

# FREAK WAVES II

Studie zum Auftreten von Extremwellen in der südlichen Nordsee

Abschlussbericht des Forschungs- und Entwicklungsprojekts



*basierend auf dem Kooperationsvertrag vom 01.02.2021 zwischen dem  
Bundesamt für Seeschifffahrt und Hydrographie  
und der  
Helmholtz-Zentrum hereon GmbH*

Projektlaufzeit: 01.02.2021 - 31.12.2024

## Inhaltsverzeichnis

Abbildungsverzeichnis .....	2
Tabellenverzeichnis .....	4
1 Einleitung .....	5
2 Verwendete Daten .....	6
3 Ergebnisse.....	7
3.1 Rogue Waves in mittleren Wassertiefen und der Beitrag der Modulations-Instabilität..	7
3.2 Rogue Waves in Flachwasser und die Rolle von Solitonen.....	13
3.3 Einfluss von Tideströmungen auf die Statistik von Rogue Waves.....	18
3.4 Vergleich und Eignung von Ansätzen des maschinellen Lernens.....	19
4 Zusammenfassung .....	25
5 Ausblick.....	26
6 Literaturverzeichnis .....	28

## Abbildungsverzeichnis

<b>Abb. 1:</b> Position der Bojenstationen in der südlichen Nordsee. ....	7
<b>Abb. 2:</b> Parameter $\varepsilon$ , welcher die spektrale Bandbreite beschreibt, für alle Zeitreihen an den Tiefwasser-Bojen. Aufgeteilt nach Zeitreihenkategorie. Box: Interquartilsabstand; Whiskers: 1.5x Interquartilsabstand; horizontale Linie in der Box: Median; Punkte: Ausreißer. ....	9
<b>Abb. 3:</b> Vergleich der Verteilungen von Parameter $\varepsilon$ aus Extremwellenzeitreihen und Zeitreihen ohne Extremwelle. a) Zeitreihen nach dem Höhenkriterium bzw. extreme Extremwellen; b) Zeitreihen nach dem Wellenberg-Kriterium.....	9
<b>Abb. 4:</b> Ergebnis der Monte-Carlo-Simulationen für a) 2510 und b) 485 Verteilungen des Parameters $\varepsilon$ . Blau: Mittelwert der Verteilungen; lila: Bereich innerhalb von 0.5 Standardverteilungen; rosa: Bereich innerhalb einer Standardverteilung; orange: Bereich innerhalb von zwei Standardverteilungen. ....	10
<b>Abb. 5:</b> Ergebnis der Monte-Carlo-Simulationen für a) 2510 und b) 485 Verteilungen der Richtungsstreuung $\sigma_\theta$ . Blau: Mittelwert der Verteilungen; lila: Bereich innerhalb von 0.5 Standardverteilungen; rosa: Bereich innerhalb einer Standardverteilung; orange: Bereich innerhalb von zwei Standardverteilungen. ....	11
<b>Abb. 6:</b> Paarweise Darstellung von spektraler Bandbreite $\varepsilon$ bzw. $v$ und Richtungsstreuung. Vergleich von Zeitreihen mit und ohne Extremwelle. Farbcodierung: Dichte der Datenpunkte „normaler“ Extremwellensamples. Schwarze Punkte: Werte aus Extremwellenzeitreihen. Diese beinhalten entweder „extreme“ Extremwellen oder Extremwellen nach dem Wellenbergkriterium, siehe jeweilige Legende.....	12
<b>Abb. 7:</b> Beispiel einer Messzeitreihe mit einer Extremwelle nach ungefähr 820 s (oben) und das zugehörige Soliton Spektrum basierend auf der vKdV-NLFT (unten). Die signifikante Wellenhöhe zum Zeitpunkt der Messung betrug 2,71 m, die Höhe der Extremwelle betrug 7,00 m. Die Zeitreihe wurde am 17.10.2013 an der Boje SEE gemessen. ....	14
<b>Abb. 8:</b> Beispiel einer Messzeitreihe ohne Extremwelle (oben) und das zugehörige Soliton Spektrum basierend auf der vKdV-NLFT (unten). Die signifikante Wellenhöhe zum Zeitpunkt der Messung betrug 2,46 m, die Höhe der größten Welle im Zeitfenster betrug 4,44 m. Die Zeitreihe wurde am 26.12.2016 an der Boje SEE gemessen. ....	15
<b>Abb. 9:</b> (a) Zeitreihe mit Extremwelle gemessen am 17.10.2023; (b) Vergrößerung des Zeitabschnittes mit der Extremwelle und Darstellung der reduzierten Zeitreihe (rot); (c) Soliton Spektren der Original- (blau) und der modifizierten Zeitreihe (rot). ....	16
<b>Abb. 10:</b> Box-Whiskerplots des Verhältnisses der Amplituden des zweitgrößten und des größten Solitons im Spektrum für Fälle mit und ohne Extremwellen. Bei Fällen mit Extremwellen wird weiterhin unterschieden in Extremwellen Definitionen nach verschiedenen Kriterien. ....	17
<b>Abb. 11:</b> Mittlere Werte (rote Linie) und zugehöriger Bereich von +/- einer Standardabweichung (rot schraffierter Bereich) verschiedener Seegangparameter (y-Achse) in Abhängigkeit von der relativen Strömungsgeschwindigkeit an der Station Fino 1.....	18
<b>Abb. 12:</b> (Links) Verstärkung der Überschreitens-Wahrscheinlichkeit verschiedener Grenzwerte (Farben) für die Definition von Rogue Waves in Abhängigkeit von der relativen Strömungsgeschwindigkeit an der Station Fino 1. (Rechts) Überschreitens-Wahrscheinlichkeiten bei minimaler (rot) und maximaler (blau) relativer Strömungsgeschwindigkeit im Vergleich zur Rayleigh Verteilung. ....	19

**Abb. 13:** Vergleich zwischen beobachteten Abnormitätsindex (x-Achse) und durch das RF-Modell (links) bzw. das NN-Modell (rechts) vorhergesagten Abnormitätsindex (y-Achse) für die nicht zum Training der Modelle verwendeten Testdaten.....23

**Abb. 14:** Untergruppen von Instanzen, die im Random-Forest-Modell ähnlichen Entscheidungswegen folgen und Merkmale mit einem p-Wert von weniger als 0,01 aufweisen. Ein Zielwert (target) von 1 weist auf eine potenzielle Rogue Wave hin, bei der der Abnormitätsindex gleich oder größer als 2 ist.....24

**Abb. 15:** Globale Merkmals-Bedeutung basierend auf SHAP. ....24

**Abb. 16:** Lokale Merkmalsbedeutung für die ausgewählten Schlüsselmerkmale unter Verwendung von SHAP für das reduzierte neuronale Netzwerk. Der Balken stellen die Verteilung der normalisierten Merkmalswerte dar, wobei hohe/niedrige Merkmalswerte in rot/blau dargestellt sind.....25

## Tabellenverzeichnis

<b>Tab. 1:</b> Eingabe- (Features) und Zielparameter (Target) für die verwendeten Algorithmen des maschinellen Lernens. Die Parameter beziehen sich jeweils auf die 30 Minuten, die der jeweiligen Welle vorausgehen. ....	20
<b>Tab. 2:</b> Konfusions-Matrix für das Random Forrest Modell für die Testdaten, die nicht zum Training des Modells verwendet wurden. ....	22

# 1 Einleitung

Seegang besteht aus einem Spektrum von Wellen verschiedener Höhen, Frequenzen und Richtungen, das durch charakteristische Größen wie z.B. die signifikante Wellenhöhe charakterisiert wird. Wellen die höher sind, als in einem bestimmten Seegang typischerweise erwartet, werden gemeinhin als „Rogue Waves“ (früher oft auch „Freak Waves“) bezeichnet [Bitner-Gregersen and Gramstad, 2016]. Was in diesem Zusammenhang „erwartet“ genau bedeutet ist eine Frage der Definition, die in der Literatur nicht einheitlich behandelt wird [z.B. Dysthe et al., 2008]. Ein gängiger Ansatz besteht darin, Rogue Waves als Wellen zu definieren, deren Höhe das Doppelte der aktuellen signifikanten Wellenhöhe überschreitet. Dabei bezieht sich die signifikante Wellenhöhe auf die durchschnittliche Höhe des höchsten Drittels der Wellen in einer Aufzeichnung.

Für einen gegebenen Seegangszustand, in dem sich charakteristische Größen wie die signifikante Wellenhöhe über einen gewissen Zeitraum nicht ändern, können die Häufigkeiten mit denen Wellen einer bestimmten Höhe auftreten durch bekannte Verteilungen wie etwa die Rayleigh oder die Forristall Verteilung beschrieben werden [Forristall, 1978]. Anhand dieser Verteilungen lässt sich auch abschätzen, mit welcher Häufigkeit sehr hohe Wellen wie z.B. Rogue Waves zu erwarten sind, wenn keine zusätzlichen (nichtlinearen) Mechanismen zu ihrer Entstehung beitragen.

Im Vorläufer Projekt FREAK WAVES I („Studie zum Auftreten von Extremwellen in der Nordsee im Klimawandel“) wurde ein umfangreicher Datensatz erstellt und ausgewertet. Ein zentrales Ergebnis war, dass an der Mehrzahl der untersuchten Positionen die Häufigkeit der beobachteten Extremwellen in guter Übereinstimmung mit einer Rayleigh oder einer Forristall Verteilung war [Teutsch et al., 2020]. Ausnahmen stellten einige Flachwasserstationen dar, an denen im Vergleich mit der Forristall Verteilung eine erhöhte Zahl von Extremwellen beobachtet wurde. Weiterhin wurden an mehreren Stationen Abweichungen im Bereich extremer Einzelwellen festgestellt, deren Höhe die signifikante Wellenhöhe um mehr als das 2.3-fache überstieg.

Im Projekt FREAK WAVES II wurde deshalb untersucht, inwieweit nichtlineare Prozesse zur Entstehung solcher Einzelwellen beigetragen haben können. Diese Prozesse unterscheiden sich in tiefem und in flachem Wasser.

In FREAK WAVES II wurden zunächst mögliche Beiträge der sogenannten Modulations-Instabilität („modulational instability“) bei der Bildung von Extremwellen in mittleren bis größeren Wassertiefen in der südlichen Nordsee untersucht. Die Modulations-Instabilität ist ein nichtlinearer Mechanismus zur Fokussierung der Wellenenergie, der mathematisch nachgewiesen und in Laborexperimenten bestätigt wurde. Es ist jedoch noch unklar, ob dieser Mechanismus für die Entstehung von Rogue Waves unter realen Meeresbedingungen verantwortlich ist. Eine Modulations-Instabilität tritt vor allem dann auf, wenn die Wellen ähnliche Frequenzen und Richtungen haben. An den in FREAK WAVES II untersuchten Stationen waren diese Bedingungen nicht erfüllt. Dieser Befund führte zu der Erkenntnis, dass die Modulations-Instabilität nicht der wahrscheinlichste Mechanismus zur Erzeugung von Rogue Waves in der südlichen Nordsee ist.

**Die Ergebnisse sind im Detail in Abschnitt 3.1 dieses Berichts dargestellt und führten zu folgender Veröffentlichung:**

*Teutsch, Ina; Weisse, Ralf (2023): Rogue Waves in the Southern North Sea—The Role of Modulational Instability. In: J. Phys. Oceanogr. 53 (1), S. 269–286. DOI: 10.1175/JPO-D-22-0059.1.*

In einem weiteren Schritt wurden die Beiträge möglicher nichtlinearer Prozesse zur Entstehung von Rogue Waves in flachem Wasser untersucht. Dabei wurde insbesondere die Rolle von Solitonen untersucht. Es konnte gezeigt werden, dass in der Mehrzahl der Fälle, in denen Rogue Waves auftraten, eine bestimmte Form des Soliton Spektrums beobachtet werden konnte.

**Die Ergebnisse sind im Detail in Abschnitt 3.2 dieses Berichts dargestellt und führten zu folgenden zwei Veröffentlichungen:**

*Teutsch, Ina; Brühl, Markus; Weisse, Ralf; Wahls, Sander (2023): Contribution of solitons to enhanced rogue wave occurrence in shallow depths. A case study in the southern North Sea. In: Nat. Hazards Earth Syst. Sci. 23 (6), S. 2053–2073. DOI: 10.5194/nhess-23-2053-2023.*

*Teutsch, Ina; Weisse, Ralf; Wahls, Sander (2024): Brief communication. Implications of outstanding solitons for the occurrence of rogue waves at two additional sites in the North Sea. In: Nat. Hazards Earth Syst. Sci. 24 (6), S. 2065–2069. DOI: 10.5194/nhess-24-2065-2024.*

Anschließend wurde der Einfluss der Strömung auf die Statistiken von Seegangparametern und von Rogue Waves untersucht. Dabei wurde festgestellt, dass in den überwiegend vorhandenen Zuständen mit starker Richtungsstreuung eine eher symmetrische Modulation erfolgte.

