Review of "Development of a forecast-oriented km-resolution ocean-atmosphere coupled system for Western Europe and evaluation for a severe weather situation" by Joris Pianezze et al., https://doi.org/10.5194/nhess-2021-226

This study by Pianezze et al. presents a kilometer-scale atmosphere-ocean coupled system newly developed to improve the forecasts over the northeastern Atlantic and western Mediterranean seas. Additionally, during the 12-19 October 2018 storm event, on one hand, the performances of the new system are assessed and, on the other hand, a sensitivity study of the impact of the coupling is performed.

I believe the development and use of km-scale atmosphere-ocean models should be promoted as it has been proven in many studies that these models improve both forecasts and climate projections. However, in my opinion, this particular study has failed to demonstrate the interest of such numerically expansive modelling suite and thus cannot be published without taking into account the major corrections described below.

Major comments:

(1) Design of the coupled system:

At the strategic level, it is extremely difficult to understand why the modelling system does not use similar grids in the atmosphere and ocean and thus reduce the computations by exchanging the fields between the two grids without any interpolation. In the actual configuration the size of the two grids is nearly identical and the fields are interpolated, so there is absolutely no computational gain and, in my opinion, no justification for such a strategy.

Additionally, the design of the ocean model which ignores the Baltic Sea, the Adriatic Sea and the border east of Sicily is a bit strange as it covers seas that in fact are not modelled and then could have been limited to the previous operational setup (NEATL36). Overall, it feels that atmospheric and ocean grids were designed separately and patch up together for the coupled system.

This leads to my last point. Not imposing SST from ocean model to the atmospheric model particularly on the open Atlantic boundary seems a strange choice. It will definitely create a discontinuity in the SST field imposed on the atmospheric model and avoiding these kind of discontinuities which can translate into numerical "shocks" is a basic modelling concept. In brief, I would strongly recommend to rethink a modelling strategy that fluidly allows the atmospheric and ocean models to exchange fields over their entire domains without any grey zones (as presented in Fig. 1 and A1). At the very least, the authors must merge figures 1 and A1 and discuss at length the different drawbacks of their modelling strategy and how they could be fixed and why they are not.

(2) Evaluation of the modelling system:

First, testing the model on one storm event does not represent de facto an evaluation of a complex atmosphere-ocean modelling suite. It is merely a test of the capacity/performance of the model during a specific event. However, the authors did a nice sensitivity study on

the impact of the coupled system on the results. I would thus recommend to highlight the sensitivity study aspect and not the evaluation aspect in their article. A more appropriate title for their study could be something like: "Sensitivity to coupling of a forecast-oriented km-resolution ocean-atmosphere system during a severe weather situation".

Second, the authors mention the configuration of the operational system actually in use and composed of AROME (AROME-France, 1.3 km-resolution) and NEMO over the Iberia-Biscay-Ireland (IBI) region (NEATL36, 1/36°-resolution). The comparison of the performances of this system with their newly developed system thus seems a mandatory step to show the interest of their developments. This is an important missing part of the study and some clues on how the new model outperforms (or not) the already widely use system should be provided.

Third, the sensitivity to coupling should be presented for all comparison with observations. For example, Fig. 4 should show the spatial differences also for $SST_{OCE\text{-}ifs}$-SST_{sat} & $SST_{OCE\text{-}aro}$-SST_{sat}.

Fourth, the structure of section 4 is extremely hard to follow and overall confusing. In my opinion, for each sub-section, the authors should systematically, first present the performance against observations (for all experiments) and then the comparison between experiments. (e.g. 4. Sensitivity to coupling, 4.1 Sea surface temperature, 4.1.1 performance, 4.1.2 sensitivity, 4.2 Temperature, salinity, height, currents and ocean mixed layer, 4.2.1 performance, 4.2.2 sensitivity, 4.3 wind, 4.3.1 performance, etc.).

Finally, some sub-sections do not have any comparison with observations (e.g. 4.1.2 or 4.1.4). If finding observations in the ocean is a difficult task and should be acknowledge, I believe that the authors can have access to many land-based coastal weather stations that would provide observations of rain but also wind. I would also recommend to check the availability of Argo float measurements or other ocean observations (e.g. CTD) during the time of the numerical experiments. Overall, this study could benefit from a bigger number of observations to assess the performances of the different experiments. Indeed, for the moment, some comparisons done in Fig. 6, Fig. 7, Fig. 8, Fig. 11 and Fig. 12 seem relatively pointless as we truly don't know how each experiment performs against observations for any the compared variables.

(3) <u>Conclusions of the study:</u>

From the presented results, it is clearly shown that IFS performs better than AROME (coupled or uncoupled) (i.e. better comparison for wind speed measurements and comparison with SST buoys similar for all ocean experiments) and thus (again with the limited presented comparisons with observations) it feels pointless to develop a complex atmosphere-ocean model which will use a lot of numerical resources while a NEMO model forced with IFS seems to provide better results. I truly believe that the presented results do not properly reflect the performances of the newly developed model but as it stands the study does not demonstrate the interest of such a model and this should be emphasis in the conclusions and in the abstract.

Specific comments:

The following list of specific comments is pretty succinct as major comments should first be addressed before a more detailed review can be made.

Lines 60-64: Efforts done by diverse research groups in the world to develop km-scale atmosphere-ocean coupled operational forecast models should be mentioned (e.g. Indian and western Pacific Oceans, Hawaii, UK, Adriatic Sea, etc.)

Lines 165: feedback of ocean currents to atmospheric models is not state-of-the-art coupling and may be developed a bit more here.

Figure 1. Please combine with Figure A1 for discussion of the modelling strategy concerning the grids (see major comments).

Figure 2. Scale of the different sub-plot makes it confusing to understand firsthand what the different subplot are representing. Plotting coastline and drawing sub-domains (b, c, d) on top of satellite image (panel a) may help reader identify without effort the geographical locations of interest.

Figure 3. Not sure how this figure is relevant concerning the sensitivity study done in the article. It is obviously expected that during a storm event SST will diverge from initial (i.e. background) conditions. I think this figure is not needed in the main article and may be presented as supplementary material to highlight how the area of interest correspond to the obtained results.

Figure 4. Please include comparison of satellite data with all experiments (see major comments).

Figure 5. The legend on the top of the time series as well as on the right hand side of the figure should be increased in size.

Figure 6. Please remove wind from the plot and put wind comparison in a separate figure dealing with wind comparison specifically.

Figures 7 & 8. Interesting figure only if all experiments presented are previously compared with measurements over the full "water column" (e.g. Argo floats, CTDs) in the different zone presented. Otherwise, no conclusion can be drawn except this experiment is similar or different from the coupled one.

Figure 9. Please include comparison $wind_{OA}$ - $wind_{IFS.}$ Labels on the top of figures should also be increased in size.

Figure 10. As Figure 2.

Figure 11. Please include $_{rainOA}$ - $rain_{IFS}$

Figure 12. As for Figures 7 & 8, interesting only if all experiments are previously compared to observations in the different zone presented.

Data availability: From my understanding of the EGU publication rules (I may be wrong), the specific model results (atmosphere and ocean) used in this study should be publicly available as

anyone should be able to reproduce the presented findings of the article, but authors mentioned it is only “upon request”.