Author response to RC2: "Evaluating the impact of atmospheric forcing resolution and air-sea coupling on near-coastal regional ocean prediction" by Huw W. Lewis et al.

We thank RC2 for their particularly constructive and detailed review comments and have amended the manuscript in response. Their contribution has also been acknowledged in the revised manuscript. In addition to correcting the list of 'Other comments' provided (see below), and a review of the full document in light of RC1 and RC2 comments in general, the following substantive changes have been addressed:

- Better justification for only considering one month in summer in this study,
- Improved linkage of relevance of this finer-scale ocean work with suggested references from the regional climate modelling community,
- More explicit reference to the positive but secondary added value of wave coupling,
- An updated presentation and discussion comparing SST simulations with OSTIA,
- Discussion of the initial condition bias, and opportunities for use of ocean analyses in future research,
- Better attribution of some heat budget differences to the different SST state in forcing and coupled atmosphere simulations,
- Expansion of the discussion on partially coupled results,
- More careful reference to resolution and physics changes between the global NWP and km-scale regional atmosphere forcing.

Author Response to RC2 *General comments*:

1. *The justification for only considering one month in summer is however missing.*

We agree, and have briefly provided a better explanation of the motivation in the last paragraph of the Introduction on p3. This change is also in line with a similar request from RC1 to more clearly articulate these choices. In brief, we selected to assess the July 2014 results and focus on a relatively small part of the model domain in order to provide a more detailed discussion of the impact of atmospheric forcing and coupling on near-coastal SST results, for a period identified by the overview discussion of Lewis et al. (2018b) as being most sensitive to coupling.

2. *Such kind of evaluation is quite often done in the regional climate modelling community (e.g. Béranger et al., 2010; Akhtar et al. 2018a,b) and it could be relevant here to put forward the novelty of considering ocean forecasts with a very fine coupled system and the inherent difficulties.*

The role of atmospheric forcing and coupling is indeed more routinely discussed in the context of typically coarser-scale regional climate modelling activities and it would be useful to contrast this with the present study. This has been addressed in an updated Introduction in the revised manuscript.

3. *The added value of the wave coupling is also not so well highlighted.*

We summarise in Sect. 3.1 that “there is some additional value evident from coupling information of the wave state to ocean and atmosphere components in CPL_AOW (MD = 0.20 K), although this is of secondary importance to the impact of either changing atmosphere resolution or ocean-atmosphere coupling”, and in Section 3.3 that “the influence of wave coupling feedbacks is generally small at this time of year”. The aim of presenting both CPL_AO and CPL_AOW results is both to demonstrate the performance of a fully coupled regional system, i.e. with wave coupling as an important component of the earth system at these scales, but also provide a more traceable comparison of the impact of coupling relative to the ocean-only results. This also addresses one of the comments of RC1. In light of the comment above, a new summary sentence has been added to the Conclusions to again highlight the relatively minor impact of wave coupling for this region and time of year.

4. *The key point in the ocean flux forcing which is the SST inconsistency between the one simulated in the ocean model and the one used at the surface boundary of the atmospheric model. It obviously controls the differences in several heat budget terms and very likely the differences in wind between CPL_xx and FOR(_HI) but it is never mentioned here.*

This is a fair challenge and an omission of the reviewed manuscript. The new comparison of SST against OSTIA (which provided the SST surface boundary of the atmospheric model) in the revised Figure 3 helps to highlight this point, and RC2 is correct to highlight the close spatial distributions of changes in SST and the change in sensible heat and latent heating in particular. This is now addressed in the revised manuscript in discussing the heat budget results of modified Fig. 6.

5. *The robustness of the SST improvement (with the higher resolution forcing and coupling) that appears a little altered by the fact that it seems to be more a spinup effect, with a reduction of the initial bias during the first days of simulation.*

We agree that longer-term simulations of the fully coupled simulation would be required to evaluate how robust the improvement is. However, we argue from Fig. 2 that the improvement becomes well established and is relatively constant by the second half of the month at least. This motivates us to discuss the 10 day period considered in most detail as being representative of a relatively steady state. Another interpretation of the comment on spinup, is that the simulations diverge relatively quickly in the first days of simulation, driven only by a change of atmospheric forcing or introduction of the atmosphere-ocean feedbacks.

