014: The physics of anomalous (‘rogue’) ocean waves. *Reports on Progress in Physics*, **77** (10), 105 901, doi:10.1088/0034-4885/77/10/105901.
- Apley, D. W., and J. Zhu, 2019: Visualizing the Effects of Predictor Variables in Black Box Supervised Learning Models. *arXiv:1612.08468 [stat]*, arXiv: 1612.08468.
- Behrens, J., J. Thomas, E. Terrill, and R. Jensen, 2019: CDIP: Maintaining a Robust and Reliable Ocean Observing Buoy Network. *2019 IEEE/OES Twelfth Current, Waves and Turbulence Measurement (CWTM)*, 1–5, doi:10.1109/CWTM43797.2019.8955166.
- Boccotti, P., 2000: *Wave Mechanics for Ocean Engineering*. Elsevier, google-Books-ID: 1319kgDa8GUC.
- Bradbury, J., and Coauthors, 2018: JAX: composable transformations of Python+NumPy programs. URL <http://github.com/google/jax>.
- Casas-Prat, M., and L. H. Holthuijsen, 2010: Short-term statistics of waves observed in deep water. *Journal of Geophysical Research: Oceans*, **115** (C9), doi:10.1029/2009JC005742.
- Didenkulova, E., 2019: Catalogue of rogue waves occurred in the World Ocean from 2011 to 2018 reported by mass media sources. *Ocean & Coastal Management*, 105076, doi:10.1016/j.ocecoaman.2019.105076.
- Didenkulova, E. G., T. G. Talipova, and E. N. Pelinovsky, 2021: Rogue Waves in the Drake Passage: Unpredictable Hazard. *Antarctic Peninsula Region of the Southern Ocean: Oceanography and Ecology*, E. G. Morozov, M. V. Flint, and V. A. Spiridonov, Eds., Advances in Polar Ecology, Springer International Publishing, Cham, 101–114, doi:10.1007/978-3-030-78927-5\_7.
- Dudley, J. M., G. Genty, A. Mussot, A. Chabchoub, and F. Dias, 2019: Rogue waves and analogies in optics and oceanography. *Nature Reviews Physics*, **1** (11), 675–689, doi:10.1038/s42254-019-0100-0, number: 11 Publisher: Nature Publishing Group.
- ECMWF, 2021: Part VII: ECMWF Wave model. *IFS Documentation CY47R3*, IFS Documentation, ECMWF, URL <https://www.ecmwf.int/node/20201>.
- Fedele, F., 2015: On the kurtosis of deep-water gravity waves. *Journal of Fluid Mechanics*, **782**, 25–36, doi:10.1017/jfm.2015.538, publisher: Cambridge University Press.

- Fedele, F., J. Brennan, S. Ponce de León, J. Dudley, and F. Dias, 2016: Real world ocean rogue waves explained without the modulational instability. *Scientific Reports*, **6**, 27 715, doi:10.1038/srep27715.
- Fedele, F., J. Herterich, A. Tayfun, and F. Dias, 2019: Large nearshore storm waves off the Irish coast. *Scientific Reports*, **9** (1), 15 406, doi:10.1038/s41598-019-51706-8, number: 1 Publisher: Nature Publishing Group.
- Fedele, F., and M. A. Tayfun, 2009: On nonlinear wave groups and crest statistics. *Journal of Fluid Mechanics*, **620**, 221–239, doi:10.1017/S0022112008004424, URL <https://www.cambridge.org/core/journals/journal-of-fluid-mechanics/article/on-nonlinear-wave-groups-and-crest-statistics/CF946526383D92D4E12A0F950CE4FB1C>, publisher: Cambridge University Press.
- Fenton, J. D., 1988: The numerical solution of steady water wave problems. *Computers & Geosciences*, **14** (3), 357–368, doi:10.1016/0098-3004(88)90066-0.
- Gemmrich, J., and C. Garrett, 2011: Dynamical and statistical explanations of observed occurrence rates of rogue waves. *Natural Hazards and Earth System Science*, **11** (5), 1437–1446, doi:10.5194/nhess-11-1437-2011.
- Goda, Y., 2010: Reanalysis of Regular and Random Breaking Wave Statistics. *Coastal Engineering Journal*, **52** (1), 71–106, doi:10.1142/S0578563410002129, publisher: Taylor & Francis \_eprint: <https://doi.org/10.1142/S0578563410002129>.
