Review of “Km-scale regional coupled system for the Northwest European shelf for weather and climate applications: RCS-UKC4 ” by Berthou et al for Geoscientific Model Development (GMD).

This manuscript presents an updated version of the Regional Coupled Suite – UK Coupled (RCS-UKC4) which includes numerous updates of the different models and parameterizations used with notably new addition of river model and biochemistry model as well as the ability to do ensemble forecast. The papers aims to evaluate the new coupled system compared to its predecessor and/or observations available through validation of different weather scenario, i.e. hindcast, forecast, specific weather events like storm, marine heatwaves and meteotsunamis. The paper is generally well organized and figures supports the results described in the text and I believe it fits very well the scope of GMD. Pending some revisions I would consider this manuscript for publication. Overall, I think the authors might want to carefully read the paper as some details or descriptions are not always very clear, to eliminate any remaining inconsistency, typos and special care to figures where labels and captions could be much improved. See more specifics comments below.
My main comment here is regarding the Figure 6a of the drag coefficient against wind speed. How does the modeled Cd in Fig 6a compares to actual observation/widely used parameterization i.e. COARE3.6, ECMWF or else? Up to 20 m/s the mean values of Cd are relatively well known (in open ocean conditions). Somehow RAL3 and UKC4 on Fig 6a seem to have unrealistically high values of Cd on average reaching 0.004~0.005 while RAL2 seem to have more acceptable range of values. While Cd can have a relatively large scatter due to wave modulation, generally speaking Cd is usually not greater than 0.003 on average below 20 m/s, which I believe would be consistent also with Gentile et al.’s study cited here. So, higher drag will give you weaker winds here which might helps with biases, but I wonder if this bias correction happens for the wrong reasons ? are there any drag/stress observations in this region or any reasoning that would justify the Cd used in UKC4 or RAL3 over RAL2 ? I think this here needs a bit more clarification.

Minor comments:

-Line 160: It would be nice to more clearly associate cases listed in table 2 to each of these evaluation points, while some are easy to see (heatwaves, meteotsunamis) the ones in ensemble forecast are less straightforward to associate with the different bullet points.
Also, I found it hard to understand what were exactly the run start and end periods, because there are multiple runs/periods it is a bit confusing, some clarification would help, i.e. in table 2 the ‘date’ column refers to the actual event date or a run period as the title states “experiments run”.

-Line 196: “is a foundation SST product (Good et al., 2020), so it should be compared to the model’s minimum diurnal temperature” : So does this means the bias mentioned in the text are done this way ?

-Figure1: Maybe it would be good to indicate names on Fig1b (Celtic Sea, Channel, Southern North Sea and so on) as they are used throughout the paper but for someone unfamiliar with the region and these specific delimitations it could be hard to follow.

-Figure2: ‘averaged over the Northwest European shelf’, is this the full domain in Fig 1, delimited by the purple line or just inside the orange dotted line ? I don’t think this has been clearly stated. Also y-axis says "wakelin_shelf" but it’s not clear what it is. Also, would the results of this figure be very different if taken over the different hydrodynamic regions ? Please indicate the line colors in the caption too (OSTIA (black lines) etc…). Why is there no CLIM in the panel (a) ?

-Figure 5: The caption could be improved here, (e) is indeed sea surface temperature but (a-d) are SST biases, which is not very well stated. Also, DJF, MAM, JJA and SON acronyms are not mentioned in the caption.

- On the drag coefficient part: Additionally to my main comment, it is not very clear what are the differences in drag here between the three ? RAL2 has a constant Charnock, RAL3 variable Charnock with the wind and UKC4 has a Charnock from the wave model, is that right?

-Figure 6: Please revise caption. (a) is the drag coefficient not (e). What are the black circles in (a,b), and longitude/latitude would be nice on the map. Plotting mean lines on 6a would be useful as well. Titles of subplots say RAL3 and UKC4 but caption says RAL3.2 atm-only or coupled, not very clear..

Figure 9 and further: So far, figures have used (a,b,c,d…) to refer to subplots but from figure 9 and after it got lost. For consistency and better clarity of caption and in-text description I would suggest to keep using (a,b,c,d…) for subplots.
Specifically for Figure 9, maybe it would be nice to locate the 6 gauges on a map.

-Line 357: “five storms”: maybe include the names here for clarity.

-Line 378: “six cases”: what are the six cases? The paragraph before mentioned looking at 5 storms, here it only shows 3 storms (which ones ?) and three other cases ? Why ?

-Figure 13: axis says wind speed but text and caption seems to refer to wave height ? This need to be corrected if needed in both text and figure whether it shows wave height or wind speed...?

-Line 455: maybe just need clarification here, it doesn't impact the cost of the coupled system while increasing the model cost ? how are those different ?

-Line 456: Just wording here “growth and decay” terms usually apply for Hs I think, not sure what it means for Ws and SST here.

-Figure 15: x-axis only shows ‘00’ ticks, it doesn’t make it really easy to read, labels could be better.

-Figure18: “Wakelin regions **(Figure 1b)**”. Also, why only a subset of the Wakelin regions and not all ?

-Figure 19: Similarly why not all regions and why not the same as in Figure 18 ?

-Line 558: “simple mixed layer heat budget computed on the averaged mixed layer depth”: how do you calculates mixed layer heat budget and mixed layer depth ?

-Line 563: “having a systematically steeper slope than the other regions in the SST/flux relationship”, is this accurate? it’s not particularly obvious to me from the figure.

-Line 600/601: maybe these locations can be placed on the map Figure 21 as for Portsmouth?

-Figure 20 : Please indicate the gray line in the caption as well.

-Line 646: “UKC4-1h (dashed lines) clearly misses this pressure drop,” it seems on the plot that the dashed lines are getting the drop in MSLP for every lines no?

-Figure 21/22: longitude/latitude would be nice on the maps.