Reply to review #1

We thank Referee #1 for the constructive comments that have helped us to clarify and improve many aspects of our manuscript. In the following, we explain how we plan to address the individual issues raised by the reviewer in the revised manuscript.
In blue color, the lines, in which changes in connection with the respective reviewer comment have been made, are indicated. These refer to the *tracked changes version.*

1. General Comments

1A We thank the reviewer for this comment. We will add a proper definition of rogue waves and discuss their relevance, as suggested. (line 2, lines 25-29)

1B We thank the reviewer for pointing out this ambiguity in the manuscript. We will draw a clear connection of Rayleigh and Weibull distributions to linear superposition and add an additional reference for second-order models to the text.
(lines 2/3, 30-33, 181/182, 470, 599)

1C We appreciate the critical review of the paragraph. We agree that some formulations are misleading and we will re-formulate the passage in agreement with the suggestions of the reviewer. (lines 37-45)

1D The word "respective" will be removed, as suggested by the reviewer. (line 43)

1E We thank the reviewer for pointing out additional important conclusions in the literature. We will add these to our introduction. (lines 49-53)

1F We agree with the reviewer. We will add the suggested references accordingly.
(lines 46-49)

1G We agree with the reviewer. We will add the suggested references accordingly.
(lines 90-92)

1H We understand the objection of the reviewer. In our text, we will replace the expression "nonlinearity" by "nonlinear processes", where appropriate. (lines 73/74, 104, 120, 364)

1I We thank the reviewer for the additional reference. We will add it accordingly.
(lines 114/115)

1J We agree with the reviewer on the contradiction in the text. We will rephrase the sentence in question accordingly. (lines 122/123)

1K We thank the reviewer for several additional important references concerning rogue wave occurrence on varying bathymetry. We will extend the discussion and include the suggested references. (lines 491, 495-508)

1L We thank the reviewer for this remark. We will replace the unpublished reference by publically available literature. (line 380)

2. Scientific Improvement

2A We thank the reviewer for this detailed assessment and the comparison with the previous study of Teutsch et al. (2020). The reviewer is completely right that Figures 7 and 9 in Teutsch et al. (2020) do not support the conclusion that rogue wave frequencies are overestimated by the Forristall distribution. However, in these Figures data from different types of instruments (radar and wave buoys) are considered jointly. Figure 2 in Teutsch et al. (2020) clearly shows that data from both instruments show different behaviour with rogue wave frequencies in the radar/buoy data set being higher/lower than that derived from the Forristall distribution. The only exception here were the results from the buoy SEE off Norderney, which showed results comparable to those derived from radar data. This rendered the station SEE outstanding and provided the motivation for this study.

2B We thank the reviewer for his/her discussion of the Ursell number formulation. As suggested, we will rewrite the equation and state in the text that different definitions of the Ursell number exist, which will lead to different threshold values. (lines 366-385)

2C For the agreement between the Forristall distribution and the measurements, please see our reply to comment 2A. In l. 369 of the discussion, we will add the information that the "nearby stations" were buoy stations as well, to avoid confusion of buoy and radar measurement results from the previous study.
Regarding the second part of the comment, we acknowledge that there is a clear definition of what represents shallow/deep/intermediate water for a wave, while the terms were used here in a broader sense to distinguish sites. We will revise the manuscript to make this clearer and we will use a different terminology. We will also replace second-order theory by Forristall distribution. (lines 216/217, 468/469)

2D The mentioned citation refers to non-Gaussianity in decreasing water depth, here in the context of wave run-up. The article is referred to by Sergeeva et al. (2011) to emphasise nonlinear behaviour of waves above a varying bathymetry. We do, however, agree with the reviewer that our formulation concerning the reference is misleading, as the referred article does not concern rogue waves. We will therefore re-formulate the sentence to read "... described e.g. by Huntley et al. (1977) *in the context of wave run-up. It has* gained increased attention in the context of rogue wave occurrence *(e.g. Sergeeva et al.* (2011)*)*".
(lines 478-489)

3. Scientific Issues

3A We thank the reviewer for raising the issue of the shallow water definition. We will present the ranges of kh in our data, as well as the value of the slope, as suggested. The reviewer is right that our article does not solely concern waves in shallow water as defined by $h < L/20$. Since we investigate waves in the context of the KdV equation, we follow the definition of the applicability of the KdV equation as given by Osborne and Petti (1994), p. 1731, and Osborne (1995), p. 2629. We acknowledge that, therefore, our definitions of shallow water ($kh < 1.36$ or $h/L < 0.22$) and deep water ($kh > 1.36$ or $h/L > 0.22$) are different from the definitions of shallow (h/L < 0.05), intermediate (0.05 < h/L < 0.5) and deep water (h/L > 0.5) that are used in the engineering context, and that this may lead to confusion. We will therefore state this difference more clearly in the text, and we will clearly define the terminology for 'shallow' that is used in the paper. Furthermore, we will change the title to include "shallow depths" instead of "shallow water", as suggested by the reviewer. (lines 106-108, 129/130, 203-205, 215-219, 227, 230-233, 234, 472-474, Table 3, 521-525, 588)

3B We apply the definition of the KdV equation as given in Osborne (2010) , p. 9, which defines the linear phase speed as $c_0 = \sqrt{gh}$. We agree that the term 'shallow-water wave celerity' is misleading when applying the KdV equation to relative depths larger than h/L = 0.05. Nevertheless, the linear phase speed is used in the KdV equation within the range of applicability. We will use the term 'linear phase speed' instead. (line 223)

3C We thank the reviewer for pointing out that the category "normal" does not account for the fact that waves slightly below the threshold H/H_s = 2.0 are influenced by nonlinear processes and can become highly dangerous, similarly to waves with H/H_s > 2.0. We will change the category "normal" to "non-rogue", to emphasise that these samples do not include waves according to the definitions $H/H_s \geq 2.0$ or $C/Hs \geq 1.25$. (entire document)

3D We agree with the reviewer that the number of samples is a vague quantity for the reader and that the total number of waves should be more informative. We will include the precise number of measured waves in Table 1. (Table 1)

3E We thank the reviewer for the valuable suggestion that the results from the crucial papers Osborne et al. (1991) and Bruehl and Oumeraci (2016) should be explained to the reader more thoroughly. For explanation, we will additionally refer to earlier crucial references, the original numerical studies by Zabusky and Kruskal (1965) and Osborne & Bergamasco (1986). Osborne et al. (1991) applied the approach to ocean measurement data and Bruehl and Oumeraci (2016) performed an experimental study. Note that there is a technical difference between our approach and the approach in the three last mentioned works. We use the NLFT for vanishing boundary conditions, while Osborne & Bergamosco (1986), Osborne et al. (1991) and Bruehl and Oumeraci (2016) apply the NLFT for periodic boundary conditions. The relation between the two transforms is somehow similar to the relation between the linear Fourier transform and the linear Fourier series (which is what the FFT computes). They are related, and both have their advantages and disadvantages, but not all results can be directly compared for this reason. We will add paragraphs accordingly, additionally referring to earlier crucial references that explain the behaviour of solitons for the NLFT employed in our paper, like Hammack and Segur (1974) and Ablowitz & Kodama (1982). In our context, the sea surface elevation is described by a discrete spectrum indicating solitons and a continuous spectrum indicating a dispersive wave train. Of these two parts, we only discuss the soliton spectrum further in this article. However, it is known that for vanishing boundary conditions the soliton spectrum completely describes the behaviour of the wave train in the far field. After the complete dissipation of the dispersive waves, only the solitons are left in the far field. When the distance between these solitons is sufficiently large, no interactions occur between them and all solitons are clearly visible with their characteristic shapes. Assuming frictionless propagation, their amplitudes can already be read from the nonlinear spectrum of the initial time series. Therefore, we prefer to add the equation for the surface elevation in the far field, resulting from the solitons, given e.g. by Equation (4) in Prins & Wahls (2019), with reference