- Gramstad, O., and K. Trulsen, 2007: Influence of crest and group length on the occurrence of freak waves. *Journal of Fluid Mechanics*, **582**, 463–472, doi:10.1017/S0022112007006507, publisher: Cambridge University Press.
- Harris, C. R., and Coauthors, 2020: Array programming with NumPy. *Nature*, **585** (7825), 357–362, doi:10.1038/s41586-020-2649-2.
- Heek, J., A. Levskaya, A. Oliver, M. Ritter, B. Rondepierre, A. Steiner, and M. van Zee, 2020: Flax: A neural network library and ecosystem for JAX. URL <http://github.com/google/flax>.
- Hessel, M., D. Budden, F. Viola, M. Rosca, E. Sezener, and T. Hennigan, 2020: Optax: composable gradient transformation and optimisation, in JAX! URL <http://github.com/deepmind/optax>.
- Hornik, K., 1991: Approximation capabilities of multilayer feedforward networks. *Neural Networks*, **4** (2), 251–257, doi:10.1016/0893-6080(91)90009-T.
- Hunter, J. D., 2007: Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, **9** (3), 90–95, doi:10.1109/MCSE.2007.55.
- Häfner, D., J. Gemmrich, and M. Jochum, 2021a: FOWD: A Free Ocean Wave Dataset for Data Mining and Machine Learning. *Journal of Atmospheric and Oceanic Technology*, **-1** (aop), doi:10.1175/JTECH-D-20-0185.1, publisher: American Meteorological Society Section: Journal of Atmospheric and Oceanic Technology.
- Häfner, D., J. Gemmrich, and M. Jochum, 2021b: Real-world rogue wave probabilities. *Scientific Reports*, **11** (1), 10 084, doi:10.1038/s41598-021-89359-1, number: 1 Publisher: Nature Publishing Group.
- Janssen, P., 2018: Shallow-water version of the Freak Wave Warning System. Technical memorandum 813, ECMWF. URL <https://www.ecmwf.int/en/elibrary/18063-shallow-water-version-freak-wave-warning-system>.
- Janssen, P. A. E. M., 2003: Nonlinear Four-Wave Interactions and Freak Waves. *Journal of Physical Oceanography*, **33** (4), 863–884, doi:10.1175/1520-0485(2003)33<863:NFIAPW>2.0.CO;2, publisher: American Meteorological Society.
- Jomar, D., 2020: PyALE: A Python implementation of accumulated local effect plots. URL <https://github.com/DanaJomar/PyALE>.
- Kluyver, T., and Coauthors, 2016: Jupyter notebooks - a publishing format for reproducible computational workflows. *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, F. Loizides, and B. Schmidt, Eds., IOS Press, Netherlands, 87–90.
- Korteweg, D. J., and G. De Vries, 1895: On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, **39** (240), 422–443.
- Longuet-Higgins, M. S., 1952: On the statistical distribution of the height of sea waves. *JMR*, **11**, 245–266.
- Maddox, W., T. Garipov, P. Izmailov, D. Vetrov, and A. G. Wilson, 2019: A Simple Baseline for Bayesian Uncertainty in Deep Learning. *arXiv:1902.02476 [cs, stat]*, arXiv: 1902.02476.
- Mallory, J. K., 1974: Abnormal Waves on the South East Coast of South Africa. *The International Hydrographic Review*, URL <https://journals.lib.unb.ca/index.php/ihr/article/view/23802>.
- McAllister, M. L., S. Draycott, T. a. A. Adcock, P. H. Taylor, and T. S. v. d. Bremer, 2019: Laboratory recreation of the Draupner wave and the role of breaking in crossing seas. *Journal of Fluid Mechanics*, **860**, 767–786, doi:10.1017/jfm.2018.886.
- McAllister, M. L., and T. S. van den Bremer, 2019: Lagrangian Measurement of Steep Directionally Spread Ocean Waves: Second-Order Motion of a Wave-Following Measurement Buoy. *Journal of Physical Oceanography*, **49** (12), 3087–3108, doi:10.1175/JPO-D-19-0170.1, publisher: American Meteorological Society.
- Miche, M., 1944: Mouvements ondulatoires de la mer en profondeur constante ou décroissante. *Annales de Ponts et Chaussées*, 1944, pp(1) 26-78, (2) 270-292, (3) 369-406, publisher: École nationale des ponts et chaussées.
- Molnar, C., 2020: *Interpretable Machine Learning*. URL <https://christophm.github.io/interpretable-ml-book/>.