6. *The use of partially coupled sensitivity experiments seems very promising, but their results are too briefly discussed.*

We appreciate this encouraging comment, and have provided some expansion of both the motivation for running the partially coupled sensitivity results and their assessment in the revised manuscript (Section 3.4). We balance this with addressing the concern of RC1 that they had less value given that the heat and wind terms are by definition not in equilibrium in these simulations. They aim to help better attribute the previous results described.

7. *I'm finally interrogative about the large impact of the higher resolution which is always highlighted by the authors instead of the impact of the physics (qualified as a smaller impact). But for me this is connected, especially over sea, far offshore. I suggest to clarify or discuss more this point.*

This is a valid concern, and a topic of discussion for the authors in the original assessment of the results in this study. The conflation of both resolution and physics changes between global NWP and what is characterised as the 'high resolution' forcing makes this a particular challenge. We have been encouraged by this comment to be more precise where possible in the revised manuscript as describing the global-scale and regional-scale forcing as being indicative of two readily available sources of atmosphere information, with the regional-scale also able to be applied with feedbacks. Changes have been made where relevant in the revised text. The paper title has also been updated in view of this comment.

The main reason for quoting the spatial grid resolution as dominating over physics changes originates from considering the larger spatial variability of forcing terms in FOR_HI than FOR_GL for example. We aim to be more careful in the revised manuscript that it is not clear we can attribute this directly to a resolution change alone.

Author response to RC2 Specific comments:

8. *Page 6, lines 2-3: "This indicates improved SST prediction can be achieved for the NWS when applying the high resolution." In my opinion, this conclusion is too rapidly set. Figure 2 shows mostly the stronger cooling during the first days of simulation (till 5 July) in FOR_HI, CPL_AO and CPL_AOW, correcting more efficiently the initial warm bias. Considering the overestimation of the wind in the higher resolution forcing (coupling), this is a possible ocean response to the initial shock with a larger effect of the vertical mixing. How do the mixed layer's depth and thermal content evolve?*

While the statement as written is correct (i.e. the SST Mean Difference for FOR_HI is lower than found for FOR_GL results), the tone of this line has been modified in the revised manuscript to be less definitive at that point, as we accept that it can be read as too definitive a conclusion.

The vertical profile in revised Fig. 4(b) shows the FOR_HI, CPL_AO and CPL_AOW simulations to have deeper mixed layers, consistent with RC2's comment, and with the persisting warm bias in FOR_GL. However, later discussion of the pCPL_RAD results show that by only applying the higher resolution (overestimated) winds does not diminish the warm bias in the same way, rather it increases over the first days of simulation and settles at order 2 K warmer than observations (Fig. 12). This further supports the value of the partially coupled simulations in drawing conclusions from the study (see response to comment 6 above).

9. *If possible I suggest to test new initial conditions, more realistic, such as such as ocean analysis that are available in the CMEMS catalogue or at least a larger discussion about the relative importance of the forcing compared to the model initialisation.*

We expand on this valid point in the revised Conclusions. While it is not practical to test new initial conditions in the present study (e.g. covering the period of interest), the relatively recent implementation of the 1.5 km resolution AMM15 ocean model configuration to provide CMEMS NWS MFC products (e.g. Tonani et al., 2018) does now offer a source of ocean analysis from the same system as used here, which should be valuable to support future research work.

10. *Between 18 and 24 July, it seems there is a warming in FOR_GL whereas SST is stable in FOR_HI and CPL_xx. How is it explained?*

The period highlighted by RC2 is apparent both for the domain-wide results in Fig. 2 and to some extent reflected in the location-specific comparison with observations at L4 in Fig. 4. While we do not provide detailed consideration of the evolution of FOR_GL results, the difference between FOR_GL and FOR_HI net shortwave radiation over period 20-30 July 2014 in the revised Fig. 5(e) highlights the relatively higher solar heating in FOR_GL described in the paper. Considering only the Celtic Sea region, the difference in Fig. 5(e) is focussed towards the south-west approaches, which coincides with the region of largest warm bias over the same period illustrated in the revised Fig. 3(d).