to Schuur (1984), Eq. 17, Schuur (1986), p. 83, Eq. 33 and Ablowitz & Kodama (1982), Eq. 2.20a.
We further agree to add plots that explain the meaning of the soliton spectrum (Figure 1 in this response). In addition to an exemplary time series (blue line in the first plot, zoomed in to the rogue wave), we will add its linear FFT spectrum (second plot) and, in addition to the soliton spectrum (last plot), the nonlinear continuous spectrum (third plot, which will not be analysed further in this article). Each of the solitons in the soliton spectrum would be a physical soliton if the signal is propagated according to the KdV equation. After sufficiently long propagation, each of these solitons will appear isolated with its characteristic shape. Within the time series, the solitons are close together; they overtake and interact with each other. For visualisation of the role of the solitons in the time series, the first plot shows the soliton train (red line) that is obtained by nonlinear superposition of the solitons (considering their interactions) using the algorithm from Prins & Wahls (2021). Note that inverting large soliton spectra is numerically very difficult (Prins & Wahls, 2021). We therefore had to use a shortened time series for the figure. (lines 139-146, 152, 175-177, 183-185, 275-281, 284-290, Figure 5)

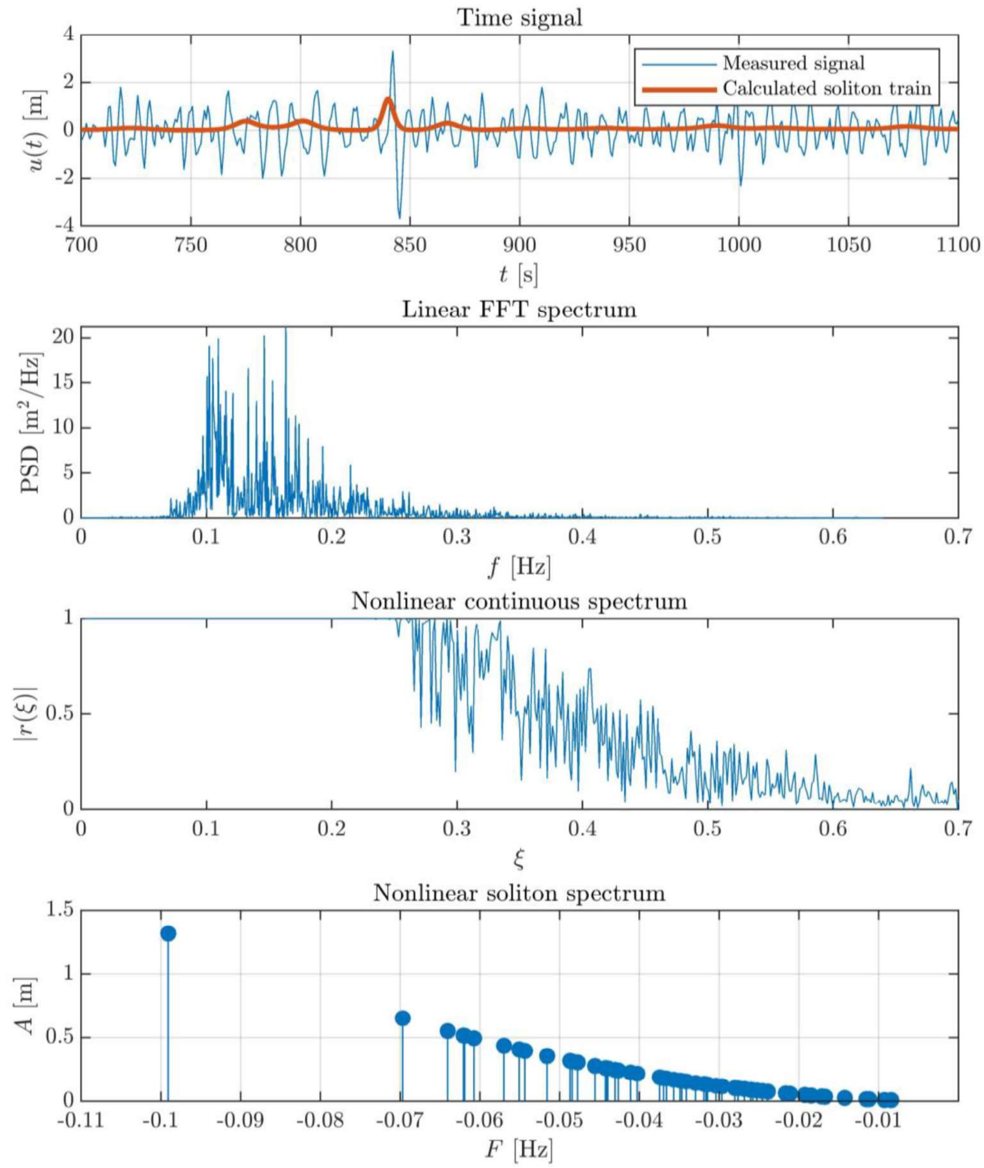

Figure 1: a) time series and soliton train, as calculated from the inverted soliton spectrum. b) linear Fourier spectrum. c) continuous spectrum. d) discrete soliton spectrum.

3F The reviewer is right in that the NLFT is currently employed as a trace method. The reason is that we do not know how the nonlinear spectrum changes during propagation around the buoy. If the KdV equation describes the propagation around the buoy reasonably well, then the nonlinear spectrum would be approximately constant during propagation, and the method could single out certain time series that lead (or led) to extreme rogue waves. This aspect however requires more research that is beyond the scope of the current paper. Therefore, we use the approach as a trace method, which has the additional potential to provide further information in future applications. Our work is nevertheless a necessary step in the direction of recognising potentially dangerous time series. If the method does not work for visible rogue waves, there is little hope for it to work for hidden rogue waves. We demonstrate for the first time that certain distinctive patterns in the nonlinear spectrum of real-world time series indicate extreme rogue waves (at a specific measurement site). Finally, we note that even if the KdV does not describe propagation well, the NLFT could still be a better transform to analyse data in this area than the linear Fourier transform (where the spectrum also only develops in a simple way if the propagation is linear and the depth is constant, but it is nevertheless applied in different contexts). We will discuss these points in the revised manuscript. (lines 186/187, 544-552)