**Die Ergebnisse sind im Detail in Abschnitt 3.3 dieses Berichts dargestellt und führten zu folgenden zwei Publikation, die sich zurzeit im Begutachtungsprozess befinden:**

*Teutsch, Ina; Mendes, Saulo; Kasparian, Jérôme (2024): Direction Symmetry of Wave Field Modulation by Tidal Current. In: J. Phys. Oceanogr., [Under Review, Preprint, arXiv:2407.04723]. DOI: 10.48550/arXiv.2407.04723.*

*Mendes, Saulo; Teutsch, Ina; Kasparian, Jérôme (2024): Symmetric influence of forward and opposing tidal currents on rogue wave statistics. In: J. Fluid Mech., [Under Review, Preprint, arXiv:2408.06371]. DOI: 10.48550/ARXIV.2408.06371.*

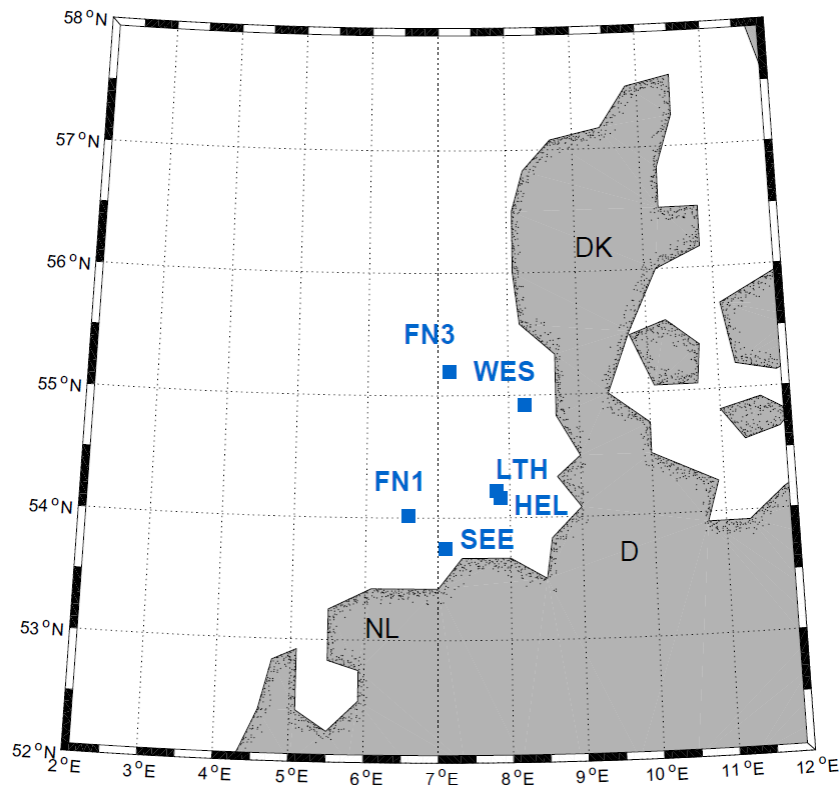
Neben einem verbesserten Verständnis der Entstehung von Rogue Waves in der südlichen Nordsee war ein Ziel des Vorhabens zu untersuchen, inwieweit Kurzzeitprognosen möglich erscheinen. Aufbauend auf den Erkenntnissen der oben beschriebenen Arbeiten wurden deshalb verschiedene Ansätze aus dem Bereich des maschinellen Lernens verwendet, um zu testen, inwieweit es möglich ist Rogue Waves automatisch zu detektieren, relevante Parameter (Features) zu identifizieren und ggfs. Vorhersagen zu machen.

**Die Ergebnisse sind in Abschnitt 3.4 dieses Berichts dargestellt. Eine zugehörige Publikation befindet sich derzeit in Vorbereitung.**

## 2 Verwendete Daten

In FREAK WAVES II wurde der im Vorgängerprojekt FREAK WAVES I erstellte und qualitätskontrollierte Datensatz von Seegangsmessungen in der südlichen Nordsee verwendet. Dieser besteht aus Rohdaten und abgeleiteten Parametern von sechs Messbojen, die innerhalb

des gemeinsamen Zeitraums 2011-2016 vorlagen. Die Lage der Messbojen ist in **Abb. 1** dargestellt. Zusätzlich standen für die Arbeiten Daten an verschiedenen Radarstationen zur Verfügung. Der Datensatz und die daraus abgeleiteten Rogue Wave Statistiken sind ausführlich im Abschlussbericht des Projekts FREAK WAVES I sowie in *Teutsch et al. [2020]* dargestellt.



**Abb. 1:** Position der Bojenstationen in der südlichen Nordsee.

## 3 Ergebnisse

### 3.1 Rogue Waves in mittleren Wassertiefen und der Beitrag der Modulations-Instabilität

Ein in der Literatur viel diskutierter Mechanismus, der die Entstehung von Rogue Waves erklären kann, ist die Benjamin-Feir-Instabilität [*Benjamin and Feir, 1967*], bei der sich aufgrund einer modulierenden Instabilität ein gleichmäßiger Wellenzug in tiefem Wasser in Gruppen auflöst, die anschließend eine Welle erzeugen, die auf Kosten der umliegenden Wellen wächst. Der Mechanismus wurde sowohl theoretisch als auch in Wellentankversuchen nachgewiesen [*Lake et al., 1977; Onorato et al., 2004*] und u.a. von *Janssen [2003]* als mögliche Ursache für die Entstehung von Rogue Waves im offenen Ozean vorgeschlagen.

Der Beitrag des Mechanismus zur Entstehung von Rogue Waves im realen Ozean wird nach wie vor kontrovers diskutiert. Eine Ursache ist, dass als Voraussetzung für die Entwicklung einer Benjamin-Feir-Instabilität ein schmales Spektrum sowohl im Frequenz- als auch im

Richtungsraum erforderlich ist. Dabei bedeutet ein enges Spektrum im Frequenzraum, dass sich die Seegangenergie in einem schmalen Bereich nahe der Peakfrequenz konzentriert. Über die Frequenzbandbreite hinaus besteht unter Experten Konsens darüber, dass die Wirkung der modulierenden Instabilität stark von der Ausbreitungsrichtung eines Wellenfeldes abhängt, und zwar in dem Sinne, dass sie bei unidirektionalen Wellen (schmales Richtungsspektrum) stärker ist als bei kurzkämmigen Wellen (breites Richtungsspektrum) [Gramstad and Trulsen, 2007; Janssen and Bidlot, 2009]. Da diese Bedingungen im offenen Ozean oft nicht gegeben sind, wird der mögliche Beitrag der Benjamin-Feir-Instabilität zur Entstehung von Rogue Waves in der Natur von verschiedenen Autoren unterschiedlich eingeschätzt.

Während die ursprünglichen Arbeiten sich auf Tiefwasser bezogen, wurde der mögliche Tiefenbereich für das Auftreten des Mechanismus mittlerweile auf mittlere Wassertiefen ausgedehnt [Karpadakis et al., 2019]. Aus diesem Grund wurde untersucht, welche Rolle die Benjamin-Feir-Instabilität bei der Entstehung von Rogue Waves an den in diesem Projekt untersuchten Positionen in mittleren Wassertiefen spielen könnte.

Unidirektionale Wellen können oberhalb einer Schwelle von  $kh = 1.363$  eine Benjamin-Feir-Instabilität entwickeln, wobei  $k$  für die Wellenzahl und  $h$  für die Wassertiefe stehen [Benjamin and Feir, 1967]. Unter der Annahme von Tiefwasserwellen lässt sich daraus eine Bedingung für die Peakperiode  $T_p$  ableiten

$$T_p < \sqrt{\frac{\pi h}{0.11g}} \quad (1)$$

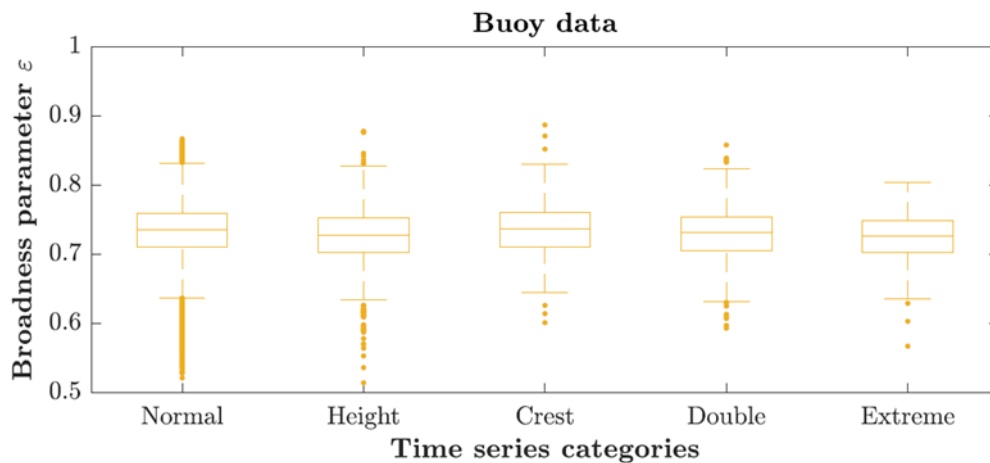
wobei  $g$  die Gravitationsbeschleunigung repräsentiert. Aus dem zur Verfügung stehenden Datensatz wurden deshalb zunächst die Bojen selektiert, an denen sich bei hohen signifikanten Wellenhöhen ( $H_s$  größer als das lokale langjährige 70%-Perzentil) potenziell Modulations-Instabilitäten erwarten lassen, d.h. deren Peakperioden in 99% aller Fälle unterhalb des aus Gleichung (1) ermittelten Grenzwertes lagen [Teutsch and Weisse, 2023].

Für diese Daten wurden anschließend Parameter zur Charakterisierung der spektralen Bandbreite ermittelt. Diese umfassen den „Broadness Parameter“  $\epsilon$ , den „Narrowness Parameter“  $\nu$ , Goda's Peakness Parameter  $Q_p$  und den Benjamin Feir Index  $BFI$  (Details s. Teutsch and Weisse, 2023). Anschließend wurde jede 30-minütige Stichprobe einer der folgenden Kategorien zugeordnet:

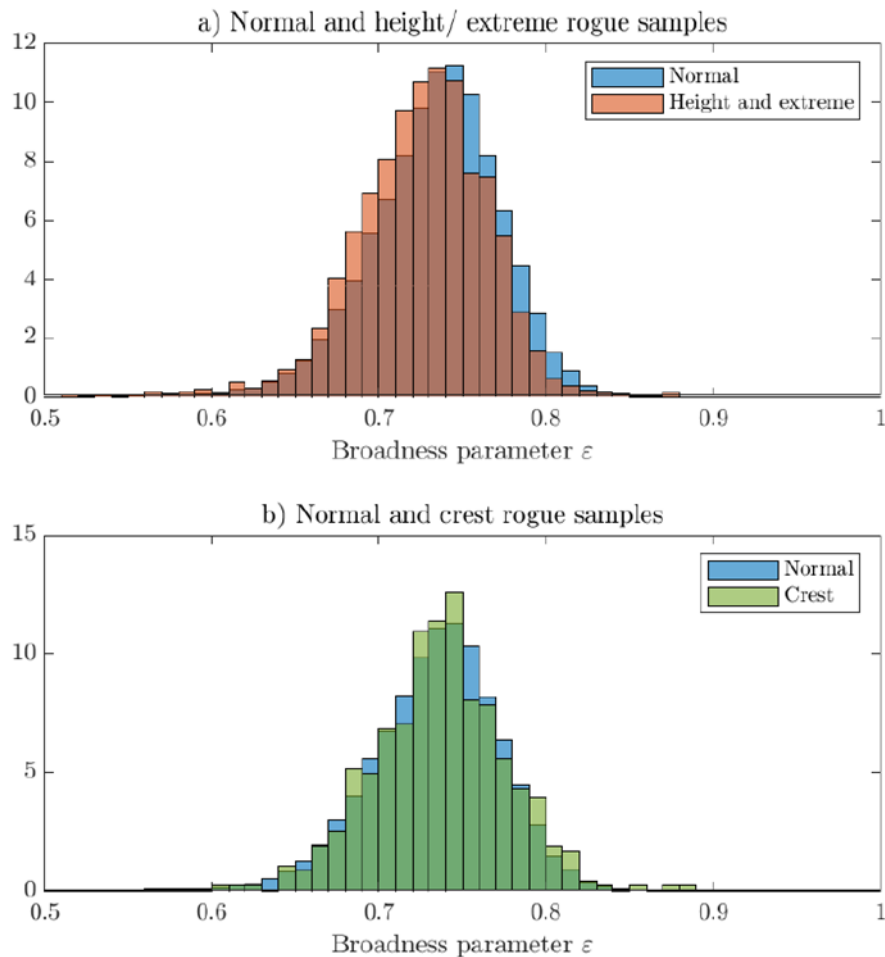
- **Normal:** alle Zeitreihen ohne Extremwelle (Rogue Wave)
- **Height:** alle Zeitreihen mit mindestens einer Extremwelle ( $2.0 \leq H/H_s < 2.3$ )
- **Extreme:** alle Zeitreihen mit mindestens einer Extremwelle ( $H/H_s \geq 2.3$ )
- **Crest:** alle Zeitreihen mit mindestens einer Extremwelle ( $C \geq 1.25 H_s$ )
- **Double:** alle Zeitreihen mit Extremwellen nach zwei Kriterien

Als Ergebnis umfasste die Kategorie Normal etwa 94% aller Zeitreihen sowie Height/Extreme/Crest/Double etwa 3.6/0.3/0.8/1.3% aller Zeitreihen.