11. *Page 8, lines 8 to 15: "(. . .) The impact of coupling on (. . .) QLW is dominated by random changes in the spatial distribution of convection. (. . .) There is also some evidence that the latent heat flux is increased in those near-coastal regions identified as being cooler in CPL_AOW than FOR_HI, where the coupled simulation SST was in closer agreement with observations in Fig. 3(c)." To well consider the differences in the heat budget terms between the coupled runs and FOR_HI, the comparison of the CPL_xx and OSTIA SST field(s) must be shown. I think it explains at first order most of the differences found in the long wave upward radiation, latent and sensible heat fluxes. The differences in the convective cloud location play also, but at a second order. The last sentence is particularly confusing for me as it mixes information about LH, differences in SST simulated by CPL and FOR_HI and the validity of the CPL SST against observations. But what about the comparison between OSTIA and the in-situ observation in this region? Please revise.*

Comparisons between the FOR_GL, FOR_HI and CPL runs relative to OSTIA are now presented in a substantially revised Fig. 3. Time series comparisons relative to OSTIA are also now provided in Fig. 2 and Fig. 4 following this encouragement. The results discussion in Sect. 3 has been amended where relevant to describe these comparisons. We consider this provides a more coherent discussion than the original manuscript in line with the comment from RC2 above.

12. *Figure 6 (i-l): Please, adjust the scales to better show the differences. To be fair, it might be shown as relative differences (in %) instead. Very likely, the differences in the wind field are also controlled by the differences in SST. See Chelton and Xie (2010) or for example Lebeaupin Brossier et al. 2015 (Fig. 8a), Meroni et al. 2018 (Fig. 6).*

A version of Fig. 6 considering % differences was also prepared for the original manuscript, but changes were disproportionately dominated by regions where mean fluxes approached 0 Wm^{-2}. The impact of changing atmospheric forcing and coupling has now been separated (following RC1) across updated Fig. 5 and Fig. 6, where the comparison of heat budget terms are presented on a clearer scale. We concur on the difference in wind field being controlled by differences in SST. See also RC2 comment and response on p11, line 13-15 below.

13. *Page 9, lines 20-24: "This provides some evidence that the differences…is driven mostly by the change in grid resolution and the change from parameterised to explicitly represented convection,…."*

    Page 10, lines 14-16: "The contrast between the spatial variability of wind speed between FOR_GL and FOR_HI further supports the assessment in Sect. 3.2…"

    I cannot really capture where the contribution of the high-resolution can be separated from the physical behaviour/parametrisations of MetUM between the FOR_GL and FOR_HI forcings. I mean, far from the coasts, there is no reason for these differences apart the MetUM physics? In addition, connections between resolution and physics exist. Some physical parametrizations may depend on the grid resolution (and time step). Please, clarify how you distinguish the relative importance of physics compared to the benefit of a finer grid mesh.

This reflect the RC2 Comment 7 discussed above, and is a valid query. Some of the key atmosphere physics differences are outlined in Sect. 2.1 As described in the response to Comment 7, the manuscript has been modified to take more care in describing the change of atmospheric forcing in terms of 'global-scale' and 'regional-scale', noting the link between grid resolution and physics choices. In particular, as noted in the paper, the main difference is in the treatment of convection explicitly at 1.5 km whereas it is parameterised in the global atmosphere model.

Author response to RC2 Other comments:

• p1, lines 14-15: "Observations. . . data". Please, revise the sentence as you do not only consider L4 observations. . .

Revised – we aim to highlight use of both the 'routine' operational observations along with use of L4 as having co-located observations of atmosphere and ocean.

• p1, lines 21-22: "...by global-scale NWP (0.7 K in the model domain) is shown..."

Corrected in the revised manuscript.

• p1, line 23: "...reduced (by 0.2K)."

Corrected in the revised manuscript.

• p2, lines 28-29: "A number of studies. . . (Lewis et al., 2018a)": revise citation.

It is not clear what is intended by this request. The intention of this citation was to really indicate Lewis et al. (2018a) as a source of further references. We have modified the citation to reference the more obvious review paper by Pullen et al. instead, mentioned elsewhere in the Introduction. We hope this might be what was intended by RC2 here.

• p3, line 7: "...for one of those periods in July 2014." The motivation(s) to dedicate this study to this reduced period must be given here.

