Dear Reviewer

Thank you again for your very constructive comments. Your comments on revision#1 of the paper are addressed below and the line numbers addressing the modifications in the manuscript are presented. In the manuscript the revisions are highlighted with light blue:

Comments and Responses

#26 It's rather the shortcomings of various source terms, not only the balance

Line #26 was modified to reflect the above fact.

#57 Correct reference to Van Vledder et al., 2016.

The reference was corrected (Ln#56)

#65 As discussed in W007 his new source term was primarily developed for coastal and inland waters with relatively short fetches.

"Coastal and inland waters" was mentioned in Ln#65

#78-#81 This is incorrect. The OMP version of SWAN4120 contains ST6. Please drop this fake argument. Implementation issues can never be a proper argument for not doing something. Please concentrate on scientific/physical arguments. A stronger argument is twofold. There is limited experience with ST6 in coastal waters, and the range of applicability of W007 for larger fetches with inhomogeneous wind fields is poorly known. As said before, the crucial difference between Komen type and W007 whitecapping source term is the use (or not) of a mean wave steepness, which gives modelling errors in multi-peaked seas. Note that W007 did not really test his source term for multi-peaked seas.

The sentences were modified based on the above comment (Ln #76-80)

#82 please state which complications may arise. This is now too vaguely formulated.

The sentence was rephrased to avoid confusion Ln#81)

#87 The apparent contradiction is just another argument that no generally applicable source term setting exists requiring tuning. In the North Sea also shallow water effects like bottom friction are a significant component of the total source term balance.

The above statement was added to make it more clear(Ln#87-90)

#104 If you know about the gustiness effect, why did you not use it to 'correct' the CFSR winds?

This clarification was added to the manuscript: (Ln#105-107)

"To be consistent with the practical modeling efforts and real applications, we did not correct the wind field for gustiness. Therefore, in our simulation we used CFSR wind field with 1 hour temporal resolution that cannot include such short term variations."

#108 Now it becomes clear why you used W007. When you stick to the available SWAN-ADCIRC executable you run into unnecessary problems.

We had to use SWAN-ADCIRC because we needed the ADCIRC ability for domain decomposition of the unstructured grid which is necessary for implementing high performance computing over a very large numbers of computational grids (more than 4,300,000). It is mentioned in page 4 Ln#112 of the revised manuscript.

#112 what are these?: … each Komen method.. This is not yet introduced. I can imagine choosing delta=0 or delta=1 may be choices.

In the revised manuscript we referred to section 2 for more details on whitecapping formulations : Ln#116-117: "(see section 2 for details of formulations)"
Details on whitecapping parameters are mentioned in section 5, Ln#253-254 of the revised manuscript.

#116 You should add also the relatively large fetches and instationary large wind fields as an argument. No reason to state that no new form is suggested. You may put that in the discussion.

Was modified according to the comment (Ln#119-121)

#124 Only using January, although representative for the winter months, is still a bit meager. You can also reverse the argument, because you want to study some features of the wind fields, you choose this month for studying. Moreover, this also fits better in view of the limitations of a fetch-limited approach. This is a stronger argument and better fits in the purpose of this manuscript.

The following statement was added in Ln#129-132 to show what specific features we need to study by choosing January:

“The persistent offshore-ward wind field during this month (Allahdadi et al., 2019) along with large fetch lengths over the modeling region provide an appropriate condition to study different features of wind field and waves including fetch-limited and fully developed sea states.”

#134 It is better to discuss the physical aspects, viz. the impact on lower or higher frequencies, see Rogers et al. (2003) for a discussion

The conclusion from Rogers et al(2002) was mentioned in Ln #141-142

#186 typo: July 2009
The Typo was fixed (Ln#193)

#195 typo: wind roses (2 words)

The typo was fixed (Ln#203)

#208 …was developed… suggest that a new model setup was created. This seems to contradict the statement that you were forced to use an existing SWAN-ADCIRC approach.

We actually used unstructured swan that is coupled to ADCIRC for domain decomposition used for high performance computing. We added the clarification to section 4, Ln#215-216

#216 unit of bias is wrong, should be m/s. Lower panels in Figure 4 shows systematic trend of under-prediction with higher wind speeds. This should be noted, as well as the question whether you corrected for this or not.