3G

- As suggested, we will support the explanation of the results by describing the insights from Bruehl and Oumeraci (2016) and the relationship between the soliton spectrum and the far field behaviour under KdV, in a meaningful connection with our reply to 3E.
- The use of the term “determined” was supposed to imply that the soliton was determined by the use of NLFT. For clarity, we will replace the terms “determined” and “specific” with “individual”.
- The reviewer criticises that scaling down a rogue wave to 80% is not linked to any physical explanation. We would like to reply that an established method for the treatment of NLFT spectra does not exist. In contrast to the linear case, where the impact of a window on the spectrum can be expressed analytically in a way that is easy to interpret, no such result is known for the nonlinear case. Windowing of the time series and calculating separate nonlinear spectra is common, but does not have any theoretical grounding. In contrast to windowing, which is a general purpose technique that impacts large parts of the time series, our method is local and aims specifically at rogue waves. By scaling only the rogue wave, the changes in the time series are as small as possible. The hope is thus that the danger of evoking additional, unrelated changes in the nonlinear spectrum is minimised by this approach.
  We intend to localise the influence of a change in rogue wave height in the soliton spectrum to establish a connection between a measured rogue wave and individual solitons. The underlying idea of the method is that if local changes of the rogue wave lead to local changes in the spectrum, the changing soliton components are associated with the rogue wave. Since a rogue wave is a particular wave event, it is reasonable to explore changes in the spectrum when only this wave is changed. Furthermore, also the (hidden) solitons in the data are localised components and changes to the particular rogue-wave event are expected to have effects to the soliton spectrum only when a soliton is located sufficiently close to the modified region. The changes in the soliton spectrum only affect a few solitons, whereas all other solitons remain constant. Since only a few solitons are modified, we can conclude that these solitons are located in the modified rogue-wave region within the time series. Regarding the request to remove the rogue wave from the time series, we prefer reducing its height as opposed to cutting it out, which would introduce an artificial gap to the time series. The method shows that gradually reducing the height of the rogue waves leads to the gradual reduction of individual solitons. A change of the rogue wave will not have an impact on all soliton components in the spectrum. By this straightforward approach, solitons that are directly linked to the rogue wave are easily identified.
- Solitons linked to the rogue wave are not always the largest solitons in the spectrum. As also pointed out in the manuscript, the soliton alone is not sufficient to explain the rogue wave. Only by interaction with components from the continuous spectrum, the rogue wave is formed. In order for this to happen, the soliton and the other components must interact constructively. When the interactions are not constructive, it is very well possible that a larger soliton leads to a smaller hump in the time series. Hence, the dispersive waves and nonlinear interactions have a strong impact, and the largest soliton is not necessarily associated with the largest wave in the time series. For a visual illustration, we would like to refer to Figure 1 (a) in

Osborne et al. (1991), in which the largest soliton is also not associated with the highest wave in the time series. In contrast, in our example given above, the largest soliton is located close to the position of the rogue wave.

- Against this background, changing Figure 8, in which rogue wave heights are compared with the associated solitons, would not make sense. Comparing with the highest soliton in the spectrum would include some solitons that are linked to wave groups without rogue waves. While we do not plan to change this in the paper, we nevertheless followed the suggestion of the reviewer and calculated A_{max} with respect to H_{max}. We present the updated Figure 8 in comparison with the original Figure 8 below (Figure 2-Figure 4). Here, grey dots show cases, in which the highest attributed soliton is identical with the maximum soliton in the discrete spectrum ($A_s^1 = A_{max}$). For rogue samples, this is true in most cases (extreme: 87%, double: 85%, crest 78%, height 71%). For non-rogue samples, this is true in 42% of the cases. The figures show that the results are in a comparable range when rogue wave (or maximum wave) heights are related to the maximum instead of the highest attributed soliton.

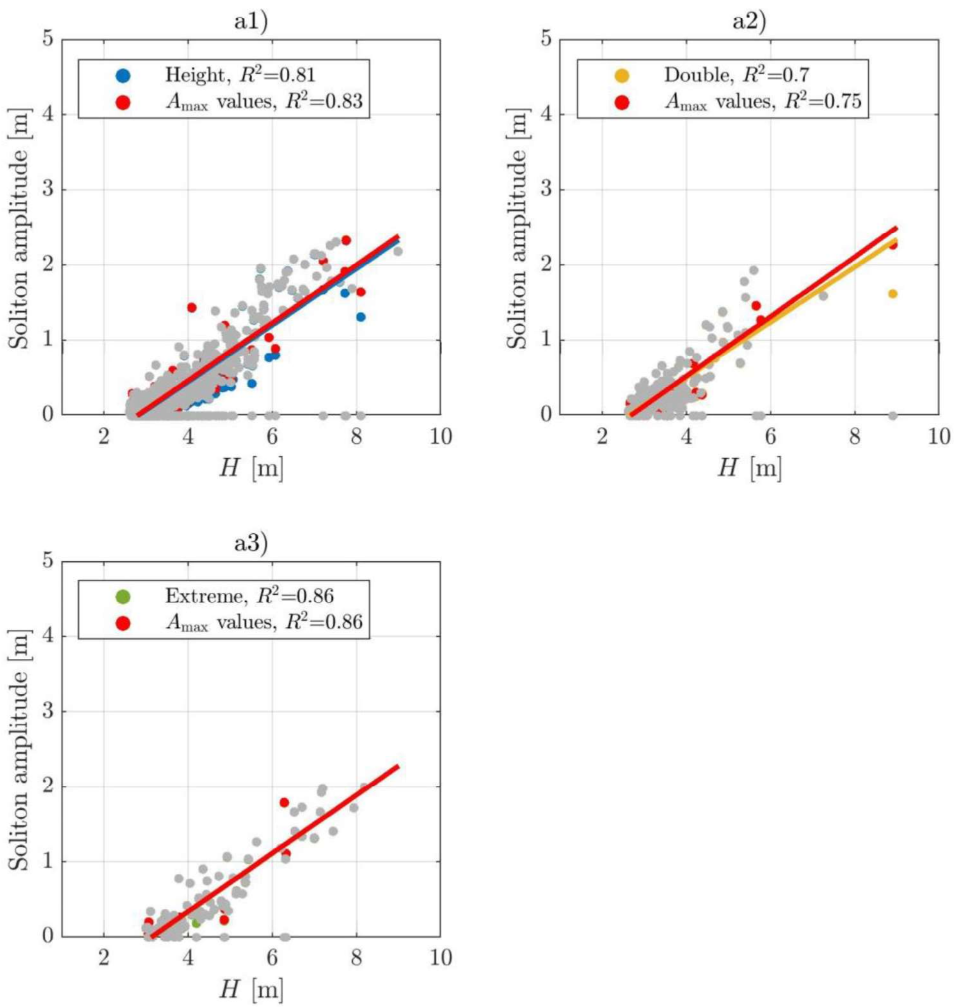

Figure 2: Update of Figure 8a in the preprint. Blue, yellow and green dots, in the legends referred to as "Height", "Double" and "Extreme", represent values from the original figures. When using maximum soliton amplitudes instead of the

amplitudes of the attributed solitons, these values change to values represented by red markers. Grey markers show values of attributed soliton amplitude that are identical with the maximum soliton amplitude.