We agree, and have briefly provided a better explanation of the motivation in the last paragraph of the Introduction on p3. This change is also in line with a similar request from RC1 to more clearly articulate these choices.

• p3, line 30: "...describing the surface heat and water budget..."

Corrected

• p3, line 31 "...NEMO using the 'flux formulation'...": Where (and how) is computed the wind stress?

In the configuration used in this study, key_shelf is used in NEMO, and the wind stress is computed within NEMO based on 10 m wind components rather than applying the atmosphere model computed stress directly. This is clarified in the revised manuscript and a reference provided.

• p4, line 8 (and lines 17-18): "the wave-dependent roughness Charnock parameter of 0.085 is used.": Could you precise if it is α or z0? If it is a constant, it is not wave dependent. . .?

This is *α*. This sentence has been revised to clarify we mean a constant value used.

• p4, line 17: "...assumed zero and a constant value..."

Coorrected

• p4, line 31: Valcke et al. (2015) is missing in the references list. Moreover, I think the citation for OASIS3-MCT is Craig et al., 2017.

This correction has been applied in the revised manuscript.

• p4, line 31: "...all information exchanged...". A brief list of the exchanged fields could be useful.

This has been clarified with additional text at the end of Section 2.1

• p5, lines 7-9: I am happy to see here this comment concerning the 'double penalty' effect that is indeed of primary importance when comparing high-resolution modelling results with observations.

This is indeed an issue for evaluation of all such systems, and thank you for the supportive comment.

• p5, end of section 2.2: I am a little surprise there are only GTS data considered. Some other kinds of data could be available on the CMEMS website, in particular profilers to examine the vertical stratification or satellite data that allows a 2D coverage at the surface. Was it a choice to exclude them? And if yes, why?

We have focussed the analysis in this paper on the Celtic Sea region, and aim to make most use of the L4 buoy observations given the rare co-location of ocean and atmosphere observations, along with the radiation measurements. We also considered the co-located CTD observations from 28 July sufficient to provide some indication of the vertical profile in this region. The in-situ data on the CMEMS website (e.g. http://www.marineinsitu.eu/dashboard/) are in general consistent with those displayed in Fig. 1. We appreciate the encouragement to compare SST results with OSTIA, based on satellite data, which are now included in the revised manuscript (e.g. Fig. 2, Fig. 3).

• p5, lines 26-27: "Figure 2 demonstrates that all ocean simulations had the same initial conditions.." This is not something that must be demonstrated. That must be said before in section 2.

This sentence has been updated in the revised manuscript, although we consider it useful to remind readers of this from Fig. 2, particularly given that the later analysis focusses on the later period when the 4 experiments have diverged. The initialisation is indeed referenced in Section 2.1 to indicate all simulations have the same initial condition.

• p6, line27-28: "On several days (e.g. 20, 21, 23, 26 and 29 July) a tidally-forced heating signal of about 1 K is apparent." Well, it is not so apparent it is a tidally-induced heating or if it is a diurnal cycle.

This sentence has been revised to be 'tidally-dominated', while we agree there will be some influence of diurnal cycle at this time of year. The temperature range observed at L4 is large – greater than 1K on some days in fact, and there is observed evidence of 'double peaks' on some days through the series. We also consider the phasing of the time of maximum temperature to be progressively delayed from around noon on 20 July to late evening on 25 July for example.

• p6, lines 30-31: "The SST variability of FOR_GL is in general stronger than observed..." Where is it shown?

This sentence referred to the temporal variability of simulation results at L4 shown in the new Figure 4. In addition to being biased warm, the FOR_GL results show larger diurnal range than other simulations and than observed. This line has been revised in the updated manuscript to clarify that we mean temporal rather than spatial variability here.

• p7, line 19: "Surface heat budget..." Please change also everywhere after: 'Energy' can be potential or kinetic. . . 'Heat' is more precise.

This has been revised everywhere mentioned through the manuscript.

• p7, lines 29-30: "...and from CPL_AO and CPL_AOW coupled systems...": The flux fields for CPL_AO and CPL_AOW are not shown in Figure 6.

The comparison of coupled results with FOR_HI have been separated a little from the FOR_GL vs FOR_HI comparisons, following the suggestion of RC1, as reflected in splitting out new Figure 5 from new Figure 6. The manuscript has been revised to reflect the updated Figures, and the required correction identified here has been removed as part of this.