**Abb. 2** zeigt beispielhaft die Verteilung des Broadness Parameters  $\epsilon$  für die unterschiedlichen Kategorien. Es ist zu erkennen, dass die Breite der Spektren in allen Kategorien stark streut und sich die Verteilungen auf den ersten Blick nicht auffällig voneinander unterscheiden. Beim Blick auf die Histogramme (**Abb. 3**) lassen sich leichte Unterschiede erkennen, so dass die spektrale Bandbreite im Mittel in der Kategorie der Height/Extreme Fälle im Vergleich zur Referenz (Normal) etwas geringer, bei der Kategorie der Crest Fälle dagegen etwas erhöht zu sein scheint.

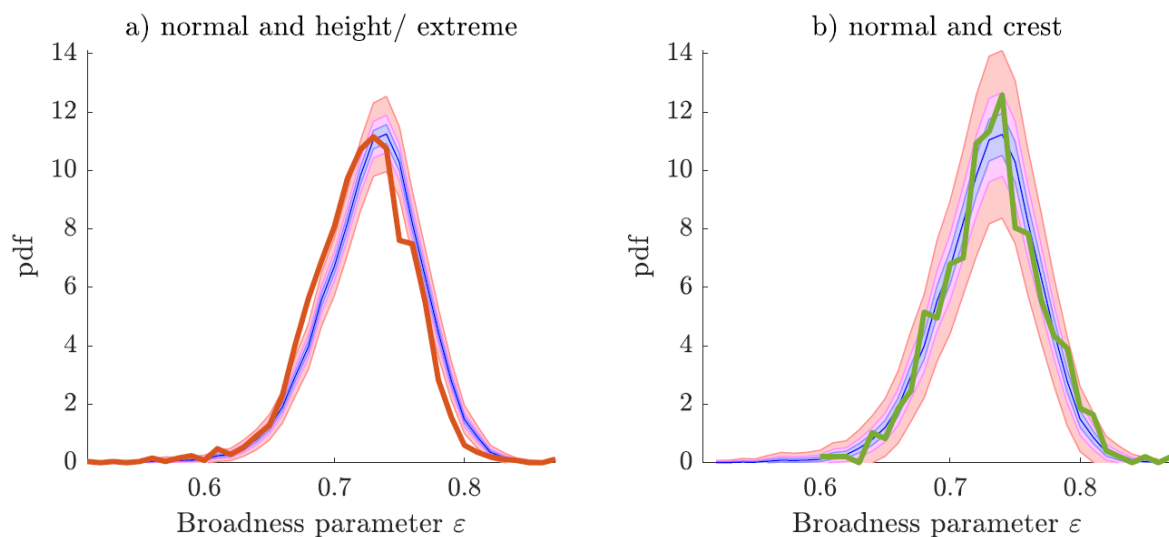


**Abb. 2:** Parameter  $\varepsilon$ , welcher die spektrale Bandbreite beschreibt, für alle Zeitreihen an den Tiefwasser-Bojen. Aufgeteilt nach Zeitreihenkategorie. Box: Interquartilsabstand; Whiskers: 1.5x Interquartilsabstand; horizontale Linie in der Box: Median; Punkte: Ausreißer.



**Abb. 3:** Vergleich der Verteilungen von Parameter  $\varepsilon$  aus Extremwellenzeitreihen und Zeitreihen ohne Extremwelle. a) Zeitreihen nach dem Höhenkriterium bzw. extreme Extremwellen; b) Zeitreihen nach dem Wellenberg-Kriterium.

Bei der Interpretation der Verteilungen ist zu beachten, dass die Schätzungen in den einzelnen Kategorien auf sehr unterschiedlichen Stichprobenumfängen beruhen. Während die Schätzungen in der Kategorie **Normal** auf etwa 59.000 Stichproben beruhen, sind es beispielsweise in der Kategorie **Height** nur etwa 2.300. Um einen Vergleich und Test auf signifikante Unterschiede zwischen den Kategorien zu ermöglichen, wurde eine Reihe von Monte-Carlo-Simulationen durchgeführt. Dazu wurden aus der Kategorie **Normal** 10.000-mal Stichproben gezogen, deren Größe jeweils dem Stichprobenumfang in den Extremwellenkategorien entsprach. Anhand dieser Stichproben konnte getestet werden, wie wahrscheinlich es ist, dass die in den einzelnen Kategorien beobachteten Verteilungen signifikant von denen der Kategorie **Normal** abweichen. Für den Broadness Parameter  $\varepsilon$  konnte für die Kategorien Height/Extreme ein signifikanter Unterschied detektiert werden, wonach nach diesen Kriterien definierte Rogue Waves üblicherweise in etwas schmalere Frequenzspektren auftreten als im Mittel vorherrschen. Für Rogue Waves, die nach dem Crest Kriterium definiert werden, ließ sich dagegen kein Unterschied nachweisen (**Abb. 4**).

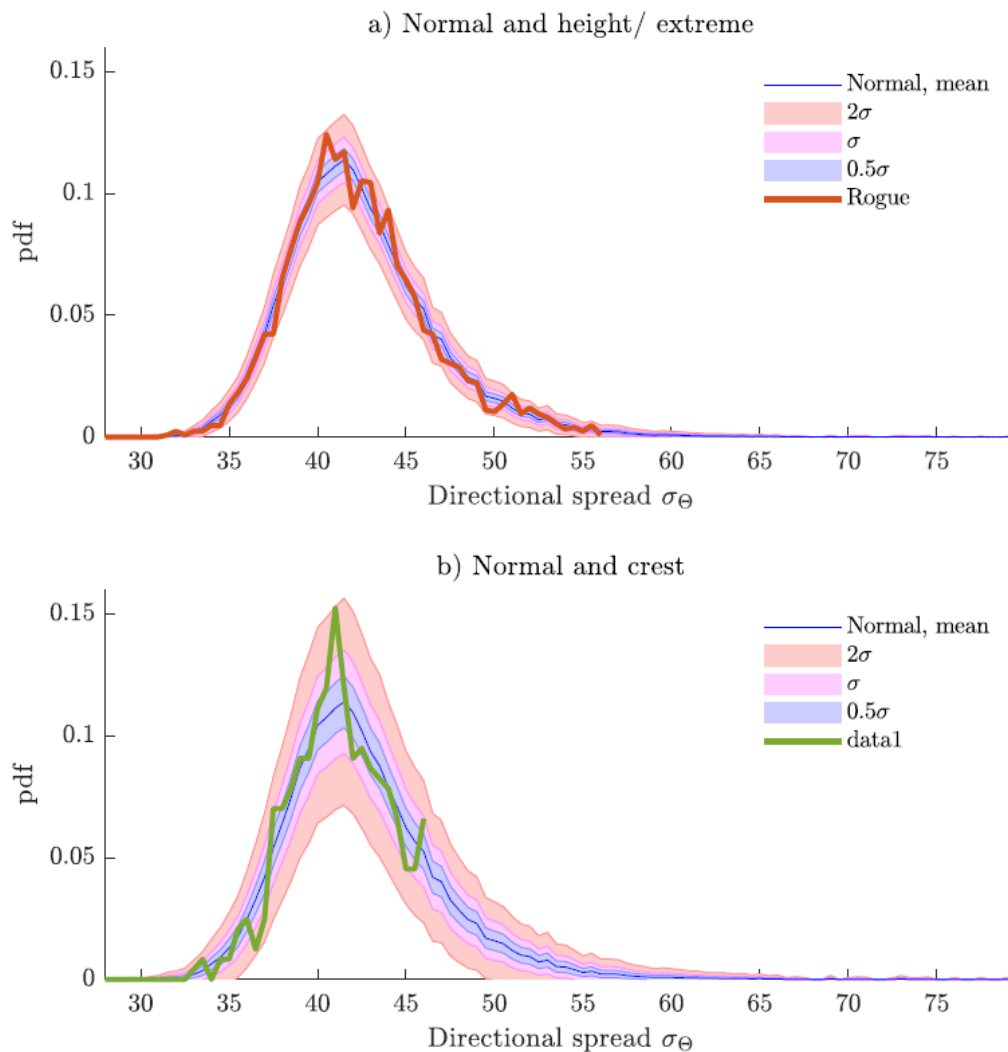


**Abb. 4:** Ergebnis der Monte-Carlo-Simulationen für a) 2510 und b) 485 Verteilungen des Parameters  $\varepsilon$ . Blau: Mittelwert der Verteilungen; lila: Bereich innerhalb von 0.5 Standardabweichungen; rosa: Bereich innerhalb einer Standardabweichung; orange: Bereich innerhalb von zwei Standardabweichungen.

Die Untersuchungen und Analysen wurden analog für den „Narrowness Parameter“  $\nu$ , Goda's Peakness Parameter  $Q_p$  und den Benjamin Feir Index  $BFI$  durchgeführt. Im Gegensatz zum Broadness Parameter  $\varepsilon$  konnten hier jedoch keine signifikanten Unterschiede nachgewiesen werden. Unter der Hypothese, dass die Modulations-Instabilität die an den betrachteten Stationen beobachteten Rogue Waves verursacht hat, wäre ein höherer BFI in den Stichproben mit Extremwellen zu erwarten gewesen. Dieses konnte anhand unserer Analysen nicht bestätigt werden.

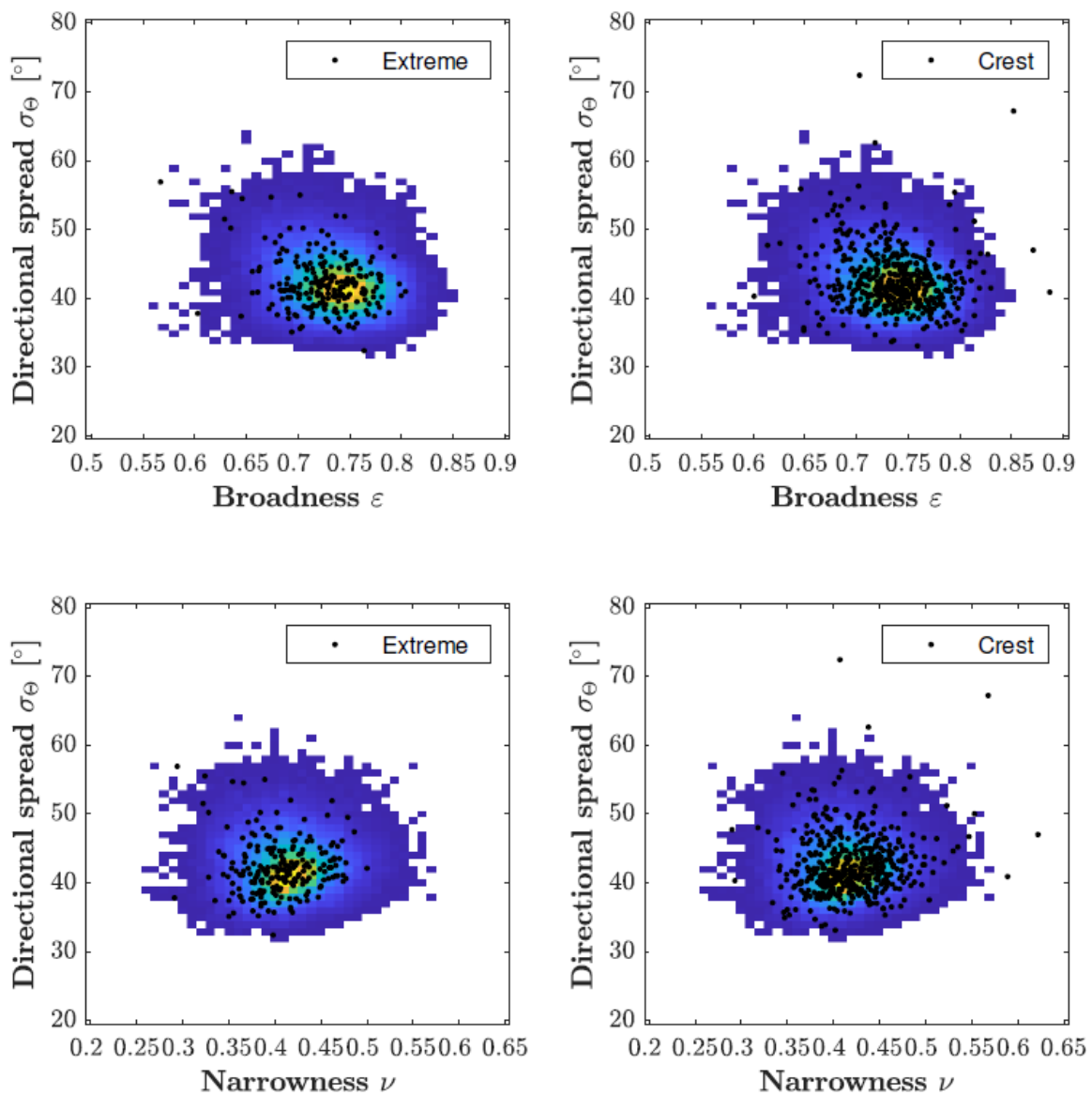
Neben einem schmalen Frequenzspektrum ist eine geringe Richtungsstreuung („directional spread“) eine Voraussetzung, damit sich eine Modulations-Instabilität entwickeln kann. Es wurde deshalb für alle Stationen die Richtungsstreuung aus den gemessenen Spektren berechnet und die jeweiligen Daten entsprechend der oben beschriebenen Vorgehensweise einer der Kategorien Normal, Height, Extreme, Crest bzw. Double zugewiesen. Anschließend wurden wiederum Monte-Carlo-Simulationen entsprechend der oben beschriebenen Vorgehensweise

durchgeführt. Anhand dieser Untersuchungen ließen sich keine signifikanten Unterschiede in der Richtungsstreuung zwischen Stichproben mit und ohne Rogue Waves nachweisen (**Abb. 5**).