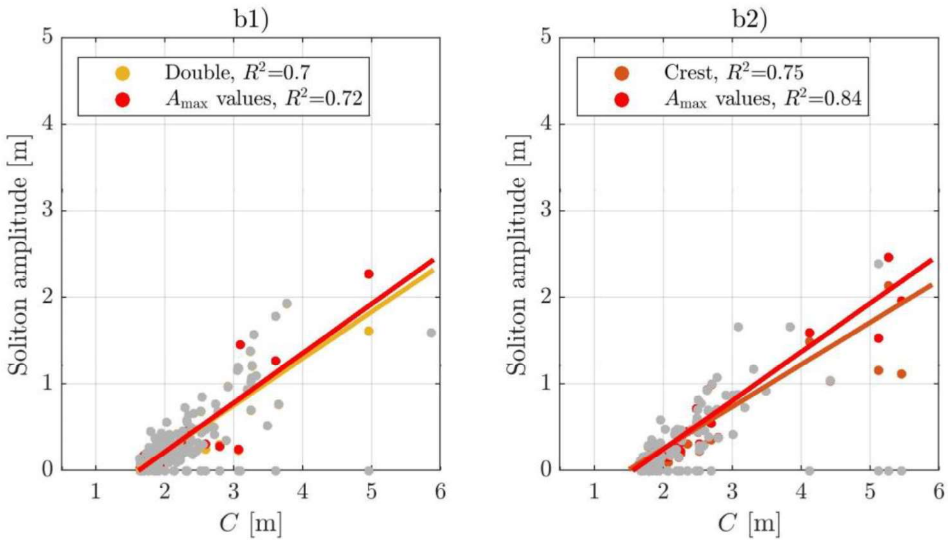

Figure 3: Update of Figure 8b in the preprint. Yellow and orange dots, in the legends referred to as "Double" and "Crest", represent values from the original figures. When using maximum soliton amplitudes instead of the amplitudes of the attributed solitons, these values change to values represented by red markers. Grey markers show values of attributed soliton amplitude that are identical with the maximum soliton amplitude.

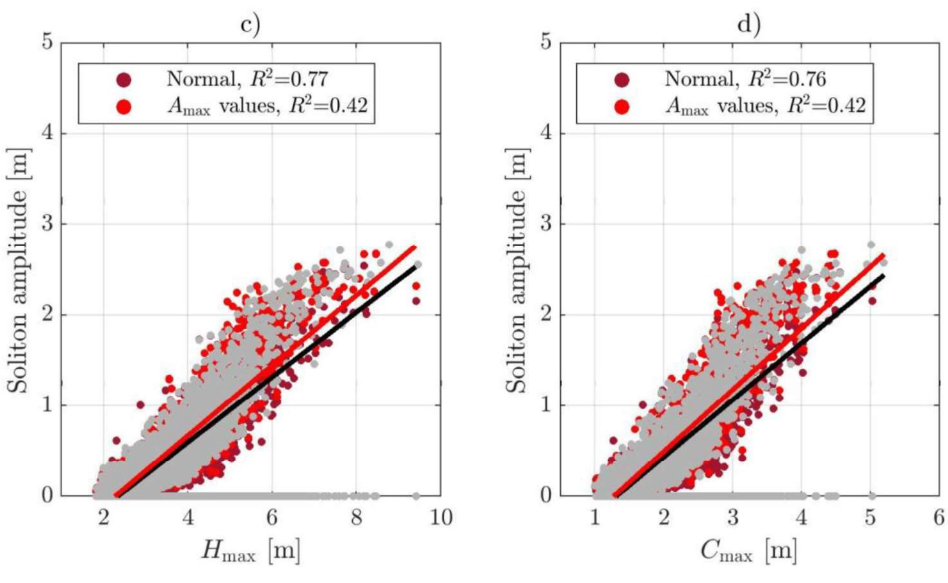

Figure 4: Update of Figures 8c and 8d in the preprint. Dark red dots, in the legends referred to as "Normal", represent values from the original figures. When using maximum soliton amplitudes instead of the amplitudes of the attributed solitons, these values change to values represented by red markers. Grey markers show values of attributed soliton amplitude that are identical with the maximum soliton amplitude.

- In our study, soliton amplitudes were always smaller than rogue wave crests, which is in agreement with Figure 1a in Osborne et al. (1991).
- We understand that the formulation “To remove the influence of the underlying sea state” in line 261 is misleading. We agree with the reviewer that the influence of the sea state is not only characterised by the significant wave height. Our intention with normalising by H_s is to create dimensionless values, to be able to compare different samples. Since rogue waves are defined on the basis of the significant wave height, we find this parameter suitable for the normalization. The influence of the sea state in terms of the parameters suggested by the reviewer (steepness, Ursell number, *kh*, bandwidth) affect the wave components in the continuous part of the nonlinear spectrum, which we do not discuss further in this article. The soliton spectrum is not affected. Through the continuous spectrum, the sea state parameters possibly influence the wave distribution. We have tested the method that was suggested by the reviewer. As an example, Figure 5 shows the exceedance probability of H/H_s in all samples of a defined Ursell number range. It is seen that the distribution behaves differently in the different ranges. However, this cannot be stated for certain, as the results rely on few data, due to the binning into ranges. The few rogue waves in the samples are distributed randomly, which leads to uncertain results. Together with the consideration that the sea state may indeed affect the continuous part of the spectrum, and thus the probability distribution may change with the sea state parameters, we have come to the conclusion that it is not possible to deduce the influence of solitons from these exceedance probability plots. Therefore, we have decided not to include the plots in the revised article.

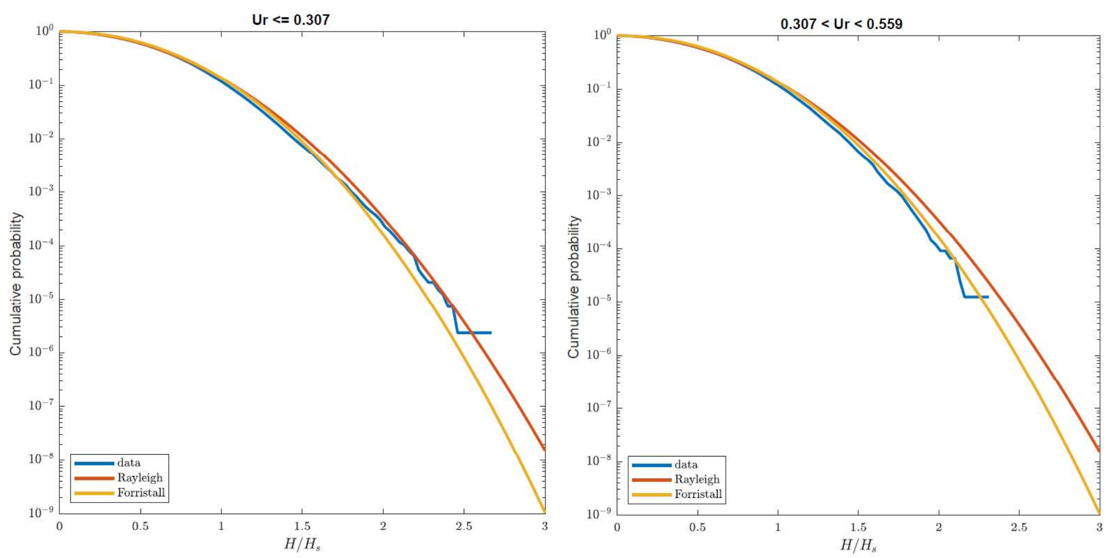

Figure 5: Cumulative exceedance probability of H/H_s in all samples that belong to the category of $0.2 < A_1^S/H_s \leq 0.3$, and two different Ursell number ranges.