• p8, line 20: ". . . increased cloud cover on 24 July..."

This has been updated in the revised manuscript.

• p8, lines 26-27: "The warm surface temperature bias of FOR_GL at L4 appears to result despite rather than because of this difference however." Maybe mixing (i.e. cooling by entrainment) is also lower in FOR_GL?

Rather than offer a detailed discussion here, we are simply noting that the SST results cannot be well explained by looking at the local energy balance terms within a relatively small area around the L4 location, as shown in Fig. 7(b). Rather, the results become a little clearer when assessing the atmospheric forcing over the broader Celtic Sea region (Fig. 8). This sentence has been revised to clarify this.

• p9, line 8: "...consistently higher during night time...". Could you explain more this result?

The result described in line 7 and line 8 – i.e. higher net radiation from FOR_GL (contributing to higher SST) is resolved later on p9 from around line 20, where we relate the mean differences to lower spatial variability (lower standard deviations). This section has been revised further in the updated manuscript to attempt to clarify these discussions, noting earlier comments.

• p9, lines 18-19: "...both day and night..." ?? "...but typically of order 20-50% lower. . ." Where the '50%' comes from?

As requested by RC1, the time series of spatial standard deviation plots are now included in a revised Figure 8. This illustrates the substantially reduced standard deviation of radiation in FOR_GL relative to other configurations. The difference between daytime maxima through the period shown is considered to be of order 20-50% reduced.

• p9, line 27: Delete "the accumulated ".

Figure 8 has been revised, and now provides results as accumulated heat budget terms. This provides a clearer illustration of the differences between FOR_GL and FOR_HI forcing than mean values.

• p10, line 17: "The atmospheric forcing..."

This has been corrected. The original intention of "ocean forcing" was "The forcing of the ocean….", but this suggestion is clearer.

• p10, line 25: ". . . than FIX_HI..." FOR_HI?

Corrected in the revised manuscript.

• p11, lines 13-15: "Some evidence of the link between SST and near-surface atmosphere conditions within the coupled system was discussed by Lewis et al. (2018b) in considering the relationship between near-surface stability and wind speed over the ocean." How this relates to the sentence before? More details or a summary of Lewis et al. (2018a)'s conclusions about the SST/stability/wind relationship would be useful.

This section has been developed further in the revised manuscript, noting in particular RC2's comment 6 noting this section was too briefly discussed in the original. In summary, we argue that maintaining a feedback between SST, near-surface stability and near-surface winds is required.

• Tables 2/3: Replace FIX_xx by FOR_xx

Thank you, this has been corrected in the updated manuscript.

• Please revise Figure 1: The colour scale for bathymetry in a is blank. What is the 'UKV' atmosphere grid? What are the small black and red points in b?

The original Figure 1 colour scale in (a) was attempting to reference the contour lines off-shelf. These have now been made thicker in the revised manuscript. The caption text has been updated to clarify what was meant by the 'UKV grid', and the size of symbols referenced in the caption – the small points indicating points where there are a limited number of observations available over the selected period.

• Figure 2: If possible add the OSTIA SST error time-series.

Thank you for this suggestion. The comparison between daily OSTIA SST with in-situ observations is now included in the revised Figure 2. The OSTIA SST error has a strong diurnal signal given that it is a daily SST product, but comparisons with in-situ observations are hourly to be consistent with the model vs observation comparison. OSTIA data have also now been used as a reference in the revised Figure 3, and an OSTIA SST time series at the L4 location has been added to the revised Figure 4(a).

• Figure 6: Please, adjust the colour bars in i, j, k, l.

Figure 6(i-l) have now been pulled out into a new Figure 6 in the revised manuscript, focussing only on the impact of coupling (CPL_AOW-FOR_HI), with revised colour bars and clearer plots.

• *Figure 7: Add the colour legend for b (which simulation is the blue line?). Larger plots can also help to distinguish more the time-series in c.*

Corrected in revised Figure 7(b), and updated Figure 7 to have larger and clearer plots.

• *Figure 12: "...between 20 and 30 July 2014..."*

Corrected in revised manuscript.