Weather Clim. Dynam. Discuss., author comment AC1
https://doi.org/10.5194/wcd-2021-46-AC1, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on wcd-2021-46

Jennifer Saxby et al.

Author comment on "Simulations of Bay of Bengal tropical cyclones in a regional convection-permitting atmosphere–ocean coupled model" by Jennifer Saxby et al., Weather Clim. Dynam. Discuss., https://doi.org/10.5194/wcd-2021-46-AC1, 2021

Reply to the reviewers' comments

We thank both the reviewers for taking time to consider our manuscript and providing detailed and helpful comments. We are pleased that both reviewers thought the manuscript was well written and we note the particularly positive response from Reviewer 1.

Response to reviewer #1

G1: The description of the coupled modelling system is done by referring to an in-preparation paper (Casillo et al). Some of the details in that paper could be important for the interpretation of the coupled results but since it is not available yet then these details are obviously missing. I would like to see a bit more details of ocean model like for instance the vertical resolution in the upper ocean.

We had hoped that the Castillo paper currently in preparation would be published by the time our manuscript was ready for publication, although we anticipate the manuscript to be citeable from GMD Discussions shortly. We are in any case happy to add more details of the model set up to our manuscript. The ocean model has 75 vertical levels and a non-linear free surface, with a vertical grid using a hybrid terrain following "z-s coordinate" system.

G2: There is also very little information on ocean spin-up used to produce the ocean initial conditions besides the reference to the in-preparation paper. Since we are looking at tropical cyclones then the quality of the winds used as forcing for this spin-up run could influence the ocean stratification in ocean initial conditions leading to differences in coupled model performance.

We will provide more details regarding spin-up in a revised manuscript rather than only cite from the Castillo paper in preparation. In-depth analysis of the ocean model was carried out by a colleague at the University of Reading (Valdivieso et al., 2021), and we refer to her results written up as a project report in our manuscript so as to focus this paper on the atmospheric response. However, this report is not published in peer reviewed literature. Therefore, we will include more information from the analysis of the ocean model such as ocean stratification to fill in the gaps.

G3: It is not clear to me what the SST used in the coupled model actually is. Is it the top layer from the ocean model, the mean of the levels in the first 10 m or something else? Is there a diurnal cycle in the coupled SST or not?

The SST used in the coupled model is the top layer from the ocean model, nominally within 1 m of the ocean surface. This is in contrast to the foundation SST provided in the OSTIA observation-based analysis used as a daily-updating surface boundary condition in the atmosphere-only control simulations. This difference can be highlighted in a revised manuscript. Given that the ocean model component of the coupled model is evolving the ocean state through the simulation, there is a diurnal cycle in the coupled SST, which is apparent in the SST comparison to RAMA buoys shown in Figure 8 for example. Further analysis of the ocean is shown in the referenced report Valdivieso et al 2021.

G4: Related to G3. The authors chose to compare with ARGO observations rather than surface drifters which are reporting more frequently. I accept the argument that the profiles used have almost constant temperature in the uppermost 10 meters so it fit well with the prescribed SST used for the ATM runs, but how does it relate to the SST of the CPL runs?

The paper includes comparison of SST against both ARGO and RAMA moored buoys, which report more frequently (see Table 3). Figure 8 provides a comparison of simulated and analysis SST relative to RAMA buoys, to give some assessment relative to more frequently reporting in-situ observations, while the use of ARGO observations in Figure 9 is selected to increase the data coverage within range of each TC. We would agree that a revised manuscript should clarify that the prescribed SST used in the ATM runs is representative of a foundation SST whereas the SST of CPL runs represents the uppermost NEMO ocean model level. This indeed further complicates comparisons with in-situ and profile observations and should be highlighted as a known limitation in a revised manuscript.

G5: I think that the authors could have compared the ocean stratification of the ARGO floats and the RAMA moorings to both the ocean initial conditions and coupled integrations to gain some insight on the ocean performance of their system, but maybe that is a separate paper?

This analysis was performed and discussed in the cited reports by Valdivieso et al. (2021), based on an assessment of the ocean component over a longer period of time linked to a research observational campaign. The details are beyond the scope of our manuscript. However, we will briefly comment on their findings in a revised manuscript.