**Abb. 5:** Ergebnis der Monte-Carlo-Simulationen für a) 2510 und b) 485 Verteilungen der Richtungsstreuung  $\sigma_\theta$ . Blau: Mittelwert der Verteilungen; lila: Bereich innerhalb von 0.5 Standardabweichungen; rosa: Bereich innerhalb einer Standardabweichung; orange: Bereich innerhalb von zwei Standardabweichungen.

Abschließend wurde noch die Kombination aus spektraler Bandbreite und Richtungsstreuung untersucht. Dies geschah, weil als günstige Bedingung für die Entstehung einer Modulations-Instabilität ein schmales Spektrum sowohl in Richtung als auch in Frequenz angesehen wird [Alber, 1978; Stansberg, 1995]. **Abb. 6** zeigt die gemeinsame Häufigkeitsverteilung aus Bandbreite und Richtungsstreuung für alle Stichproben ohne Extremwelle (Farbcodierung). Zusätzlich sind jeweils die Werte für die zugehörigen Extremwellenstichproben als schwarze Punkte eingetragen. Es ist zu erkennen, dass sich diese um die Mittelwerte der Verteilung der Zeitreihen ohne Extremwelle clustern und breit innerhalb der auch in normalen Zeitreihen beobachteten Werte streuen. Auch hier lässt sich kein deutlicher Unterschied zwischen Zeitreihen mit und ohne Extremwelle erkennen. Insbesondere ist keine Häufung bei niedriger Bandbreite und Richtungsstreuung zu erkennen, was auf günstige Bedingungen für die Entwicklung von Modulations-Instabilitäten hinweisen würde.



**Abb. 6:** Paarweise Darstellung von spektraler Bandbreite  $\varepsilon$  bzw.  $\nu$  und Richtungsstreuung. Vergleich von Zeitreihen mit und ohne Extremwelle. Farbcodierung: Dichte der Datenpunkte „normaler“ Extremwellensamples. Schwarze Punkte: Werte aus Extremwellenzeitreihen. Diese beinhalten entweder „extreme“ Extremwellen oder Extremwellen nach dem Wellenbergkriterium, siehe jeweilige Legende.

Die Modulations-Instabilität tritt vor allem dann auf, wenn die Wellen ähnliche Frequenzen und Richtungen haben. In unseren Daten waren diese Bedingungen nicht erfüllt, so dass die Modulations-Instabilität nicht den wahrscheinlichsten Mechanismus für die Entstehung von Extremwellen in der südlichen Nordsee darstellt.

## 3.2 Rogue Waves in Flachwasser und die Rolle von Solitonen

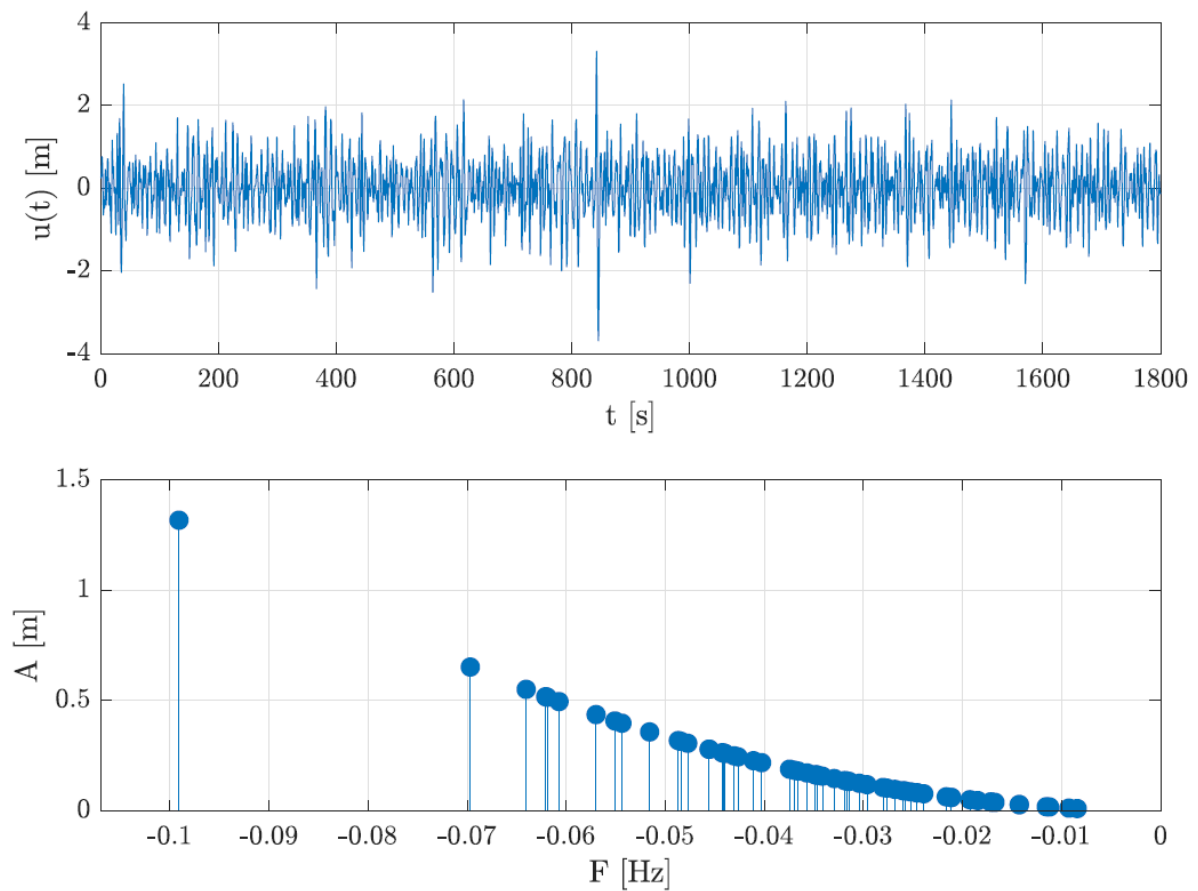
Im Vorgänger Projekt FEAK WAVES I wurde festgestellt, dass an Flachwasserstationen wie z.B. der Boje SEE vor Norderney auffällig viele Extremwellen gemessen wurden. Die Entwicklung von Wellen im Flachwasser wird typischerweise mit Hilfe der KdV Gleichung [Korteweg and Vries, 1895] beschrieben. Lösungen der KdV Gleichung unterteilen das Spektrum in Anteile mit (nichtlinearen) oszillierenden Wellen und in Anteile mit Solitonen. Unter einem Soliton wird dabei ein Wellenpaket verstanden, das sich ohne Änderung seiner Form durch ein dispersives und zugleich nichtlineares Medium bewegt.

Im Küstenbereich bilden sich Solitone häufig bei der Passage von Wellen über Bereiche starker Tiefegradienten [Sergeeva et al., 2011]. Da im Bereich der Boje SEE vor Norderney erhebliche Änderungen der Wassertiefe erfolgen, wurde die Rolle von Solitonen bei der Bildung von Extremwellen in diesem Bereich untersucht [Teutsch et al., 2023].

Dazu wurden die Zeitreihen an der Boje SEE einer nichtlinearen Fouriertransformation basierend auf der KdV Gleichung mit verschwindenden Randbedingungen (vKdV-NLFT) unterzogen. Diese unterteilt das Spektrum in Anteile mit oszillierenden Wellen und in Anteile mit Solitonen. **Abb. 7** zeigt als Beispiel eine Messzeitreihe und das zugehörige Soliton Spektrum. Um die Solitone von oszillierenden Wellen zu unterscheiden, werden die Solitone im Folgenden auf einer negativen Frequenzachse dargestellt. Technisch gesehen hat die Frequenzachse keine physikalische Bedeutung, da bei einem Soliton die Oberflächenauslenkung die Ruhewasserlinie nicht unterschreitet [Brühl and Oumeraci, 2016]. Die Frequenz  $f$  wurde aus der Lösung der KdV mittels

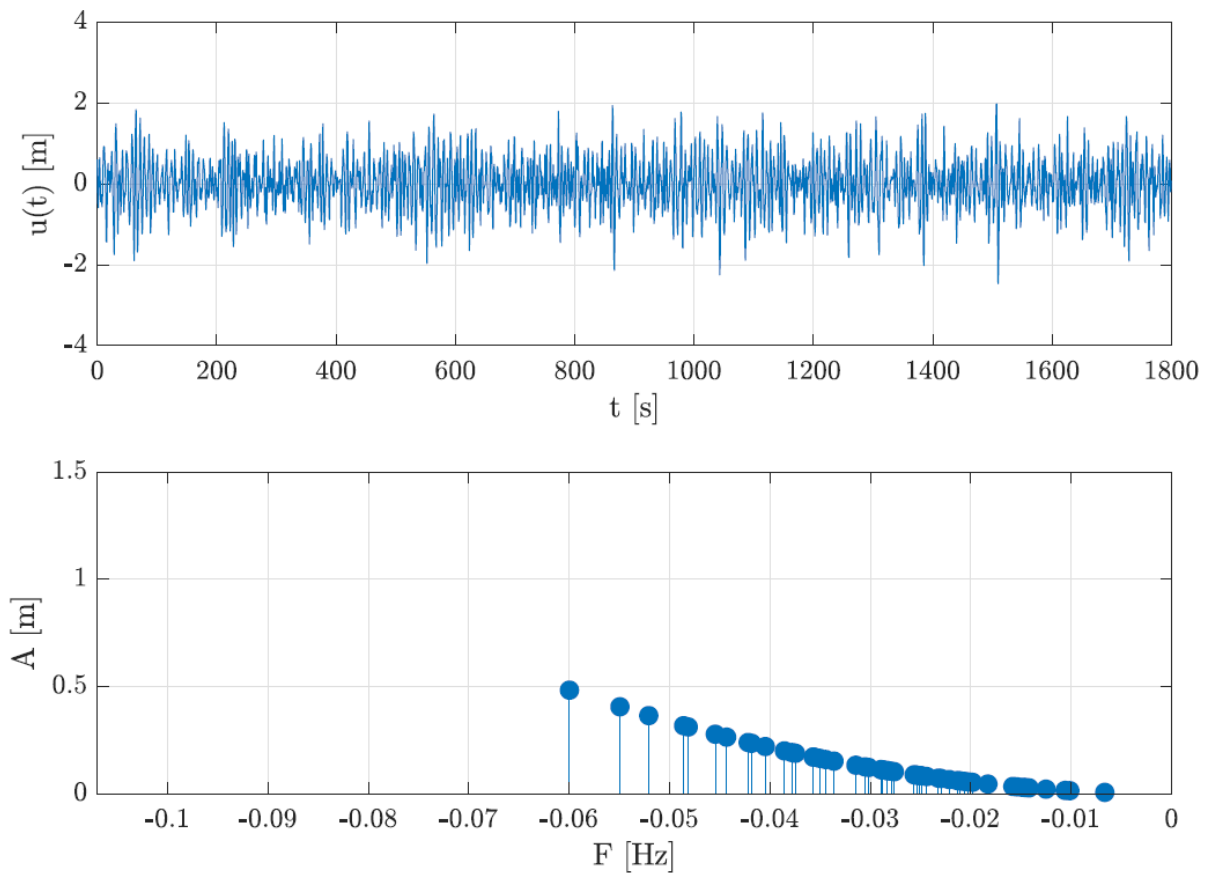
$$\omega = 2\pi f = \sqrt{\frac{3Ag}{4h^2}}$$

bestimmt, wobei  $\omega$  die Kreisfrequenz,  $h$  die Wassertiefe,  $g$  die Gravitationsbeschleunigung und  $A$  die Amplitude des Solitons beschreibt. Durch diese Gleichung wird die Frequenz in eine Beziehung mit der Amplitude des Solitons gesetzt, so dass die Solitone entlang der Frequenzachse nach der Größe ihrer Amplitude sortiert werden.



