COMMENTS FROM REVIEWER #1:

Review of "Accuracy of numerical wave model results: Application to the Atlantic coasts of Europe"a paper by Alday and 3 co-authors.

Summary:

The authors design an unstructured grid implementation of WW3 for the European coastline. They evaluate the model against buoys and altimetry. They perform sensitivity analyses to determine impact of various things. Topics studied include: simple wind adjustments, swell dissipation settings, tidal currents effects, directional resolution, and bottom friction sensitivity.

Recommendation:

I'm not very familiar with the expectations of the journal, especially re: what types of papers are OK. I have a generally positive impression of the paper. So, I'd like to give the authors the benefit of the doubt, and recommend "accept with minor revisions"

General comments:

Some of the comparisons seem rather workaday, like the sort of thing that would go into a tech report or dissertation. But it's OK, I think. And they are not just dry comparisons: the authors put substantial thought into them, so nothing comes across as especially unnecessary.

The literature review is sufficient.

The comparisons are careful, the discussion is straightforward/honest, and the authors don't over-sell the outcomes.

The findings are potentially useful for people doing similar modeling.

The use of English is very good. There are a few awkward or broken sentences (e.g. lines 310, 393-394) and minor style errors (e.g. lines 21, 27, 409, 412), but not more than ~ 1 problem per page.

The purpose of the study meanders.

There are a lot of evaluations, but they are a grab-bag that are only similar insofar as they are for the same system. I put them in four broad categories:

1) differences vs. global model (geographic resolution)

2) impact of design choices, where choice is not clear a priori (swell dissipation setting, directional resolution, BC formatting, wind adjustments)

3) impact of including/excluding effect, where it is more or less accepted a priori that the effect should be included (bottom friction, tidal current)

4) general evaluation of the model in its final form (altimetry comparison).

I don't have a good suggestion for how to address this. I suppose the authors do not want to stick to a consistent type of content, since they would lose half the study.

I really like the evaluation of the model against altimetry, sorted by distance from the coast. Excellent.

OBS.1. There are a number of unexplained acronyms/initialisms. Perhaps a glossary would work better than rigorous in-line explanation. I leave this to the editor to decide.

R1: We have reduced significantly the use of acronyms in order to improve readability, for example

- **the number of occurrences of "NMD" (normalized mean difference) has been reduced from 16 to 5**
- **the 6 occurrences of WAE (Wave Action Equation) have been removed,**
- **The 3 occurrences of MSL have been removed**
- **10 occurrences of NMB have been removed**
- **10 occurrences of SI have been removed**
- **20 occurrences of SWH or H_s have been removed**

OBS.2. When printed out, some of the figures are hard to see, like Fig 7 a-d, where lines are hard to distinguish.

R2: Some figures were indeed a bit busy. We have chosen to simplify the figures to make it possible to magnify them. In the case of Figure 7 the normalized bias plots were removed.

Specific comments.

OBS.3. 52: "CFL...minimum time step". Wouldn't CFL dictate the maximum time step?

R3: This sentence was rewritten for clarity,

This coarsening allowed a lower Courant-Friedrich-Lewy (CFL) number, which makes it possible to use larger time step for wave propagation, 13~s in this case for our lowest frequencies, but it also implies that details of the Norwegian coastline are not as well resolved.

Indeed the actual time step used at run time is automatically adjusted between a maximum value and whatever it takes to keep the CFL number under 1, and this adjustment is done separately for each spectral component (higher frequencies that propagate slower are typically integrated with higher time steps).

OBS.4. Same paragraph. Why is the implicit scheme not used? It should be mentioned, at least.

R4: We have added the following paragraph after Fig.1:

An alternative to this careful editing of the mesh is the use of implicit schemes. However, using implicit schemes with CFL values much larger than 1 opens the door to both larger advection errors (stability does not imply accuracy) and larger splitting errors as the time steps for advection can be much smaller than the refraction and source term time step (Roland and Ardhuin 2014). We have preferred to stick to the explicit N-scheme because numerical efficiency is not central in the study, and it simplifies comparisons with global model results that also use explicit schemes. Implicit schemes are probably necessary when resolving regional scales and surf zones in the same mesh when CFL constraints require prohibitively small time steps in explicit schemes.

OBS5. 63: Boulders of D50=15 cm. This is too small to be boulders, which are at least 25 cm, according to wikipedia. Cobbles, maybe?

R5: The reviewer is correct. Based on the Wentworth grain size scale, sediments of D50 ~15cm are termed pebbles, thus bottom sediment size specifications in line 64 will be corrected as follows:

"...,while zones characterized as pebbles or larger elements (boulders) were represented with a D50=150 mm."

OBS.6. 95: only tidal currents are used. Maybe it is OK to neglect general ocean currents, but it should be justified/explained here.