M1 Page 5 line 114: It is not completely obvious to me that degrading the OSTIA SST using in the ATM runs with 1 day is a "fair" benchmark to compare with the CPL runs, since this could not be done in an operational settings. Maybe this could be rephrased?

It was not obvious what the best choice of atmosphere-only control experiment design would be, but we are confident we made defensible choices. We thought fixed SSTs throughout such long simulations (much longer than forecasts would be typically run for) could potentially be considered 'too harsh' a control, so we opted for daily updating SSTs, which are available at the time of daily initialisation. Discussion of relevant control simulations for such experiments could usefully be added to the revised manuscript, given relevance to similar modelling studies by other researchers.

Response to reviewer #2

The purpose of the manuscript was to investigate coupled simulations of a new coupled environmental prediction system that has been used for the first time to simulate tropical cyclones (TC) in the Bay of Bengal. The aim was to assess the capability of the model to

represent TCs and their evolution, and to understand the impact of air-sea coupling in a convection-permitting regional model. The development of km-scale coupled prediction tools is a growing area of research of interest to a range of applications, and so we consider the paper to be of interest to atmosphere and ocean modelling communities.

The intention was not to present a new operational forecast system, and so the model configurations were not run in the same way as they would have been if they were in 'forecast mode'. In particular, the SSTs in the atmosphere-only simulations were from daily updated observation-based analysis and the atmospheric lateral boundary conditions from updated global forecasts, both of which would not be possible in operational running. The simulations were also run for up to 13 days to cover the whole lifetime of the TCs, considerably longer than it would be sensible to run free-running forecasts for (3-5 days is more typical for this). We have stated these points in the manuscript (see Section 2.1: Model configurations), and we refer to the runs as 'simulations' rather than 'forecasts' throughout; however, we can make this clearer by modifying parts of the abstract, introduction, and the title of the manuscript. In particular, we would clarify the role of the paper in providing new insights on the sensitivity of process representation of air-sea interactions within a km-scale simulation system, rather than a forecast model verification paper.

(1) The model suffers from huge phase errors which could lead to poor predictions of timings of land-fall, critical for forecasting. While those may not be obvious in Figure 1, those errors do show up in several of the other diagnostics. Track errors of Vardah for 05/12 is on the order of 1000 Km and, in general, for Gaja and Phethai are huge as well and may not be useful for forecast purposes (please compare the tracks from the operational models such as WRF or GFS in that region for those cases). Such huge track errors could be due to some large-scale issues which have not been analyzed here. Very likely those are coming from the driving global model. Why not plot the track errors from the driving model? Very likely the regional configuration is so small that such longer rage predictions without two-way interactions may not even be possible. Very likely this regional configuration cannot predict tracks beyond 48 hours. Unfortunately, the number of cases isn't statistically significant to prove or disprove that the regional configuration is capable of predicting beyond 48-72 hours.

(2) The problem with the track and landfall timings affects intensity errors as well. If a formal intensity verification is done for Vardha, Gaja, and Roanu, all of them will be extremely huge for a forecast model. I would encourage the authors to first look at the large scale and look at why there are such huge track and intensity errors. Very likely those are coming from phase errors originating at the boundary of the regional domain. Is the domain too small for forecasting beyond 48-72 hours? I suspect the WRF or other operational models used over that region should have provided much better statistics for all these cases. I think the author needs to make an analysis of why the track and intensity errors are so huge? I doubt if such errors could be useful for the forecasters.

We acknowledge that the track errors (and therefore intensity errors) in these regional model simulations for TCs Vardha, Gaja and Roanu are very large for some of the initialisation times, but these were the simulations with the longest lead times and initialised very early in the life of the TC. The track errors did not exceed 500 km until 3 days into the earliest simulations for Vardah, until 5 days into the simulation for Gaja and until 3-4 days into the simulation for Roanu. In the case of Roanu, the TC was close to the coast and therefore a small error in the track could take the TC erroneously over land where the environmental conditions are totally different, which then had a large impact on further development of the storm. The examples given by the reviewer of track-errors >1000km in Fig. 3 only occur 6 days into the simulations and for only one of the ensemble

members of each model for Vardah and Gaja. To assert that other forecast models can better simulate those systems is not comparing like-with-like. Many simulations within their first 3 days only had track errors of the order of 200 km, similar to those found in operationally-used forecast models and similar studies in the literature. Understanding why these track errors were so large in a few cases was beyond the aim of the manuscript and not a major concern for the aims of this paper given the lead times at which they happened and that these simulations were not run as forecasts.