Unit of bias was fixed (Ln#224)
Underestimation trend of the CFSR for higher win speeds and the statements that “no correction” was applied to the wind field are mentioned Ln#225-228

#218 before -> around
Fixed(Ln#228)

#218 Make the buoy position markers and text larger for visibility

Fixed. See the modified Figure 5

#222 timestep is the wrong word here. A time step is an interval, whereas here you mean a moment in time. Please rephrase.

All “timesteps” words related to this matter were changed to “time”

#236 which two approaches? From earlier remarks I count 3. One W007 and two Komen-type formulations?. If only one Komen is used, what delta is used?

Only two approaches were used as mentioned in section 2: Komen (the default SWAN approach), and Westhuysen as a Saturated-based method (as mentioned in section 2). The default values for both method is used. So, the delta regarding the Komen method is 1. The default parameters were added to section 5, Ln #253-254

#238 as defaults may change over time (as they did), you can better explicitly state the settings used.

The default parameters were explicitly mentioned (section 5, Ln #253-254)

#242 It is still unclear which delta is used in the Komen formulation. Please note that the choice of delta may have significant effects.

Delta value was added to section 5, Ln #253-254)

#252 Crucial pieces of information are missing here. Which type of buoys are used? Over which frequency range have the buoy spectra been integrated? Now it seems that different frequency intervals have been used. In case SWAN table output for Tm02 has been used, the integration is up to 10 Hz (using the prognostic and parametric spectrum). Buoys usually deliver reliable spectra up to 0.5 Hz. This mismatch may cause significant differences for especially Tm02, see section 4.3 Akpinar et al. (2012) for a discussion on this topic. In such a comparison, one should always use the same frequency interval to derive parameters. For Hm0 the effect is often insignificant, but the higher the frequency moment, the larger the effect. This effect may partly explain the under-prediction. Please check carefully, the consequences.

We have already mentioned the frequency ranges of the buoy data (0.02-0.485) in page 9(Ln#304-305) of the manuscript (highlighted). We agree that the difference between the integration range of the buoy data in the frequency range with the Prognostic range of SWAN may cause some differences for Tm02. We will mention that in the result part of page 8. However, as Akpinar et al (2012) showed, the differences decrease for higher values of Tm02. In fact for measured wave periods larger than three seconds, the differences are generally negligible. Since the measured wave periods during our simulations at all four buoys are larger than three seconds for most cases (Figure 6e-h), we can safely neglect this discrepancy for the wave period comparison. Furthermore, we picked two events with large wave height and period for discussion (times t1 and t2, see table 2 for values). .We have added some short discussion about this effect and the probable effect on our results in page 8, Ln#267-274.

#271 Note that this under-prediction may also be due to the fact that no calibration was performed.

That is absolutely right. However, in this paper we only examined the default parameter. In the discussion part (section 6.2) we mentioned about the necessity for recalibration of models for different sea state conditions.

#300 Can you find a reason why Komen spectra are larger than those of W007. Is this due to the use of a mean wave steepness which in the presence of low-frequency wave components results in an over-prediction of the higher frequencies?

In the manuscript this behavior was attributed to the larger algebraic sum of the source terms from Komen compared to the Westhuysen Whitecapping:

“Komen simulated a larger sum of source terms at the peak frequency and all frequencies below that (Figure 12d). This result is consistent with Figures 10a that shows higher spectral energies at this time by Komen compared to Westhuysen” (page 12, Ln#386-388)

#309 Author Clyson does not exist. Check reference

The correct reference is *Kahma and Calkoen, 1992* that was corrected in the manuscript (see Ln#326).

#326 Can you estimate how much wind strength errors contribute to the total prediction error. As noted, there is an inherent prediction error in each source term, but here also the numerical determination of Tm02 may contribute to the total error. Without such a quantitative consideration, the discussion is a bit pointless.

In this paper the main purpose is comparing the performance of two models and present different reasons for these comparative differences. Since we used the same wind field for simulation using both whitecapping approach this comparison is meaningful. However, determining and quantifying the effect of wind field discrepancies can be a great idea, but needs thorough simulations. We will suggest it as a topic for further researches in the conclusion part. However, based on a calibrated model for the same area(Allahdadi et al., 2019) some preliminary quantifications on contribution of the wind on model discrepancies were presented(please see section 6.1, Ln#343-348)

Figure 9 Add proper name and symbol of wave parameters in legend ‘significant’ wave height Hm0 , .. mean wave period Tm02

The proper names and symbol were added to Figure caption.