- Mori, N., and P. A. E. M. Janssen, 2006: On Kurtosis and Occurrence Probability of Freak Waves. *Journal of Physical Oceanography*, **36** (7), 1471–1483, doi:10.1175/JPO2922.1, URL <https://journals.ametsoc.org/jpo/article/36/7/1471/10720/On-Kurtosis-and-Occurrence-Probability-of-Freak>, publisher: American Meteorological Society.
- Nair, V., and G. E. Hinton, 2010: Rectified linear units improve restricted boltzmann machines. *Proceedings of the 27th International Conference on International Conference on Machine Learning*, Omnipress, Madison, WI, USA, 807–814, ICML’10.
- Onorato, M., A. R. Osborne, M. Serio, L. Cavaleri, C. Brandini, and C. T. Stansberg, 2006: Extreme waves, modulational instability and second order theory: wave flume experiments on irregular waves. *European Journal of Mechanics - B/Fluids*, **25** (5), 586–601, doi:10.1016/j.euromechflu.2006.01.002.
- Pedregosa, F., and Coauthors, 2011: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, **12**, 2825–2830.

- Peters, J., P. Bühlmann, and N. Meinshausen, 2016: Causal inference by using invariant prediction: identification and confidence intervals. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **78** (5), 947–1012, doi:10.1111/rssb.12167, \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/rssb.12167>.
- Peters, J., D. Janzing, and B. Schölkopf, 2017: *Elements of Causal Inference: Foundations and Learning Algorithms*. Adaptive Computation and Machine Learning series, MIT Press, Cambridge, MA, USA.
- Serio, M., M. Onorato, A. R. Osborne, and P. Janssen, 2005: On the computation of the Benjamin-Feir Index. *Nuovo Cimento della Societa Italiana di Fisica C*, **28**, 893–903, doi:10.1393/ncc/i2005-10134-1.
- Stansell, P., 2004: Distributions of freak wave heights measured in the North Sea. *Applied Ocean Research*, **26** (1), 35–48, doi:10.1016/j.apor.2004.01.004, URL <http://www.sciencedirect.com/science/article/pii/S0141118704000379>.
- Tayfun, M. A., and F. Fedele, 2007: Wave-height distributions and nonlinear effects. *Ocean Engineering*, **34** (11), 1631–1649, doi:10.1016/j.oceaneng.2006.11.006.
- Trulsen, K., H. Zeng, and O. Gramstad, 2012: Laboratory evidence of freak waves provoked by non-uniform bathymetry. *Physics of Fluids*, **24** (9), 097 101, doi:10.1063/1.4748346, publisher: American Institute of Physics.
- Ursell, F., 1953: The long-wave paradox in the theory of gravity waves. *Mathematical Proceedings of the Cambridge Philosophical Society*, **49** (4), 685–694, doi:10.1017/S0305004100028887, URL <https://www.cambridge.org/core/journals/mathematical-proceedings-of-the-cambridge-philosophical-society/article/longwave-paradox-in-the-theory-of-gravity-waves/5A178FB13BD7B9C3314A49A495A860DB>, publisher: Cambridge University Press.
- Virtanen, P., and Coauthors, 2020: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, **17**, 261–272, doi:10.1038/s41592-019-0686-2.
- Waskom, M. L., 2021: seaborn: statistical data visualization. *Journal of Open Source Software*, **6** (60), 3021, doi:10.21105/joss.03021.
- Wes McKinney, 2010: Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference*, Stéfan van der Walt, and Jarrod Millman, Eds., 56 – 61, doi:10.25080/Majora-92bf1922-00a.
- Xiao, W., Y. Liu, G. Wu, and D. K. P. Yue, 2013: Rogue wave occurrence and dynamics by direct simulations of nonlinear wave-field evolution. *Journal of Fluid Mechanics*, **720**, 357–392, doi:10.1017/jfm.2013.37, publisher: Cambridge University Press.
- Ying, L. H., Z. Zhuang, E. J. Heller, and L. Kaplan, 2011: Linear and nonlinear rogue wave statistics in the presence of random currents. *Nonlinearity*, **24** (11), R67–R87, doi:10.1088/0951-7715/24/11/R01, URL <https://doi.org/10.1088/0951-7715/24/11/r01>, publisher: IOP Publishing.
- Young, I. R., 1995: The determination of confidence limits associated with estimates of the spectral peak frequency. *Ocean Engineering*, **22** (7), 669–686, doi:10.1016/0029-8018(95)00002-3.