R6: This is particularly an interesting point. Non-tidal flows are generally dominated by geostrophic currents. The impact of these currents are analyzed in (among others) Marechal and Ardhuin (2020), Alday et al. (2021) and Echevarria et al. (2021). Resolving the small scales of these currents at least 30 km is not feasible in a deterministic sense because the quasi-geostrophic dynamics are not yet resolved in observations (Ballarotta et al. 2019). Introducing modelled quasi-geostrophic currents at scales shorter than 100 km can often increase model errors (even though the result can be statistically better) because structures are not constrained by assimilation to be in the right place.

OBS.7. 113: 30 sec time step is mentioned here, but we saw 13 sec limit earlier, so I don't understand the discrepancy. Please explain.

R7: As explained in R3, the 13 s is obtained from the CFL condition for the lowest frequencies. The actual advection time step varies with spectral components and is 30 s at most.

OBS.8. 125: maximum frequency of buoy should be mentioned. And later, the frequency range used for calculation of Tm02, etc.

R8: The following specification will be added in the caption of Table 1:

"...Deploy depth obtained from model bathymetry interpolated into the buoys' position. All buoys with spectral data present a frequency range from 0.025 to 0.58 Hz, with a frequency interval of 0.005 Hz."

In relation to the wave parameters computed from the frequency spectrum, the following will be added to the caption of Fig.6:

"...Hs bin size is 0.25 m, periods' bin size is 0.2 s. Hs, Tm01 and Tm02 are computed integrating the spectra in the frequency range 0.0339-0.537 Hz."

OBS.9. Section 4.1. It should be made more explicit up front here that the impact of resolution is studied by comparison against a global model, not by applying two regional models of different resolution. (Initially, I guessed the paper was doing the latter.) Same problem occurs on line 383.

R9: In the case of section 4.1, we consider that the comparison method is properly explained from line 145 to line 148:

"A comparison between February 2011 mean significant wave heights (Hs) fields from the global model described in section 2.4 and our implemented regional model is presented in Fig. 5. To evaluate the differences between models, the output from the 0.5◦ global grid was linearly interpolated into the regional mesh nodes before computing the mean Hs and the NMD for the selected time window."

The text in lines 382 to 383 will be changed to:

"We found that higher spatial resolution adequate to solve bathymetry features and explicitly solving coastlines can introduce changes in Hs estimations of about 20% when compared to lower resolution global models"

OBS.10. Figure 7. Is mean E(f) from an average over the 14 days? This should be specified.

R10: All stats and mean E(f) in Fig. 7a-d are computed comparing 1 month simulations with buoy data. Time series presented in Fig. 7e-f are there to show the characteristics of H20 from simulations.

The following will be added in the caption of Fig.7:

"...In panels (a) to (d) 1 month modelled results compared with buoy data…"

OBS.11. Section 4.4. I don't understand the difference between 48D24BC and 48D48BC. Both are interpolating from the same global model to the same regional model, right? Is it a question of doing the spectral interpolation in WW3 vs. outside WW3? This should be clearer.

R11: To clarify the approach of the tests in Section 4.4 (specified in table 3), the following explanation will be added after the paragraph between lines 224 to 229:

"…For example, the difference between 48D24BC and 48D48BC is that the boundary conditions (BC) were created in a global model with different spectral resolution. Test 48D24BC employs boundary conditions from a global model with 24 discrete spectral directions equivalent to 15º which are then interpolated into 48 directions to match the mesh resolution (48D). On the other hand, in test 48D48BC we used boundary conditions from a global model which has a spectral directional resolution of 7.5º (48 directions) the same used in the high resolution mesh, hence, no directional interpolation of the spectrum is required."

OBS12 Figure 17. I didn't realize until this figure (at the end of the paper) that the authors ran this regional model for a period of ~16 years. Was that mentioned anywhere? The comparisons prior to that are for much shorter periods, e.g. 1 month.

R12: Indeed in Section 5 we used wave heights from a 28 years database created in the context of marine renewable energies. This database is described in Accensi et al. (2021) which can be found here: https://archimer.ifremer.fr/doc/00736/84812/91664.pdf.
The use of the dataset is mentioned in line 346. The following will be added to make it more clear:

"...Making use of the coverage and improvements in this altimeter product, we analyzed the performance of our mesh over part of the wave hindcast described in Accensi et al. (2021), which was created using the same mesh employed in the present study."

OBS13. 425: Is the satellite altimetry dataset mentioned here?

R13: Thanks for pointing this out, the altimeter data set is in fact missing in the list. Options to access the ESA CCI V1 and V3 (same as v2 but open to public access) will be added in the Data Availability section:

"ESA Sea State CCI altimeter dataset access: https://catalogue.ceda.ac.uk/ and https://cersat.ifremer.fr/Data/Latest-products/

Note that the CCI V2 altimeter dataset is not available to the public. Version 3 will be soon available to public access and is identical to V2 used in this study"