(lines 90-92, 183-185, 275-281, 300/301, 345/346)

3H We would like to point out that we do not intend to explain a rogue wave by one soliton alone (as can be seen by comparison of the free-surface elevation ad and the soliton train in Figure 1 in this document). Our hypothesis is that solitons *contribute* to the formation of rogue waves, and we have shown that there is/are always one or several solitons involved when a rogue wave is present in the sample. However, the surface elevation is described not only by solitons, but also by dispersive waves and by the interaction of wave components. This statement is supported by Figure 10 in the preprint, which we therefore would like to keep in the article. Bruehl et al. (2016) have shown that soliton-like waves can form waves that seem to have linear shapes. (lines 152, 475-485)

3I The soliton spectrum alone cannot reveal the formation of rogue waves in general, but solitons may be directly attributed to rogue waves and as such are involved in their presence. We would like to emphasise that it is actually the first time that this has been verified by the nonlinear Fourier analysis of real-world data. Furthermore, as discussed later in the manuscript in relation to Figure 15, certain configurations of the soliton spectrum (A_2/A_1) actually do indicate the presence of rogue waves with high probability. The attribution is shown by the method discussed in comment 3G, which we would like to retain in the paper. Furthermore, the size of a soliton is not sufficient to explain the height and the shape of a rogue wave all by itself (see for example Figure 1). As suggested by the reviewer, we will transfer these issues to the discussion of the article to present the line of argument in a straight order. (lines 336/337, 338-342, 362-363, 391-393, 475-485)

3J The soliton gap was chosen after studying the soliton spectra of many rogue samples. It is not arbitrary, and as far as we know there are no existing alternatives from the literature that would have fit our context. Please also note that many existing signal processing tools heavily rely on the linearity of the transform, and applying them with a nonlinear transform in general is meaningful only in the quasi-linear regime. We do not expect the variance of the soliton spectrum to be a better tool because it would involve solitons that are not associated with the rogue wave. We agree that the conclusion should be that outstanding solitons are not good indicators of rogue waves or large waves near the rogue wave threshold of H/H_s = 2.0, but only for extreme rogue waves. We will avoid the term “predictor” in the text, as the soliton spectrum becomes available only after the recording of the time series and the occurrence of the rogue wave (see also reply to comment 3F). (lines 434, 599)

3K The remark is correct, we will adjust the text accordingly, referring to the Forristall distribution (see replies on comments 1B and 2C). (lines 592/593)

3L We confirm that we have not investigated the interaction with oscillatory waves in the context of rogue wave formation and that we are therefore not entitled to make a statement on the exact nature of the interactions. We agree that we should mention this in the conclusions, so as not to raise wrong expectations with the reader. From Figure 1, it is seen that the soliton train alone does not account for the full height of the rogue wave. This only leaves the continuous spectrum for the explanation of the missing height. (lines 606/607, Figure 5)

References

Ablowitz, M. J. & Kodama, Y., 1982. Note on Asymptotic Solutions of the Korteweg-de Vries Equation with Solitons. *Studies in Applied Mathematics,* 66(2), pp. 159-170.

Brühl, M. & Oumeraci, H., 2016. Analysis of long-period cosine-wave dispersion in very shallow water using nonlinear Fourier transform based on KdV equation. *Applied Ocean Research,* Volume 61, pp. 81-91.

Hammack, J. L. & Segur, H., 1974. The Korteweg-de Vries equation and water waves. Part 2. Comparison with experiments. *Journal of Fluid Mechanics,* 65(2), pp. 289-314.

Huntley, D. A., Guza, R. T. & Bowen, A. J., 1977. A universal form for shoreline run-up spectra?. *Journal of Geophysical Research,* Volume 82, pp. 2577-2581.

Osborne, A. R., 1995. The inverse scattering transform: Tools for the nonlinear Fourier analysis and filtering of ocean surface waves. *Chaos Solitons Fractals,* 5(12), pp. 2623-2637.

Osborne, A. R., 2010. *Nonlinear ocean waves and the inverse scattering transform.* Amsterdam: Elsevier.

Osborne, A. R. & Bergamasco, L., 1986. The solitons of Zabusky and Kruskal revisited: Perspective in terms of the periodic spectral transform. *Physica D: Nonlinear Phenomena,* 18(1-3), pp. 26-46.

Osborne, A. R. & Petti, M., 1994. Laboratory-generated, shallow-water surface waves: Analysis using the periodic, inverse scattering transform. *Phys. Fluids,* 6(5), pp. 1727-1744.

Osborne, A. R., Segre, E., Boffetta, G. & Cavaleri, L., 1991. Soliton basis states in shallow-water ocean surface waves. *Physical Review Letters,* Volume 67, pp. 592-595.

Prins, P. J. & Wahls, S., 2019. Soliton Phase Shift Calculation for the Korteweg–De Vries Equation. *IEEE Access,* Volume 7, pp. 122914-122930.

Prins, P. J. & Wahls, S., 2021. An accurate O(N2) floating point algorithm for the Crum transform of the KdV equation. *Communications in Nonlinear Science and Numerical Simulation,* Volume 102, p. 105782.

Schuur, P., 1984. Multisoliton phase shifts in the case of a nonzero reflection coefficient. *Phys. Lett. A,* 102(9), pp. 387-392.

Schuur, P. C., 1986. *Asymptotic Analysis of Soliton Problems: An Inverse Scattering Approach.* 1232 ed. (Lecture Notes in Mathematics): A. Dold and B. Eckmann, Eds. New York, NY, USA: Springer-Verlag.

Sergeeva, A., Pelinovsky, E. & Talipova, T., 2011. Nonlinear random wave field in shallow water: variable Korteweg–de Vries framework. *Natural Hazards and Earth System Sciences,* 11(2), p. 323–330.

Teutsch, I., Weisse, R., Moeller, J. & Krueger, O., 2020. A statistical analysis of rogue waves in the southern North Sea.. *Natural Hazards and Earth System Sciences,* 20(10), p. 2665–2680.

Zabusky, N. J. & Kruskal, M. D., 1965. Interaction of "Solitons" in a Collisionless Plasma and the Recurrence of Initial States. *Physical Review Letters,* Volume 15, pp. 240-243.

Reply to review #2

We thank Referee #2 for the constructive comments that will help us to clarify and improve several points in our manuscript. In the following, we explain how we plan to address the individual issues raised by the reviewer in the revised manuscript.
In blue color, the lines, in which changes in connection with the respective reviewer comment have been made, are indicated. These refer to the *tracked changes version.*

1. On the KdV approximation

- We fully agree with the reviewer that the assumption that KdV is valid, “cannot be justified on the basis of single-point measurements” (nhess-2022-28-RC2-supplement, p. 1). We would like to point out that we applied vKdV-NLFT as a signal transform, similar to e.g. wavelets or the FFT applied to nonlinear cases. Although we do not know how well the KdV describes the propagation of the measured time series around the measurement site, the KdV does not have to be valid for most of the conclusions of this article, which investigates the results of a signal transform to rogue waves. We do not want to claim that the soliton components in the nonlinear spectrum are physical. We tried to point this out e.g. in the abstract ("Under the hypothesis that the KdV describes the evolution of the sea state around the measurement site well, these results suggest that solitons ...") and the conclusion ("Each measured rogue wave could be associated with at least one soliton in the NLFT spectrum."), but see that this should be pointed out more prominently. We will clarify this in the abstract, the introduction and the conclusion. Our study does not intend to explain the mechanism of rogue wave generation in shallow water. The method should rather be interpreted as a spectral analysis method. We would like to gain insight into the spectral characteristics based on KdV-NLFT at the available measurement site. These spectral characteristics and their differences in samples with and without rogue waves are described in this paper. We would like to point out that in our work, the vKdV-NLFT is applied to a large number of real-world time series for the first time. It is also the first time that certain characteristics of nonlinear spectra could be linked to rogue waves. We thus present a first assessment of the NLFT applied to real measurement data from shallow depths. This is only a first step and future research is needed.
- The reviewer states that “KdV solitons [may usually] be recognized [in time series] by eye” (nhess-2022-28-RC2-supplement, p. 1). We would like to object to this statement and refer to Zabusky and Kruskal (1965), who described the evolution of a sinusoidal-shaped surface elevation, in which solitons eventually form from the background, while not being immediately visible. The observation is reinforced by Brühl & Oumeraci (2016) for the evolution of a long-period cosine wave in very shallow water and in Brühl et al. (2022) for an initially trapezoidal-shaped bore. Here, the solitons that are found by KdV-NLFT, are not immediately visible in the time series, but the surface elevation eventually decomposes into a train of solitons in the far field. While the time series changes with time, the nonlinear spectrum remains invariant. This shows that time series exist, in which KdV solitons are not visible by eye, but may be identified by KdV-NLFT. Another reason for the “invisibility” of the solitons in the time series is that the water surface in the North Sea is not calm before and after the recording of the time series. This means that