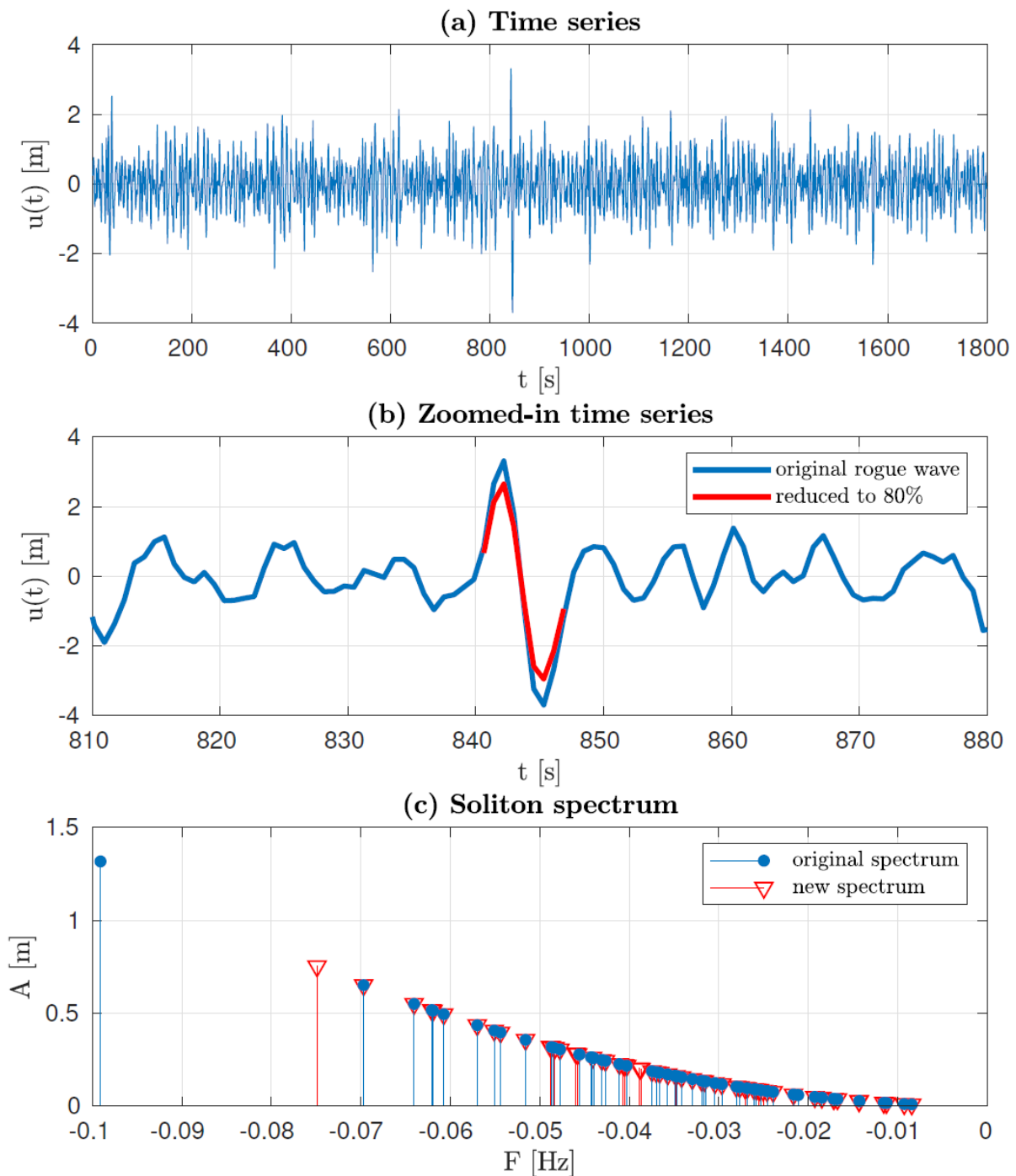
**Abb. 7:** Beispiel einer Messzeitreihe mit einer Extremwelle nach ungefähr 820 s (oben) und das zugehörige Soliton Spektrum basierend auf der  $\nu$ KdV-NLFT (unten). Die signifikante Wellenhöhe zum Zeitpunkt der Messung betrug 2,71 m, die Höhe der Extremwelle betrug 7,00 m. Die Zeitreihe wurde am 17.10.2013 an der Boje SEE gemessen.

Die Auswertung der Solitonspektren ergab, dass in der Mehrzahl der Fälle, in denen eine Extremwelle in den Messungen vorhanden war, das Spektrum durch ein einzelnes Soliton charakterisiert war, das deutlich höher als alle anderen war und deshalb im Solitonspektrum durch einen größeren Abstand zu den anderen auffiel („outstanding soliton“). Ein solcher „outstanding soliton“ ist in **Abb. 7** beispielsweise zwischen etwa -0.09 und -0.1 Hz zu erkennen. In Zeitreihen ohne Extremwelle waren solche outstanding solitons in weniger als der Hälfte der untersuchten Zeitreihen vorhanden (**Abb. 8**)



**Abb. 8:** Beispiel einer Messzeitreihe ohne Extremwelle (oben) und das zugehörige Soliton Spektrum basierend auf der  $v$ KdV-NLFT (unten). Die signifikante Wellenhöhe zum Zeitpunkt der Messung betrug 2,46 m, die Höhe der größten Welle im Zeitfenster betrug 4,44 m. Die Zeitreihe wurde am 26.12.2016 an der Boje SEE gemessen.

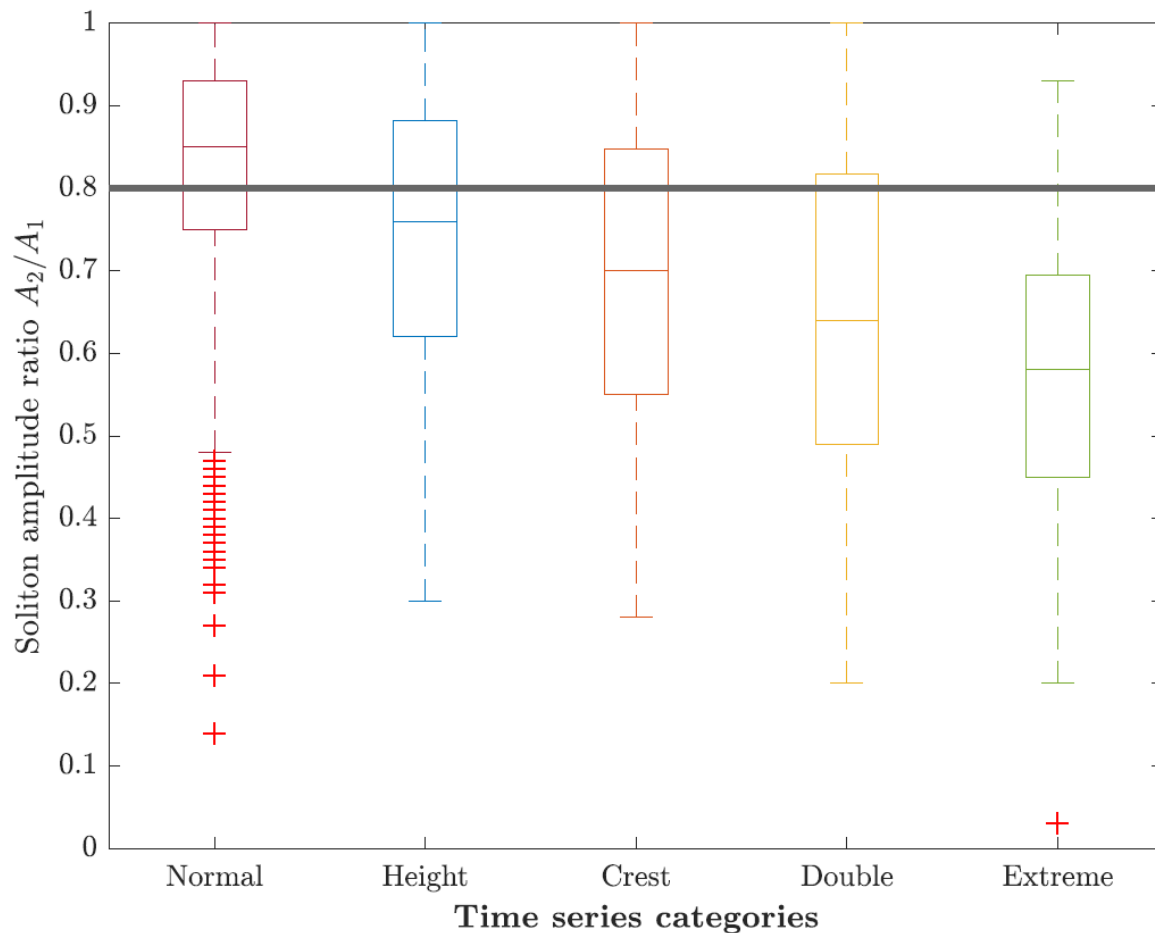
Um zu untersuchen, inwieweit sich diese outstanding solitons den beobachteten Extremwellen zuordnen lassen, wurden die Wasserstandsauslenkungen innerhalb zweier Nulldurchgänge um eine Extremwelle auf 80% der ursprünglichen Werte herab skaliert (**Abb. 9b**). Anschließend wurden erneut die Solitonspektren berechnet und mit den ursprünglichen verglichen. Für alle untersuchten Fälle zeigte sich, dass sich bei diesem Vorgehen die Amplitude des größten Solitons deutlich reduzierte, während die kleineren Solitone nahezu unverändert blieben. Da Amplitude und Frequenz gemäß obiger Gleichung zusammenhängen, verschiebt sich der der Extremwelle zugehörige Soliton und rückt näher an die übrigen Solitone heran (**Abb. 9**).



**Abb. 9:** (a) Zeitreihe mit Extremwelle gemessen am 17.10.2023; (b) Vergrößerung des Zeitabschnittes mit der Extremwelle und Darstellung der reduzierten Zeitreihe (rot); (c) Soliton Spektren der Original- (blau) und der modifizierten Zeitreihe (rot).

Um im Folgenden die Spektren einfacher klassifizieren zu können, wurde für alle Zeitreihen das Verhältnis der Amplituden des größten  $A_1$  und des zweitgrößten  $A_2$  Solitons berechnet. Dieses Verhältnis  $\frac{A_2}{A_1}$  misst den Abstand der beiden größten Solitone im Spektrum, wobei kleinere Werte auf einen größeren Abstand im Spektrum hinweisen.

**Abb. 10** zeigt die Verteilung des Verhältnisses der Soliton Amplituden für alle Fälle mit und ohne Extremwellen. Es ist deutlich zu erkennen, dass Fälle mit Extremwellen im Mittel höhere Solitonabstände (geringe Verhältnisse) aufweisen, als ohne.



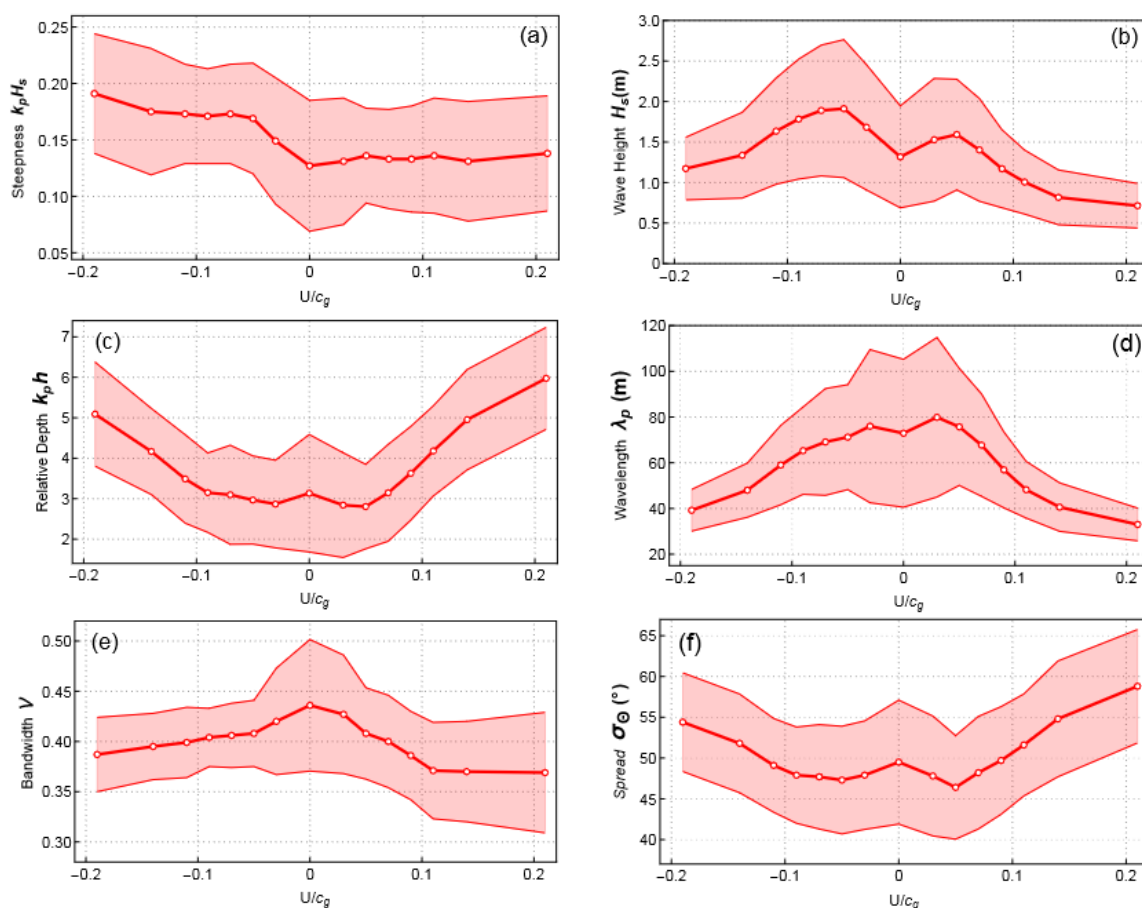
**Abb. 10:** Box-Whiskerplots des Verhältnisses der Amplituden des zweitgrößten und des größten Solitons im Spektrum für Fälle mit und ohne Extremwellen. Bei Fällen mit Extremwellen wird weiterhin unterschieden in Extremwellen Definitionen nach verschiedenen Kriterien.