The use of a wide variety of cases and initialisation times was designed to assess initial system behaviour, and to examine coupled processes in TC evolution; to report on only cases where errors were low would be to 'cherry-pick' results, although we note that cases with low errors are more useful for assessing the representation of TC dynamics in the model in comparison to observed characteristics and therefore limit some of our analysis to these cases. Given the long lead times, position errors are low and landfall times are accurate.

The reviewer suggests that we plot track errors from the global MetUM driving model. This highlights that one source of error is the global driving model (e.g. state of TC initialisation), however, we clarify that the atmosphere initialisation and boundary conditions are common to all experiments, and thereby the paper explores simulation sensitivity for a given driving model. In this context, providing a discussion of global model characteristics is considered beyond the scope or focus of the manuscript. In general, we find that the IND1 regional model produces improved tracks compared to the lower-resolution MetUM global model simulations. In the case of TC Fani, the global simulation track errors exceeded 400km after 1.5-2.5 days for the earliest 2 initialisation times and reached 600km after 6 days, whereas for convection-permitting ATM model simulations, track errors never exceeded 330km, showing the improved tracks due to increased resolution in line with previous studies (e.g. Gopalakrishnan et al 2012). We could include illustrative plots of this as Supplementary Material in the revision if required.

We also note in the manuscript that few studies are carried out on TCs in the Bay of Bengal region; landfall timing and position errors from this study do compare favourably to results from previous lower-resolution simulations of Vardah, Roanu and Fani (Nadimpalli et al., 2020; Singh et al., 2021), when we compare similar lead times (see Discussion, lines 553-556).

By not simply looking at a single lead-time, the study demonstrates the relative sensitivity to coupling vs initialisation state and that some TCs are more affected by this than others. It was not clear to us prior to the project where the largest sensitivities would lie. The aim of this work was partly to investigate the sensitivities and compare the impact of different model initialisation times and setups on TC track, intensity, and structure. We will bring this out more clearly in a revised section 3.1.

We do not think the domain size of the regional model is a problem because the lateral boundary conditions are derived from daily updating global forecasts (not a 13-day global model free-running forecast, see also response to point 4 below) and the domain is based on that of the NCUM-R operational atmosphere configuration (Jayakumar et al., 2017; Mamgain et al., 2018; Jayakumar et al., 2019) although the domain extends slightly eastward and southwards. We will add more details of the model set up given that the Castillo paper is not yet published. The runs are convective-scale and so the high resolution means there is a limit on the domain size that would be feasible to run in terms of computational expense and run times.

(3) I am also concerned about the model comparisons that have been made between the atmospheric only (one-way interactive, to be precise) and the coupled versions. While the atmosphere only version is run at 4.4 km, the coupled version is run at 2.2 km. While we

expect a higher resolution to provide improved intensity predictions, this has been the reverse here. The coupled version is producing excessive cooling. I don't see an analysis of why the ocean model is reducing excessive cooling. Is there a problem with the coupling? That also begs the question about the surface layer parameterization scheme used in the model. What are the Cd and Ck functions used? That could have a huge impact on the pressure/wind relationship as well.

We use the same atmosphere configuration and resolution across all experiments. The CPL model configuration was run at the same atmospheric resolution (4.4 km) as the ATM model resolution. It was only the ocean model component of the CPL system that was run on a grid with 2.2km spacing. We will make this clearer in the manuscript. Table 1 provides a summary of component details, but this could be further clarified in the abstract and the model setup description. Longer-term assessment of the ocean part of the CPL system was assessed by a co-author. We mention that the thermal profile of the ocean was found to be incorrect (too cold in the upper ocean), and a negative bias was found in the net surface heat input into the ocean (Valdivieso et al., 2021) and that this would impact the wake induced cooling. We will add more details to our manuscript from the Valdivieso report given that this report is not peer-reviewed literature.