all existing solitons will continuously interact with the surrounding waves, which makes them difficult to identify by visual inspection. Figure 1 in Osborne et al. (1991) e.g. demonstrates that solitons do not have to be clearly visible in a real-world measurement.

- The reviewer points out that “the estimated soliton amplitudes are not very large, the solitons do not dominate” (nhess-2022-28-RC2-supplement, p. 1). We have reconstructed a soliton train underlying a time series, by nonlinear superposition of solitons, using the algorithm from Prins & Wahls (2021) (Figure 1 in this document). The time series corresponds to the example in Figure 5 of the preprint. (Note that inverting large soliton spectra is numerically very difficult (Prins & Wahls, 2021). We therefore had to use a shortened time series for the figure.) Figure 1 in this document supports findings by Osborne et al. (1991) that show that the solitons are much lower in amplitude than the maximum waves in the time series. Therefore, we agree with the reviewer, and we have stated so in the conclusion, that in our rogue-wave samples solitons alone cannot be responsible for the formation of the measured rogue waves. The continuous spectrum of the vKdV-NLFT, which actually contains most of the energy in our time series, must account for the remaining parts of the exceptional heights. However, the soliton contribution is not negligible and has the potential to turn a non-rogue wave into a rogue wave. This is concluded from differences in the discrete spectra of samples with and samples without rogue waves.

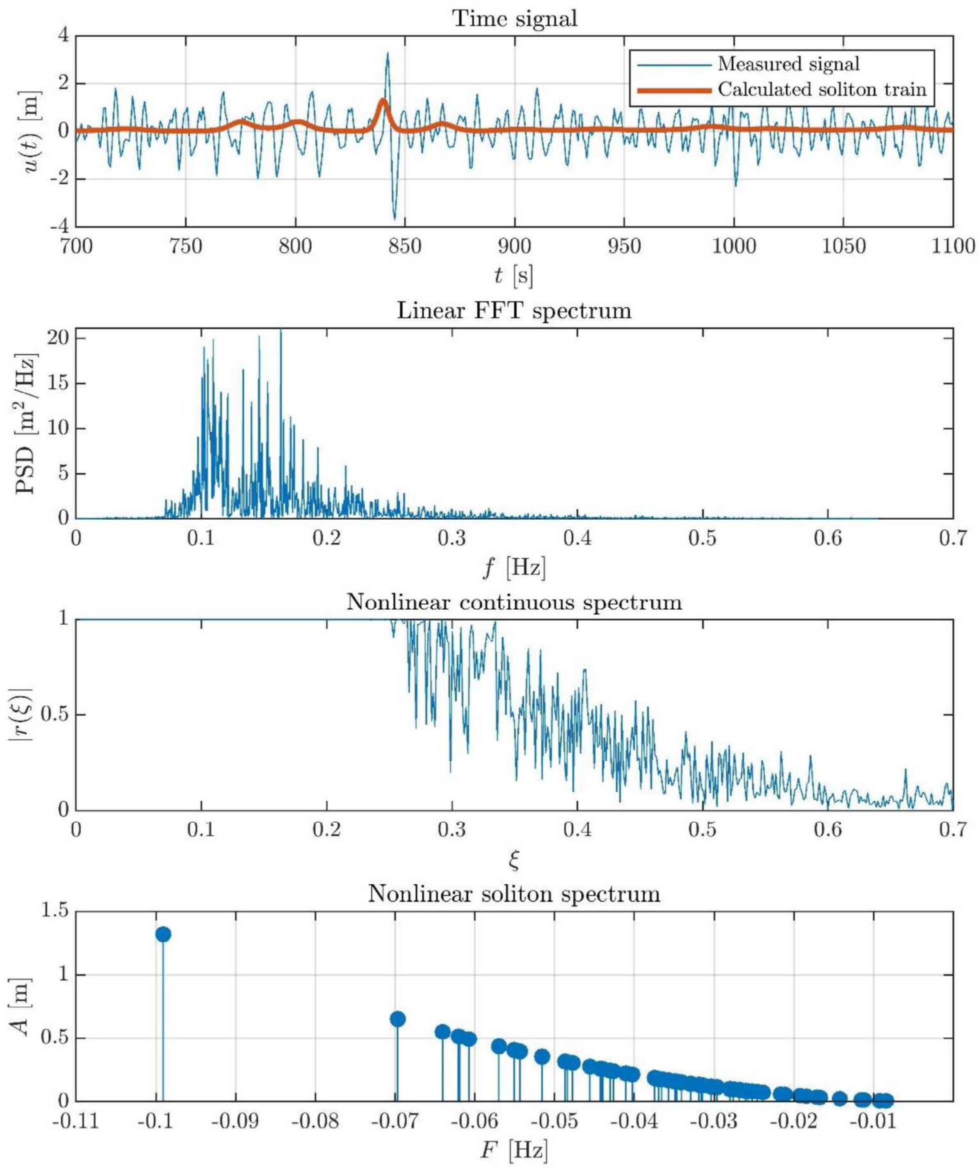

Figure 1: a) time series and soliton train, as calculated from the inverted soliton spectrum. b) linear Fourier spectrum. c) continuous spectrum. d) discrete soliton spectrum.

- We agree with the reviewer to add in the discussion and conclusion that the assumption that waves around our measurement station may be approximated by the KdV equation, cannot be proven based on the available data.
- The reviewer suggests the presence of envelope solitons at varying depths and draws our attention to a study in which "envelope solitons are shown responsible for the wave amplification when the depth increases" (nhess-2022-28-RC2-supplement, p. 1). We agree that varying bathymetry might be an explanation for the enhanced rogue wave occurrence, and will add the corresponding references.
- The reviewer suggests "the oblique interaction of KdV solitons" as an "essentially directional effect" (nhess-2022-28-RC2-supplement, p. 1) as a reason for the enhanced rogue wave occurrence at Norderney. As in the previous comment, we agree with the reviewer that this reason is conceivable, and also, that we cannot assess it in the frames of the KdV equation. All we offer in this article is a signal transformation by NLFT, which

suggests an influence of the presence of shallow-water solitons on rogue wave generation.

- We thank the reviewer for the references on NLS solitons and rogue waves measured in deep water, the accurracy of the NLS equation for strongly nonlinear, the observation of long-lived NLS solitons in the field of strongly nonlinear waves, and KdV soliton content in sinusoids. We will include these in our introduction of the revised article.