Um zu untersuchen, inwieweit die Ergebnisse der Untersuchungen an der Boje SEE verallgemeinert werden können, wurden die Untersuchungen für zwei zusätzliche Flachwasserstationen durchgeführt [Teutsch et al., 2024b]. Auf der Grundlage dieser zusätzlichen Daten konnten die zuvor anhand der Boje SEE gewonnenen Erkenntnisse bestätigt werden, dass das Vorhandensein eines großen, einzelnstehenden Solitons im NLFT-Spektrum bei Zeitreihen, die eine Rogue Wave enthalten wahrscheinlicher ist, als bei Zeitreihen ohne eine solche Extremwelle.

### 3.3 Einfluss von Tidenströmungen auf die Statistik von Rogue Waves

Theoretische Studien über die Modulation eindimensionaler regelmäßiger Wellen über einem flachen Boden durch eine Strömung weisen in der Regel eine Asymmetrie zwischen den Auswirkungen von Gegen- und Folgeströmungen auf die Entwicklung von Seegangparametern wie der signifikanten Wellenhöhe auf. Entsprechend der Theorie wird erwartet, dass die signifikante Wellenhöhe bei gegenläufigen Strömungen monoton zunimmt und bei nachfolgenden Strömungen abnimmt.

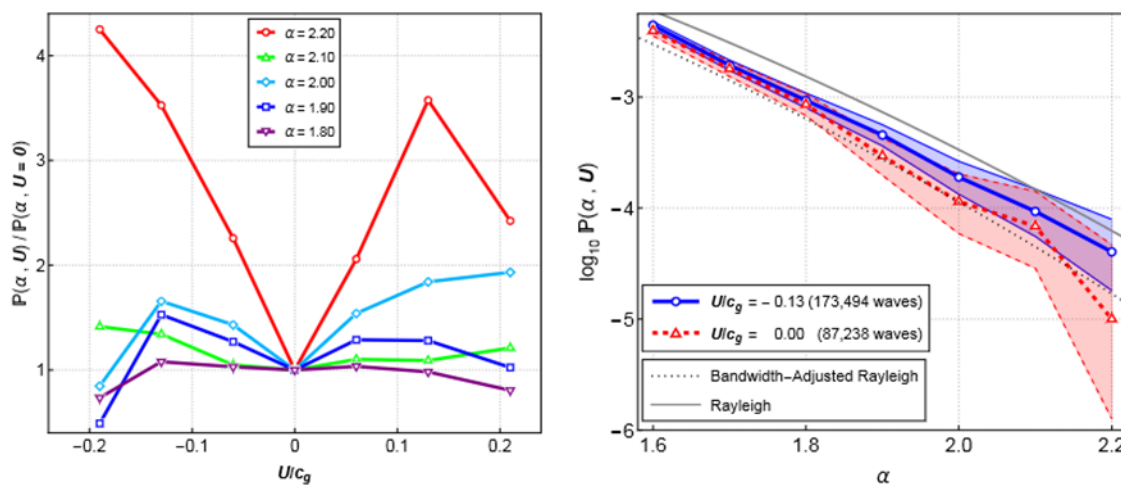
Anhand des in FREAK WAVE II verwendeten mehrjährigen Datensatzes konnte gezeigt werden, dass insbesondere in sehr breitbandigen Seegangsspektren mit hoher Richtungsstreuung, ein solches Verhalten nicht beobachtet werden konnte. Stattdessen konnte insbesondere bei hohen Werten des Verhältnisses zwischen Gezeitenstrom und Gruppengeschwindigkeit eher eine symmetrische Modulation verschiedener Seegangparameter nachgewiesen werden [Teutsch et al., 2024a] (**Abb. 11.**)



**Abb. 11:** Mittlere Werte (rote Linie) und zugehöriger Bereich von  $\pm$  einer Standardabweichung (rot schraffierter Bereich) verschiedener Seegangparameter (y-Achse) in Abhängigkeit von der relativen Strömungsgeschwindigkeit an der Station Fino 1.

In numerischen Simulationen und Laborstudien konnte nachgewiesen werden, dass solche Inhomogenitäten im Wellenfeld, die durch Untiefen, Kreuzseen oder Strömungsinteraktionen

verursacht werden, die Wellenstatistik verändern und die Wahrscheinlichkeit des Auftretens von Rogue Waves erhöhen [z.B. *Hjelmervik and Trulsen, 2009*]. In FREAK WAVES II wurden anhand von Langzeitbeobachtungen Auswirkungen von Strömungsintensität und -richtung auf die Häufigkeit des Auftretens von Rogue Waves in der südlichen Nordsee untersucht. Analog der Ergebnisse für die Seegangparameter [*Teutsch et al., 2024a*] konnte gezeigt werden, dass die Erhöhung der Häufigkeit von Extremwellen bei gegenläufigen und nachfolgenden Strömungen ähnlich ist [*Mendes et al., 2024*] (**Abb. 12**). Obwohl die Seezustände überwiegend breitbandig sind und eine große Richtungsstreuung aufweisen, liegen die Verstärkungen in der gleichen Größenordnung wie die, die in unidirektionalen Laborexperimenten für stationäre Strömungen beobachtet wurden [*Mendes et al., 2024*].



**Abb. 12:** (Links) Verstärkung der Überschreitens-Wahrscheinlichkeit verschiedener Grenzwerte (Farben) für die Definition von Rogue Waves in Abhängigkeit von der relativen Strömungsgeschwindigkeit an der Station Fino 1. (Rechts) Überschreitens-Wahrscheinlichkeiten bei minimaler (rot) und maximaler (blau) relativer Strömungsgeschwindigkeit im Vergleich zur Rayleigh Verteilung.

### 3.4 Vergleich und Eignung von Ansätzen des maschinellen Lernens

Um Zusammenhänge zwischen Seegangparametern und dem Auftreten von Rogue Waves zu untersuchen, bieten sich auf Grund des umfangreichen vorliegenden Beobachtungsmaterials ebenfalls Ansätze des interpretierbaren maschinellen Lernens an, um große Mengen von Beobachtungsdaten zu analysieren. Ausgehend von einem in *Häfner et al. [2021]* entwickelten Ansatz wurde dazu über alle Beobachtungen ein 30-minütiges Zeitfenster entlang der Zeitachse von einer Welle zur nächsten verschoben und so für jede Welle eine große Anzahl von Seegangparametern berechnet, die das jeweils vorherrschende Seegangsfeld beschreiben (**Tab. 1**). Anschließend wurde mit verschiedenen Ansätzen untersucht, inwieweit wesentliche Parameter identifiziert werden können, von denen die Bildung von Extremwellen abhängt und inwieweit eine Kurzfristprognose mit den verwendeten Daten möglich erscheint. In allen folgenden Untersuchungen ist dabei die Zielvariable maximale relative Wellenhöhe ( $H/H_s$ ) in den kommenden 10 Minuten (**Tab. 1**). Diese Zielgröße wird im Folgenden als Abnormitätsindex bezeichnet.

**Tab. 1:** Eingabe- (Features) und Zielparameter (Target) für die verwendeten Algorithmen des maschinellen Lernens. Die Parameter beziehen sich jeweils auf die 30 Minuten, die der jeweiligen Welle vorausgehen.

Parameter	Formelzeichen	Beschreibung
Signifikante Wellenhöhe	$H_s$	mittlere Wellenhöhe des höchsten Drittels der Wellen einer halben Stunde
Mittlere Wellenperiode	$T_0$	zeitlicher Abstand zwischen zwei positiven Nulldurchgängen der Wasserspiegelauslenkung
Kurtosis	$\lambda_{40}$	Maß für die Relevanz von Ausreißern
Schiefe	$\lambda_{30}$	Maß der Asymmetrie der Daten um den Mittelwert
Wellenlänge	$L_{deep}$	Mittlere Wellenlänge, bestimmt nach der Gleichung $L = g/2\pi \cdot T^2$ für tiefes Wasser
Mittlere Steilheit	$s$	$s = H/L_{deep}$
Mittlere Asymmetrie	$\mu$	Verhältnis zwischen Amplitude und Wellenhöhe einer Welle
Relative Wassertiefe	$kh$	Wellenzahl $k = 2\pi/L_{deep}$ ; Wassertiefe $h = 30$ m an FiNO1
Ursell number	$Ur$	$Ur = HL^2/h^3$ , Maß für Nichtlinearität
Peak-Periode	$T_p$	Umkehrwert der Peak-Frequenz des FFT-Spektrums
Spektrale Bandbreite	$\varepsilon$	"Broadness parameter" $\varepsilon = \sqrt{\frac{m_0 m_4 - m_2^2}{m_0 m_4}}$ mit Spektralmomenten $m_n$
Spektrale Bandbreite	$\nu$	"Narrowness parameter" $\nu = \sqrt{\frac{m_0 m_2}{m_1^2} - 1}$ mit Spektralmomenten $m_n$
Peakedness-Parameter	$Q_p$	$Q_p = \frac{2}{m_0^2} \int_0^\infty f \left[ \int_0^{2\pi} S(f, \Theta) d\Theta \right]^2 df$ des Spektrums $S(f, \Theta)$
Benjamin-Feir-Index	BFI	$BFI = \frac{\sqrt{2}\varepsilon}{\Delta\omega/\omega_0}$ mit Steilheit $\sqrt{2}\varepsilon$
Berg-Tal-Verhältnis	$r$	$r = \frac{1}{m_0} \sqrt{\rho^2 + \lambda^2}$
Dünungsanteil	swell	Anteil niedrigfrequenter Wellen in der Zeitreihe
Windgeschwindigkeit	$v_{wind}$	gemessen am Mast FiNO1
Böengeschwindigkeit	$v_{gust}$	kurzzeitiger Höchstwert der Windgeschwindigkeit

Mittlere Windrichtung	$\Theta_{wind}$	Werte als Namen codiert (z.B. N, SE, WSW)
Lufttemperatur	$T_{air}$	
Temperaturunterschied	$\Delta_T$	Bezogen auf eine Stunde
Luftdruck	$p$	gemessen am Mast FiNO1
Druckunterschied	$\Delta_{p,1h}$	Bezogen auf eine Stunde
Druckunterschied	$\Delta_{p,3h}$	Bezogen auf drei Stunden
Druckunterschied	$\Delta_{p,6h}$	Bezogen auf sechs Stunden
Druckunterschied	$\Delta_{p,12h}$	Bezogen auf zwölf Stunden
Objektive Wetterlage	Obj	nach der Definition des BSH für die Nordsee: <a href="https://www.bsh.de/DE/DATEN/Klima-und-Meer/Wetterlagen-Stuerme/Wetterlagen_Sturmklassifizierung/wetterlagen-sturmklassifizierung_node.html">https://www.bsh.de/DE/DATEN/Klima-und-Meer/Wetterlagen-Stuerme/Wetterlagen_Sturmklassifizierung/wetterlagen-sturmklassifizierung_node.html</a>
<b>Zielvariable</b>	$AI_{10min}$	Maximale relative Wellenhöhe $H/H_s$ innerhalb der kommenden 10 min

In einem ersten Schritt wurde dazu ein „Random Forest Modell“ verwendet. Random Forest (RF) ist ein populärer Algorithmus für maschinelles Lernen, der für seine Einfachheit und die relativ wenigen Hyperparameter bekannt ist, die für das Tuning erforderlich sind. Es handelt sich um eine überwachte Lernmethode, die Ensemble-Techniken einsetzt, indem die Vorhersagen mehrerer Entscheidungsbäume gemittelt werden. Normalerweise wird ein Bagging-Ansatz verwendet, um jeden Entscheidungsbaum zu trainieren. Beim Bagging werden Daten aus dem Trainingsdatensatz entnommen, um einen individuellen Entscheidungsbaum zu erstellen. Die entnommenen Daten werden dann in den gesamten Datensatz zurückgeführt. Dieser Prozess reduziert die Varianz und erhöht die Genauigkeit, indem die Vorhersagen mehrerer Entscheidungsbäume zusammengefasst werden.

Zum Vergleich wurde anschließend zusätzlich ein neuronales Netz (NN) trainiert. Dabei handelt es sich um einen Ansatz des Deep Learning, der sowohl für Regressions- als auch für Klassifizierungsaufgaben geeignet ist. Im Laufe der Zeit wurden verschiedene neuronale Netzwerkarchitekturen entwickelt. In dieser Arbeit wurde ein neuronales Feed-Forward-Backpropagation-Netzwerk verwendet, bei dem verborgene Schichten von Neuronen mit den Eingabe- und Ausgabeschichten verbunden sind und ein sogenanntes mehrschichtiges Perzeptron (MLP) bilden. Diese Netzwerke sind sehr effizient, selbst bei Datensätzen mit bis zu einer Milliarde Einträgen. Für diese Studie wurde Keras, eine Open-Source-Python-Bibliothek und High-Level-API für das TensorFlow-Framework verwendet, um die MLP-Modelle zu entwickeln.

Keras gilt weithin als einfaches, flexibles und leistungsfähiges Werkzeug im Bereich des maschinellen Lernens.

In einem nächsten Schritt wurden anschließend zwei Methoden zur Interpretierbarkeit der Modelle angewandt. Bei der ersten Methode, dem Forest-Guided Clustering (FGC) [<https://forest-guided-clustering.readthedocs.io>] handelt es sich um eine modellspezifische Erklärungsmethode für entscheidungsbaumbasierte maschinelle Lernmodelle. FGC gruppiert die Daten in Untergruppen, für die dieselben Entscheidungsregeln gelten. FGC berechnet dabei die Relevanz der einzelnen Eingabeparameter für die Entscheidungen des Random Forest bei der Vorhersage in den einzelnen Untergruppen. Ein entscheidender Vorteil von FGC ist, dass es nicht auf der Annahme der Merkmalsunabhängigkeit beruht. Folglich bleiben die Ergebnisse unbeeinflusst von korrelierten Merkmalen.