#332 Specify the kind of source terms shown in Figure 11. I guess they are integrated source term magnitudes. This should be explained.

Yes. They are integrated. Clarification was added in Ln# 354)

Figure 11 This figure seems inconsistent with what I expect in relation to Figure 9.
1) The shape of the spatial distributions seems inconsistent with those in Figure 9.
2) The upper panels have a color band in yellow, whereas this is not present in the lower panels.
3) Using the unit w/m2 is wrong, as these are integrated source terms of the rate of change of wave variance. No multiplication with rho g has to be done.
4)

1. Figure 9 shows the simulated maps of wave height and period for times t1 and t2. Figure 11 shows the source term variation over the modeling area only for time t1. If you compare panels in Figure 11 with figures 9(a,b) and 9(e,f) that are onlr related to time t1 they would be consistent.
2. Lack of yellow color in the lower panel(Westhuysen source terms) is due to the fact that the whitecapping and wind input and consequently the quadruplet terms in Westhuysen approach are scaled based on larger factors that results in larger source terms than Komen(have been mentioned in section 6.1).

3. The unit for the source terms was changed to $\boldsymbol{m^2/s}$ (Ln#360)

#336 As mentioned earlier, the source terms are different because the underlying variance density spectra are also different. Still, I agree that for the same spectrum, the trend is similar. This was easily checked for a JONSWAP spectrum.

Thank you for the confirmation.

#344 This choice is related to keeping the total balance appropriate.

The statement was added to the text in page 11, Ln#367.

#357 The chosen spectrum and related wcap source terms do not really show some effect for small frequencies. I did a test for a JONSWAP spectrum and then subtle for significant differences popup.

This is consistent with W007 as the observed the same differences for energy at low frequencies.

#360 A comment on the strange shape of Snl4 is missing

The comment was added to the bottom of page 11, Ln#378-380.

#368 How can a model be more than reality? Please rephrase

The sentence was modified (Ln#393-394)

#393 are resulted -> result

We suspect that “resulted” maybe the right word.

#430 result -> results
Fixed, Ln#455

#442 Such a distinction will be difficult in practice. This is another indication that these formulations are not generally applicable. This is an interesting conclusion. As noted in Bingolbali, tuning for a large basin may give different results when tuning is done for a local area.

The phrase “ if possible” was added before this statement to show the difficulty of implementing this suggestion (Ln#466-467)

#447 rephrase … too much…, -> occur often?

Correction was applied (Ln#472)

#475 directional spectra cannot be observed, only reconstructed from the Fourier coefficients that can be distilled from the buoys time series. So, state how the buoy spectra were reconstructed

It was mentioned in the manuscript that the directional specra is “reconstructed from Fourier coefficients” (Ln#504-505)

#489 How do you quantify directional spreading?
We estimated the directional spreading as “the total angle for which wave energy exists within the scale to 360 degrees”. This clarification was added to page 15, Ln#519-520

#468 Note that the quadruplet term here is the crude DIA. See Ardhuin et al. (2007) and Bottema and Van Vledder (2008) for possible effects on slanting fetch situations when using the accurate Snl4.

The DIA method used in this research and differece with results from the exact method (Xnl) based on Bottena and Vledder(2008)was mentioned in pages 15, Ln#499-503.

#515 This is a welcome addition to the manuscript, leading to a nice conclusion about the inclusion of stability effects in wave modelling.

Thanks for approving.

#557 calibrated by whom and in what way? This is an important notion
.
We meant “calibrated models based on growth curve of Kahma and Calkoen(19912)”. We added this to page 17 of the manuscript Ln#588-589

#571 The logic here is a bit flawed. Variabilities in the wind field itself are no cause for slanting fetch effects. Please rephrase.

This phrase was modified and the wind field effect was removed (Ln# 601-602)

#585 What are ‘real values’?

The word “Unrealistic” was used at the beginning of the sentence and “than real value” was removed to avoid complications (Ln# 615)

#588 technical -> physical ?

We think “scientific” maybe a better word, so used that instead of physical in the page 18 (Ln# 619)

#590 omit the distinction between serial and parallel, that is not relevant

“serial and parallel” modes were removed (Ln# 620-621)