(changes in lines 6-12, 21/22, 93-103, 139, 148-151, 152-154, 186/187, 189-192, 284-290, 475-485, 544-552, 593, Figure 5)

2. Comments regarding the text

- l. 44: The reviewer points out that the NLS equation is not limited by *kh* = 1.36. We agree that the formulation in the paper might be misleading and we will clarify this in the revised paper. We will reformulate line 44 in the text to show that the condition is not oriented towards the applicability of the NLS equation, but towards the applicability of the KdV equation. (lines 64/65, 129/130)
- l. 83: The reviewer points out that the KdV equation is not the equivalent to the NLS equation, since the former is a wave displacement equation, while the latter is a wave modulation equation. We agree with the reviewer that this formulation is misleading and we will change it accordingly. (lines 126/127)
- l. 48: We thank the reviewer for the remark that the most unstable perturbations are longitudinal, not oblique. We will replace the word “oblique” by “side-band”, as to omit the direction of the perturbation and imply that disturbances arise in the form of side-band modes $\omega(1 \pm \delta)$ to the frequency ω of the initial wave train. (line 69)
- l. 53: The reviewer points out that “the BFI parameter was introduced for characterization of irregular wave statistics for the first time in [Onorato et al, (2001)].” (nhess-2022-28-RC2-supplement, p. 2) We agree and we will include this information in the text. (line 73)
- l. 55: The reviewer points out that breather solutions do not explain the physics of the modulational instability, but describe its dynamics. We agree and we will change the text accordingly. (lines 76/77)
- l. 61: Thank you, we will insert the correct date of the reference. (line 87)
- l. 67-73: We thank the reviewer for raising the issue of the shallow water definition. In the context of this study, we have termed the range of applicability of the KdV equation as given by Osborne and Petti (1994), p. 1731, and Osborne (1995), p. 2629, shallow water. We acknowledge that, therefore, our definitions of shallow water (*kh* < 1.36 or *h/L* < 0.22) and deep water (*kh* > 1.36 or *h/L* > 0.22) are different from the definitions of shallow (*h/L* < 0.05), intermediate (0.05 < *h/L* < 0.5) and deep water (*h/L* > 0.5) that are used in the engineering context, and that this may lead to confusion. We will clearly define the terminology for ‘shallow’ that is used in the paper and repeat the *kh* value in line 67. Furthermore, we will change the title to include “shallow depths” rather than “shallow water”, as suggested by Referee #1, as to avoid confusion of different shallow water terms. (lines 106-108, 129/130, 215-218, 227, 230-233, 472-474, 521/522527, Table 3, line 588)
- l. 78: The reviewer points out that rogue waves in variable depths have been discussed in several studies. We will adjust this text accordingly. We thank the reviewer for the additional reference, which we will include. (lines 111-113, 122/123)
- l. 97-99: We will refer to the book of Ablowitz and Segur (1981), where in Chapter 1.7c the asymptotic behaviour is discussed for both the continuous and the discrete spectrum. Since

the content in this book is somewhat scattered, we also plan to refer to the paper of Ablowitz and Kodama (1982), who correctly analysed the asymptotic behaviour for the first time. (lines 142/143, 157)

- l. 102: To solve this issue and avoid misunderstandings, we will include the original citation of Pelinovsky et al. (2000): "the "nonlinear" train should include a soliton". (lines 159/160)
- l. 103, 257: We agree with the reviewer that the interaction of KdV solitons alone does not lead to the formation of the observed rogue waves. We have stated so in lines 104-105. To make this statement more clear, we will reformulate the words "the interaction between one or multiple solitons with oscillatory waves" as "between one or in principle, multiple solitons, with dispersive waves". In line 257, we presume that the nonlinear interaction of solitons with dispersive waves is the probable cause. For clarification, Figure 1 in this document will be added to the revised version. The figure shows an exemplary time series (blue line in the first plot), its linear FFT spectrum (second plot), the nonlinear continuous spectrum (third plot, which will not be analysed further in this article), and the soliton spectrum (last plot). For visualisation of the role of the solitons in the time series, the first plot shows the soliton train (red line) that is obtained by nonlinear superposition of the solitons (considering their interactions) using the algorithm from Prins and Wahls (2021). This example clearly shows that solitons and their interactions are not solely responsible for the generation of the observed rogue wave. (lines 162, 482-485)
- l. 117: We thank the reviewer for the additional references, which we will include. (lines 81-84, 146-148)
- Figure 2: We agree with the reviewer to add a scale for the distance to the map. (Figure 2)
- Equation 1: According to the reviewer, the applied shallow-water threshold is related to the BFI and not to the applicability of KdV. We refer to Osborne and Petti (1994), p. 1731 and Osborne (1995), p. 2629 for the shallow-water threshold in KdV and will add the references to the text. (line 218)
- l. 155: "1" will be removed from the equation. (line 223)
- Table 1: The caption will be corrected. We agree with the reviewer that due to the recording of waves at discrete sampling points, there is a possibility that the exact crest of a wave is missed. Therefore, it is conceivable that some of the rogue waves termed "height rogue" are actually underestimated and should belong to the category of "extreme rogues". On the other hand, a possible misinterpretation is conservative, thus, all spectral characteristics attributed to extreme rogue samples are assigned correctly. Furthermore, we think that the impact of such effects is small because the sampling frequency is sufficiently large. For the sufficiency of the sampling frequency of our wave buoy, we would like to cite the sampling theorem (Shannon, 1949): if a continuous time signal contains no frequency components higher than W Hz, it may be completely determined by uniform samples taken at the Nyquist rate of f_s = 2W samples per second. A typical (raw) FFT spectrum of a rogue wave time series from our data is shown in the second plot of Figure 1 in this document. It is seen that its components approach zero at approximately 0.5 Hz and have fully decayed at W = 0.64 Hz. The Nyquist rate f_s = 2W = 1.28 Hz is the measurement frequency of the wave buoy. The signal is therefore oversampled by a factor of more than two. We will add a short discussion of this point in the revised paper. (Table 1, lines 564-570)
- l. 211: The reviewer points out that the frequency axis of the discrete soliton spectrum indeed has a physical meaning, which is the inverse duration of a soliton. We agree that the formulation may be misleading: although, theoretically, a soliton has an infinitely long duration, since it does not cross the surface, a mathematical definition of the angular frequency can be established (Equation 10). We refer to the soliton solution of the tKdV (see

e.g. Equation (12) in Bruehl et al. (2022)), from which the angular frequency may be obtained. We will therefore reformulate the misleading line 211. The reviewer correctly points out that the amplitude and the frequency of a soliton are related. Hence, the frequency gap in the soliton spectrum may be described either in terms of an amplitude ratio or in terms of a frequency ratio. As stated by the reviewer, the relation between the amplitude and the frequency is quadratic. Consequently, a soliton possessing 80% of the amplitude of the maximum soliton, has a frequency of $\sqrt{0.8} = 0.89 = 89\%$ of the frequency of the maximum soliton. In other words, a reduction in soliton amplitude by 20% corresponds to a reduction in frequency by only 11%. Since the relative differences in the amplitudes are easier to observe in the amplitude-frequency plots, and the soliton amplitude is a more descriptive parameter than its frequency, furthermore, rogue waves are defined in terms of amplitudes, we have selected to use the amplitude ratio instead of the frequency ratio for the definition of outstanding solitons. (lines 290-294)

- l. 267, 269: The word “right” corresponds to the statistical term used for a distribution drawn over an x-axis. The reviewer is right that this may be misleading, as the distribution is presented as a box plot, with the vertical axis as a reference. We will therefore rephrase the term “to the right” as “towards higher normalised soliton amplitudes”. (lines 352, 354)

- l. 276/277 and l. 428/429: As already discussed in the beginning, the term "solitons" was mostly used in the manuscript to refer to components in the discrete spectrum of a time series, which might have been confusing. We will explicitly discuss this issue in the abstract, the introduction, and the conclusion.