Die Nähe zwischen Datenpunkten wird mit Hilfe eines Proximity-Maßes bestimmt, das angibt, wie oft zwei Instanzen in denselben Endknoten eines Baumes innerhalb des RF-Modells erscheinen. Anhand dieses Maßes werden Datenpunkte, die ähnlichen Entscheidungspfaden im RF-Modell folgen, mit der K-medoids-Technik geclustert. Obwohl diese Algorithmen die Beiträge von Merkmalen zur Zielvariablen nicht direkt aufzeigen, können die Ergebnisse des Clusters helfen, Schlüsselmerkmale zu identifizieren. Diese Merkmale können dann verwendet werden, um ein einfacheres ML-Modell für die Vorhersage des Abnormitätsindex zu erstellen.

Als zweite Methode zur Interpretierbarkeit einzelner Vorhersagen eines maschinellen Lernmodells wurde die sogenannte Shapley Additive Explanations (SHAP) Methode eingesetzt [<https://shap.readthedocs.io>]. SHAP basiert auf Ansätzen aus der Spieltheorie und bietet eine Möglichkeit zur Interpretation von Machine Learning (ML)-Modellen. Im Gegensatz zu FGC ist SHAP modellunabhängig, d.h. es kann auf jedes maschinelle Lernmodell angewandt werden, um zu verstehen, wie die Merkmale zur Ausgabe des Modells (Zielvariable) beitragen. Außerdem liefert SHAP im Gegensatz zu FGC lokale Merkmalsbedeutungswerte, die den Beitrag jedes Merkmals zur Vorhersage eines bestimmten Datenpunkts quantifizieren. Der SHAP-Wert eines Merkmals stellt seinen Beitrag zum Gesamtergebnis dar, gewichtet und summiert über alle möglichen Merkmalswertkombinationen.

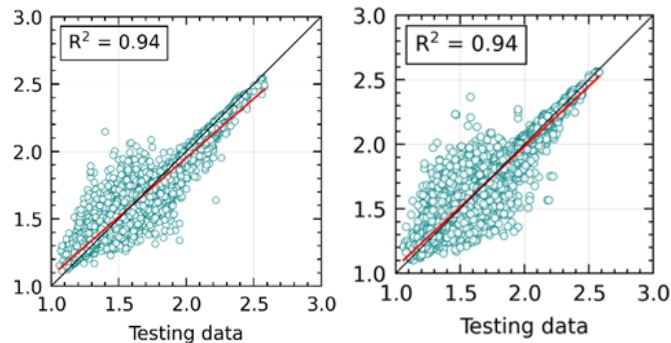
Beide Modelle des maschinellen Lernens schneiden bei den Testdaten (ungesehene Daten) prinzipiell gut ab und zeigen zufriedenstellende Vorhersagefähigkeiten (**Abb. 13**). Werden die Modelle eingesetzt, um Fälle mit Abnormitätsindex größer/kleiner als 2.0 (Rogue/Nicht Rogue-Fälle) zu unterscheiden, ergeben sich sehr ähnliche Konfusions-Matrizen (**Tab. 2**).

**Tab. 2:** Konfusions-Matrix für das Random Forrest Modell für die Testdaten, die nicht zum Training des Modells verwendet wurden.

		Beobachtete	
		Abnormitätsindex (AI)	
		AI >= 2.0	AI < 2.0
Vorhergesagte	AI >= 2.0	3.080	19
Abnormitätsindex	AI < 2.0	560	25.736

Von den im Testdatensatz beobachteten 3.640 Rogue Waves wurden durch das RF-Modell etwa 84% korrekt vorhergesagt (True Positive,  $TP$ ). Das Modell zeigt eine hohe Präzision  $\left(\frac{TP}{TP+FP}\right)$  von

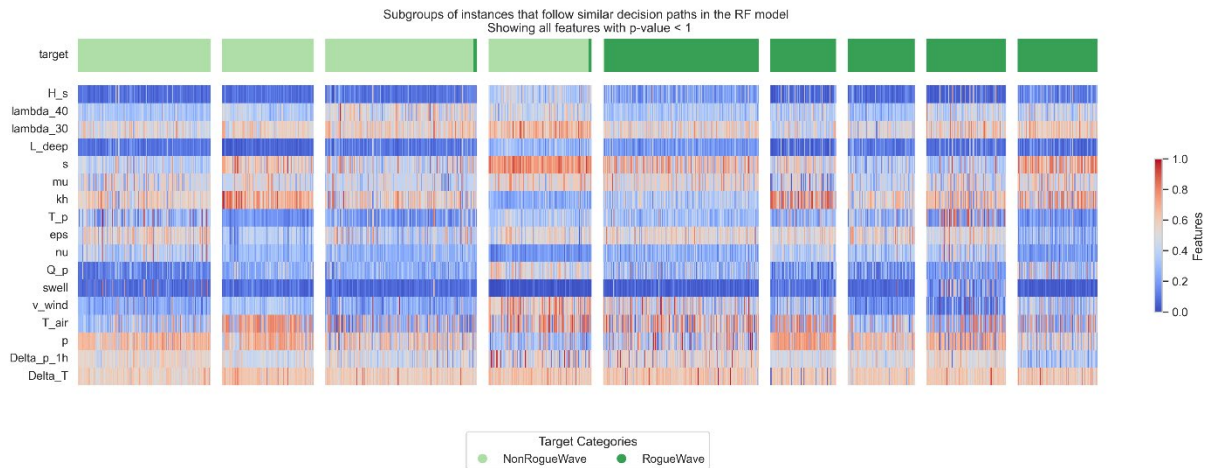
etwa 99%, d.h. es ist sehr gut im Vermeiden falsch positiver Vorhersagen ( $FP$ ). Die Genauigkeit korrekter Vorhersagen ( $\frac{TP+FP}{N}$ ) liegt bei etwa 98%. Mit Hilfe des NN-Modells konnten ähnliche Werte erzielt werden.



**Abb. 13:** Vergleich zwischen beobachteten Abnormitätsindex (x-Achse) und durch das RF-Modell (links) bzw. das NN-Modell (rechts) vorhergesagten Abnormitätsindex (y-Achse) für die nicht zum Training der Modelle verwendeten Testdaten.

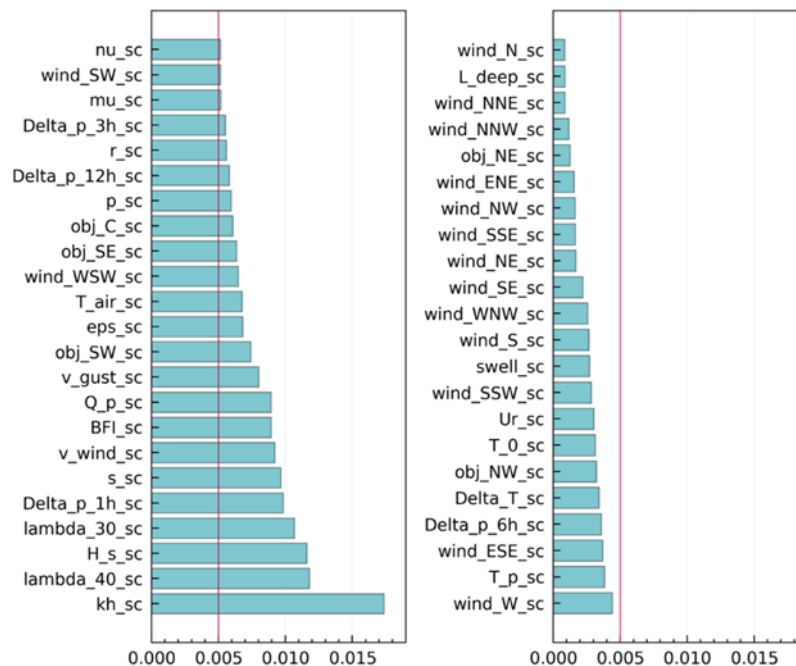
Die Ergebnisse der Algorithmen zur Interpretierbarkeit beider ML-Algorithmen können unterschiedlichen Zwecken dienen. FGC kann verwendet werden, um die Anzahl der Eingangsmerkmale zu reduzieren und ein effizienteres Vorhersagemodell mit weniger Prädiktoren zu erstellen. SHAP hingegen kann eingesetzt werden, um die Beiträge von Merkmalen zu verschiedenen Abnormitätsindizes zu verstehen, was ein besseres Verständnis der Entstehung von rogue und nicht-rogue Instanzen ermöglicht.

Die Interpretierbarkeit des trainierten RF-Klassifizierungsmodells wurde mittels FGC untersucht. Die Ermittlung der optimalen Anzahl von Clustern für FGC erwies sich als schwierig. Daher wurden größere Clusteranzahlen verwendet, in der Erwartung, dass sie unterschiedliche Bedingungen für die Bildung von Rogue Waves erkennen lassen. Die clusterspezifischen Merkmalswertverteilungen für 9 Cluster sind in **Abb. 14** dargestellt. Obwohl die Clusterung nicht perfekt ist, wenn man die Klassifizierung von Rogue- und Non-Rogue-Wellen betrachtet, sind Unterschiede in den Merkmalswertverteilungen erkennbar. Insbesondere scheint es verschiedene Cluster mit hohen Abnormitätsindizes zu geben, was auf unterschiedliche Entstehungsmechanismen für Rogue Waves hinweist.



**Abb. 14:** Untergruppen von Instanzen, die im Random-Forest-Modell ähnlichen Entscheidungswegen folgen und Merkmale mit einem p-Wert von weniger als 0,01 aufweisen. Ein Zielwert (target) von 1 weist auf eine potenzielle Rogue Wave hin, bei der der Abnormitätsindex gleich oder größer als 2 ist.

Die Interpretierbarkeit des NN-Modells wurde mit Hilfe von SHAP Werten untersucht. Demnach spielen insbesondere Parameter wie die relative Wellenhöhe, Schiefe und Kurtosis der Verteilung oder die signifikante Wellenhöhe eine Rolle bei der Erklärung des Auftretens hoher Abnormitätsindizes (**Abb. 15**).

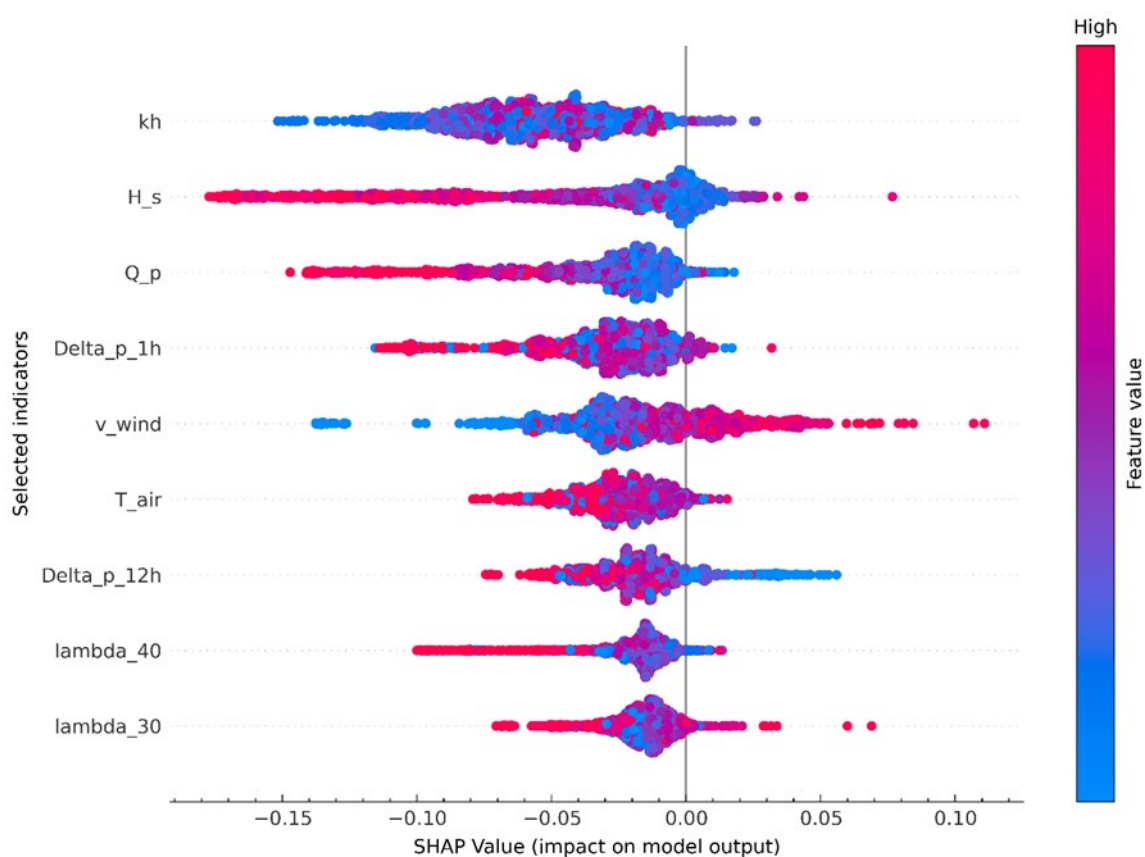


**Abb. 15:** Globale Merkmals-Bedeutung basierend auf SHAP.

Unter Verwendung der globalen Merkmalsbedeutungsanalyse aus SHAP wurden Schlüsselmerkmale auf der Grundlage eines Schwellenwerts von 0,02 ausgewählt. Unter Verwendung dieser ausgewählten Merkmale wurde dann erneut ein reduziertes neuronales Netzwerkmodell entwickelt. Basierend auf diesem Modell wurden mittels SHAP lokale Merkmalsbedeutungswerte für die einzelnen Merkmale ermittelt (**Abb. 16**). Dabei zeigte sich, dass die relative Wassertiefe ein wichtiges Merkmal darstellt, das die Ausgabe der Zielgröße

(Abnormitätsindex) wesentlich beeinflusst. Dabei scheinen hohe relative Wassertiefen im Mittel den Abnormitätsindex leicht zu reduzieren, wogegen geringe relative Wassertiefen sowohl mit niedrigen als auch hohen Abnormitätsindizes in Verbindung gebracht werden können. Hohe Windgeschwindigkeiten tragen eher zu hohen Abnormitätsindizes bei, wogegen hohe Werte des Peakedness Parameters oder der Kurtosis den vorhergesagten Abnormitätsindex eher zu reduzieren scheinen.

Insgesamt sind die Ergebnisse nicht durchgehend unmittelbar plausibel und deuten darauf hin, dass bei der Entstehung von Rogue Waves in flachem und in tiefem Wasser unterschiedliche Prozesse eine Rolle spielen und dass nichtlineare Effekte das Auftreten von Rogue Waves modulieren können. Die Ergebnisse sind insofern als vorläufig zu betrachten und benötigen weiterer Analysen vorzugsweise mit einem umfangreicheren Datenmaterial.



**Abb. 16:** Lokale Merkmalsbedeutung für die ausgewählten Schlüsselmerkmale unter Verwendung von SHAP für das reduzierte neuronale Netzwerk. Der Balken stellen die Verteilung der normalisierten Merkmalswerte dar, wobei hohe/niedrige Merkmalswerte in rot/blau dargestellt sind.

## 4 Zusammenfassung

Die wesentlichen Ergebnisse des Projekts FREAK WAVES II lassen sich wie folgt zusammenfassen:

1. Bei der Mehrzahl der in dem untersuchten Datensatz identifizierten Rogue Waves scheint es sich um seltene Realisierungen typischer Wellenhöhenverteilungen zu handeln (Ergebnis aus FREAK WAVES I).

2. In mittleren und größeren Wassertiefen ist die Modulations-Instabilität („modulational instability“) ein nichtlinearer Mechanismus, der das vermehrte Auftreten von Rogue Waves erklären kann. Modulations-Instabilitäten treten vor allem dann auf, wenn die Wellen ähnliche Frequenzen und Richtungen haben. An den in FREAK WAVES II in der südlichen Nordsee untersuchten Stationen sind diese Voraussetzungen in der Regel nicht erfüllt. Die Modulations-Instabilität ist deshalb vermutlich nicht der entscheidende Prozess, um ein erhöhtes Auftreten von Rogue Waves in der südlichen Nordsee zu erklären.
3. Im Flachwasser spielen nichtlineare Prozesse wie die Wechselwirkung von Solitonen mit dem periodischen Teil des Spektrums vermutlich eine entscheidende Rolle bei der Entstehung von Rogue Waves und können eine Erklärung für die erhöhte Häufigkeit von Rogue Waves an den untersuchten Flachwasserstationen darstellen.
4. Anhand des in FREAK WAVES II verwendeten mehrjährigen Datensatzes konnte gezeigt werden, dass insbesondere in sehr breitbandigen Seegangsspektren mit hoher Richtungsstreuung bei hohen Werten des Verhältnisses zwischen Gezeitenstrom und Gruppengeschwindigkeit eher eine symmetrische Modulation verschiedener Seegangparameter einschließlich der Häufigkeit von Rogue Waves zu beobachten war.
5. Methoden des maschinellen Lernens und deren Interpretierbarkeit können Potenzial für die Kurzzeitprognose und das Verständnis der zugrunde liegenden Prozesse bei der Entstehung von Rogue Waves besitzen. Eine eingehendere Betrachtung erfordert insbesondere den Einsatz eines umfangreicheren Datenmaterials.

## 5 Ausblick

Der in FREAK WAVES II verwendete Datensatz umfasst derzeit Daten von 11 Messstationen (Bojen, Radar) in der südlichen Nordsee für die Jahre 2011-2016 mit Messfrequenzen von 1.28-4.00 Hz. Während dieser Datensatz eine gute Grundlage für erste Untersuchungen wie der Erstellung einer Klimatologie lieferte, ist ein größerer Datenumfang insbesondere für die Anwendung von Ansätzen des maschinellen Lernens wünschenswert, da hierfür der Stichprobenumfang beobachteter Extremwellen im bisher zur Verfügung stehenden Datenmaterial nach wie vor zu gering ist.

Da sich bei den Auswertungen des ursprünglichen Datensatzes im Projekt FREAK WAVES I Hinweise auf mögliche Abhängigkeiten der Ergebnisse von verschiedenen Messsystemen und -frequenzen (Boje, Radar) ergeben haben, wäre eine Analyse von Daten von zeitgleich vorhandenen und dicht beieinanderliegenden unterschiedlichen Messsystemen (Boje, Radar) zielführend, um die Auswirkungen besser zu verstehen.

Obwohl der in FREAK WAVES II verwendete Datensatz mit einer Million Wellen bereits ungewöhnlich groß war, erwies er sich für Ansätze des maschinellen Lernens als nach wie vor zu klein. Auch der Einfluss von Parametern, die sich langsam ändern (z.B. objektive Wetterlagen) konnte damit nicht abschließend beurteilt werden. Es bleibt deshalb interessant mit Hilfe eines größeren Datensatzes zu untersuchen, inwieweit die in FREAK WAVES II entwickelten Methoden potenziell für eine automatische Detektion von Extremwellen in Messungen eingesetzt werden können und inwieweit sich anhand von Messungen unmittelbar vor dem Auftreten von Extremwellen bereits auf das nachfolgende Auftreten der Extremwelle schließen lässt. Wenn erfolgreich, könnte eine solche prototypische automatisierte schnelle Detektion (z.B. mit Bojen-

und Radarmessungen) in Offshore Windparks genutzt und für Kurzfristwarnungen eingesetzt werden.

In FREAK WAVES II wurde gezeigt, dass in Flachwasserbereichen in der südlichen Nordsee das Auftreten von Extremwellen oft bei einer bestimmten Form des Soliton Spektrums erfolgte. Umgekehrt konnte aus der Beobachtung dieser Form des Soliton Spektrums aber nicht eindeutig auf das Vorhandensein einer Extremwelle geschlossen werden. Letzteres wäre Voraussetzung, damit z.B. Vorhersagen anhand der Form des Spektrums getroffen werden können. Eine mögliche Ursache könnte darin liegen, dass im Seegebiet zwar kurzzeitig eine Extremwelle vorhanden war, aber diese mit Hilfe von nur einer Punktmessung (Boje) nicht erfasst wurde. Diese Hypothese könnte z.B. mit Hilfe parallel betriebener Messsysteme, insbesondere von flächenhaften Seegangs-Radaren und Bojen, untersucht werden. Eine weitere Möglichkeit besteht darin, mittels numerischer Propagation z.B. mit Hilfe der KdV-Gleichung zu untersuchen, inwieweit sich Extremwellen in der Umgebung einer Punktmessung bilden können.

## 6 Literaturverzeichnis

- Alber, I. E. (1978), The effects of randomness on the stability of two-dimensional surface wavetrains, *Proc. R. Soc. Lond. A* 363(1715), 525–546, DOI: 10.1098/rspa.1978.0181.
- Benjamin, T. B., and J. E. Feir (1967), The disintegration of wave trains on deep water Part 1. Theory, *J. Fluid Mech.* 27(3), 417–430, DOI: 10.1017/s002211206700045x.
- Bitner-Gregersen, E.M., and O. Gramstad (2016), Rogue Waves. Impact on ship and offshore structures., *Strategic Research & Innovation Position Paper 05-2015*, DNV GL, <https://www.dnvgl.com/technology-innovation/rogue-waves/>.
- Bühl, M., and H. Oumeraci (2016), Analysis of long-period cosine-wave dispersion in very shallow water using nonlinear Fourier transform based on KdV equation, *Applied Ocean Research* 61, 81–91, DOI: 10.1016/j.apor.2016.09.009.
- Dysthe, K., H. E. Krogstad, and P. Müller (2008), Oceanic Rogue Waves, *Annu. Rev. Fluid Mech.* 40(1), 287–310, DOI: 10.1146/annurev.fluid.40.111406.102203.
- Forristall, G. Z. (1978), On the statistical distribution of wave heights in a storm, *J. Geophys. Res.* 83(C5), 2353–2358, DOI: 10.1029/jc083ic05p02353.
- Gramstad, O., and K. Trulsen (2007), Influence of crest and group length on the occurrence of freak waves, *J. Fluid Mech.* 582, 463–472, DOI: 10.1017/s0022112007006507.
- Häfner, D., J. Gemmrich, and M. Jochum (2021), Real-world rogue wave probabilities, *Scientific reports* 11(1), 10084, DOI: 10.1038/s41598-021-89359-1.
- Hjelmervik, K. B., and K. Trulsen (2009), Freak wave statistics on collinear currents, *J. Fluid Mech.* 637, 267–284, DOI: 10.1017/S0022112009990607.
- Janssen, P. A. E. M. (2003), Nonlinear Four-Wave Interactions and Freak Waves, *J. Phys. Oceanogr.* 33(4), 863–884, DOI: 10.1175/1520-0485(2003)33%3C863:NFIAPW%3E2.0.CO;2.
- Janssen, P. A. E. M., and J.-R. Bidlot (2009), On the extension of the freak wave warning system and its verification, *Technical Memorandum 588*, ECMWF, <https://doi.org/10.21957/uf1sybog>.
- Karmpadakis, I., C. Swan, and M. Christou (2019), Laboratory investigation of crest height statistics in intermediate water depths, *Proceedings. Mathematical, physical, and engineering sciences* 475(2229), 20190183, DOI: 10.1098/rspa.2019.0183.
- 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., *Lond. Edinburgh Dublin Philos. Mag. and J. Sci.* 39, 422–443.
- Lake, B. M., H. C. Yuen, H. Rungaldier, and W. E. Ferguson (1977), Nonlinear deep-water waves, Theory and experiment. Part 2. Evolution of a continuous wave train, *J. Fluid Mech.* 83(1), 49–74, DOI: 10.1017/s0022112077001037.
- Mendes, S., I. Teutsch, and J. Kasparin (2024), Symmetric influence of forward and opposing tidal currents on rogue wave statistics., *J. Fluid Mech.*, [Under Review, Preprint, arXiv:2408.06371], DOI: 10.48550/ARXIV.2408.06371.
- Onorato, M., A. R. Osborne, M. Serio, L. Cavaleri, C. Brandini, and C. T. Stansberg (2004), Observation of strongly non-Gaussian statistics for random sea surface gravity waves in wave

flume experiments, *Physical review. E, Statistical, nonlinear, and soft matter physics* 70(6 Pt 2), 67302, DOI: 10.1103/physreve.70.067302.

Sergeeva, A., E. Pelinovsky, and T. Talipova (2011), Nonlinear random wave field in shallow water, Variable Korteweg-de Vries framework, *Nat. Hazards Earth Syst. Sci.* 11(2), 323–330, DOI: 10.5194/nhess-11-323-2011.

Stansberg, C. T. (1995), Effects from directionality and spectral bandwidth on non-linear spatial modulations of deep-water surface gravity wave trains., *Coastal Engineering* 1994, DOI: 10.1061/9780784400890.044.

Teutsch, I., M. Brühl, R. Weisse, and S. Wahls (2023), Contribution of solitons to enhanced rogue wave occurrence in shallow depths, A case study in the southern North Sea, *Nat. Hazards Earth Syst. Sci.* 23(6), 2053–2073, DOI: 10.5194/nhess-23-2053-2023.

Teutsch, I., S. Mendes, and J. Kasparin (2024a), Direction Symmetry of Wave Field Modulation by Tidal Current, *J. Phys. Oceanogr.*, [Under Review, Preprint, arXiv:2407.04723], DOI: 10.48550/arXiv.2407.04723.

Teutsch, I., and R. Weisse (2023), Rogue Waves in the Southern North Sea—The Role of Modulational Instability, *J. Phys. Oceanogr.* 53(1), 269–286, DOI: 10.1175/JPO-D-22-0059.1.

Teutsch, I., R. Weisse, J. Moeller, and O. Krueger (2020), A statistical analysis of rogue waves in the southern North Sea, *Nat. Hazards Earth Syst. Sci.* 20(10), 2665–2680, DOI: 10.5194/nhess-20-2665-2020.

Teutsch, I., R. Weisse, and S. Wahls (2024b), Brief communication, Implications of outstanding solitons for the occurrence of rogue waves at two additional sites in the North Sea, *Nat. Hazards Earth Syst. Sci.* 24(6), 2065–2069, DOI: 10.5194/nhess-24-2065-2024.