We would also like to point out that just like the usual Fourier transform is applied also to signals under nonlinear propagation, we apply the vKdV-NFT to signals that may not propagate according to the KdV. Our results show that the vKdV-NFT, when considered purely as a signal processing tool, leads to interesting new characterizations of certain rogue waves, but also demonstrates limits of this approach (at least in the form used here). How far the soliton components in the nonlinear spectrum are physical is an important question, but that question goes beyond what we can answer with our current data. We nevertheless believe that our work is an important step towards bringing the NFT to real-world data, and hope that it motivates future research in that direction.

(lines 6-12, 186/187, 286-288, 544-552, 593/594)

- Figure 10: We thank the reviewer for this hint and we will add the dimension to the axis. (Figure 10)

- Table 3: The relation mentioned by the reviewer is correct for a single soliton, but the soliton spectrum of more complicated time series changes in more complicated ways with the water depth (Figure 2). We prefer not to put this comment to avoid confusion.

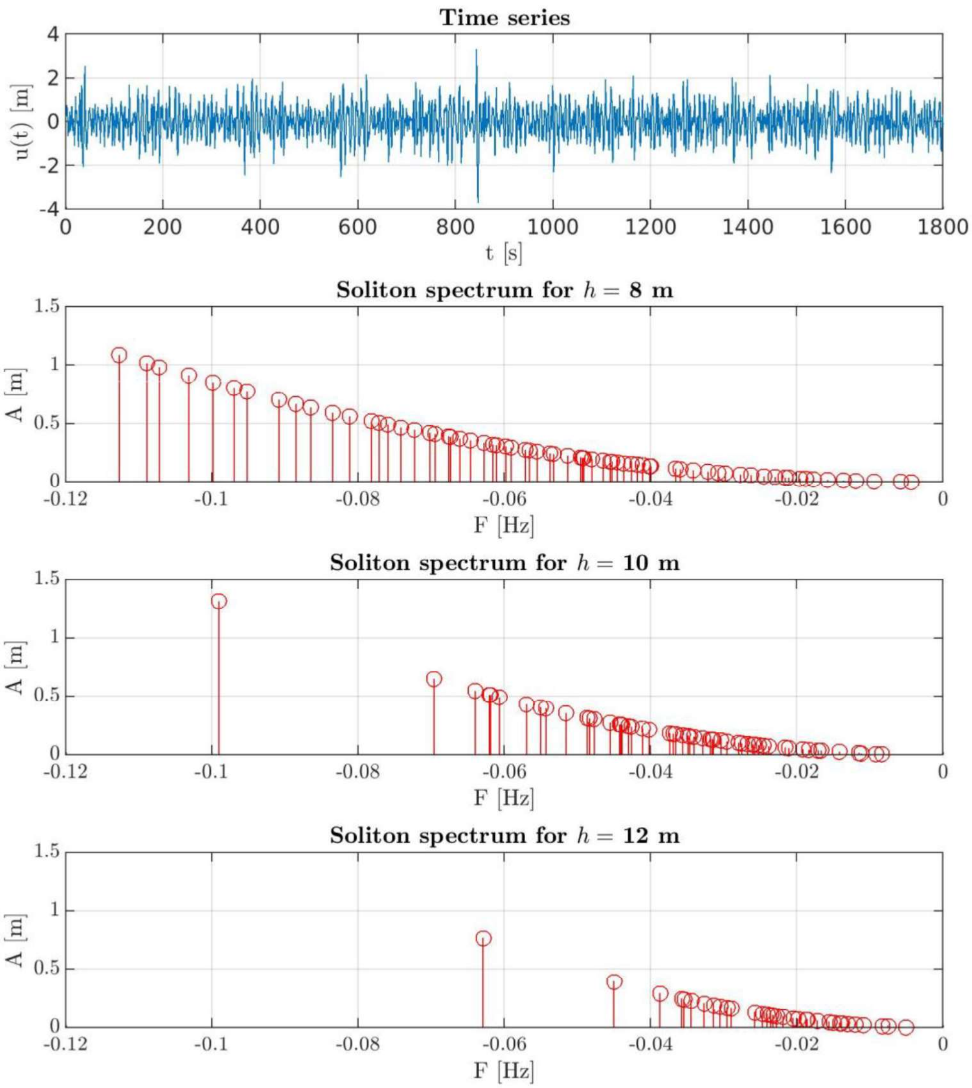

Figure 2: Soliton spectrum of a rogue wave time series for different water depths assumed in the calculation.

References

Ablowitz, M. J. & Kodama, Y., 1982. Note on Asymptotic Solutions of the Korteweg-de Vries Equation with Solitons. *Studies in Applied Mathematics,* 66(2), pp. 159-170.

Ablowitz, M. J. & Segur, H., 1981. *Solitons and the Inverse Scattering Transform.* Philadelphia: SIAM.

Brühl, M. & Oumeraci, H., 2016. Analysis of long-period cosine-wave dispersion in very shallow water using nonlinear Fourier transform based on KdV equation. *Applied Ocean Research,* Volume 61, pp. 81-91.

Brühl, M. et al., 2022. Comparative analysis of bore propagation over long distances using conventional linear and KdV-based nonlinear Fourier transform. *Wave Motion,* Volume 111, 102905.

Onorato, M., Osborne, A.R., Serio, M., Bertone, S., 2001. Freak Waves in Random Oceanic Sea States. *Physical Review Letters,* Volume 86, pp. 5831-5834.

Osborne, A. R., 1995. The inverse scattering transform: Tools for the nonlinear Fourier analysis and filtering of ocean surface waves. *Chaos Solitons Fractals,* 5(12), pp. 2623-2637.

Osborne, A. R. & Petti, M., 1994. Laboratory-generated, shallow-water surface waves: Analysis using the periodic, inverse scattering transform. *Phys. Fluids,* 6(5), pp. 1727-1744.

Osborne, A. R., Segre, E., Boffetta, G. & Cavaleri, L., 1991. Soliton basis states in shallow-water ocean surface waves. *Physical Review Letters,* Volume 67, pp. 592-595.

Pelinovsky, E., Talipova, T. & Kharif, C., 2000. Nonlinear-dispersive mechanism of the freak wave formation in shallow water. *Physica D: Nonlinear Phenomena,* Volume 147, pp. 83-94.

Prins, P. J. & Wahls, S., 2021. An accurate O(N2) floating point algorithm for the Crum transform of the KdV equation. *Communications in Nonlinear Science and Numerical Simulation,* Volume 102, 105782.

Shannon, C. E., 1949. Communication in the Presence of Noise. *Proceedings of the IRE,* 37(1), pp. 10-21.

Zabusky, N. J. & Kruskal, M. D., 1965. Interaction of "Solitons" in a Collisionless Plasma and the Recurrence of Initial States. *Physical Review Letters,* Volume 15, pp. 240-243.