(4) There is no Data Assimilation used in either of the configurations. Whereas time and again it has been proven that for improving regional forecasts of TCs, DA holds the key. Also, the behavior of the cycled system may be very different from the current system because whatever is lost in terms of the size of the domain could at least be nudged with better initialization from one cycle to another. Please note that DA is one major area that the NCMWF is focused on. The lack of DA in this configuration is surprising. Any regional models like the WRF come along with a DA system.

As highlighted in opening comments to response above, the intention of this paper is not to present results from an operational-like implementation of a new km-scale coupled system. We therefore took a pragmatic, but often used, step to initialise each TC case study from downscaling the relevant global atmosphere MetUM analysis. In that sense, the simulations are initialised from global-scale DA. The absence of a regional-scale analysis (as used operationally within NCUM-R for example), implies that there will be a period of spin-up as small-scale atmosphere features develop within the model domain. Our analysis of TC evolution over ~13 days for each case is therefore not dependent on that initial period. We also note that development of (weakly) coupled DA remains largely the focus for global scale coupled implementation and not a focus for this study, and it is preferable for the context of this study to initialise coupled and uncoupled simulations from the same initial state, and force with the same lateral boundaries. Note that the system is also kept in-check ('nudged') through use of daily updated lateral atmosphere boundaries (i.e. the global forcing is reinitialised each daily cycle based on an updated global model analysis).

The regional ocean model component is initialised in a different way, given the long spin-up time for small scale ocean features to develop relative to the atmosphere. The ocean in each case study was therefore initialised from a free-running simulation which was itself initialised on 01 January 2016, using temperature and salinity profiles from the Global_Analysis_Forecast_PHY_001_024 CMEMS product. The regional ocean model is forced at the boundaries daily by the same CMEMS product (and so is nudged to an assimilative ocean product at the lateral boundaries). This is a common experimental approach to assess case study simulations and we highlight that these simulations do not aim to emulate operational forecasts. The simulations were configured in this way to support the evaluation of modelled TC track, structure, and intensity at different lead times. This point is stated in the manuscript, and we will clarify this further in a revised version.

The suggestion to investigate the extent to which results would have been any different had we initialised atmosphere simulations from NCUM-R analysis is of interest given ongoing collaborations with NCMRWF on the development of the system, but outside the scope of the current manuscript. We envisage that results and our conclusions on the impact of coupling would be unchanged (given that both experiments would continue to be initialised from the same atmospheric state), but that we would remove the initial 6-12h spin up from global-scale features. Similarly, further extension to use a regional ocean analysis is of interest for future work with Indian collaborators and would be required to develop operational-ready capability, but beyond the scope or focus of this work.

(5) Although, the rainfall evaluation presented here is fairly unique, what else is new in this work? The model lacks DA and produces huge track and intensity errors beyond 1-2 days on several important cases, the coupled configuration produces excessive cooling and the model struggled with all the rapid intensifying cases. In that case, the authors have to think if the regional configuration described in this work is really up to the mark to produce forecast quality outputs? I think this work does not show enough evidence that the regional model is a good forecast model nor I am able to see much scientific value beyond the analysis and rainfall verification.

The manuscript was not intended to showcase a new operational regional model that is ready for operational use to "produce forecast quality outputs", but rather to understand the impacts of ocean coupling in a large-domain convective-scale model and the extent to which a convection-permitting coupled model could represent TC dynamics for TCs in the Bay of Bengal. This is a region where there is a lack of previous studies using convection-permitting models. It was also not our intention to suggest the regional model is a good forecast model for this purpose. It was not clear to us at the beginning of the project what the model's limitations would be and how this might vary at different lead times, for TCs in different seasons, of different intensity and crossing the bay in different directions, and between the coupled and atmosphere-only configurations.

We summarise the manuscript's novel aspects as:

- The first study to use a regional convection-permitting model for simulating TCs in the Bay of Bengal. To our knowledge, there is no published literature which uses convection-permitting models with similarly high resolution to simulate TCs in the Bay of Bengal, with much of the previous TC literature focusing on areas which are richer in available observations (e.g. the USA).
- The Bay of Bengal has unique atmospheric and ocean dynamics; for example, pre- and post-monsoon ocean conditions differ significantly (Krishna et al., 1993; Kumar et al., 2019; Neetu et al., 2012) and general circulation models have difficulty simulating the Bay of Bengal's salinity stratification (Chowdary et al., 2016) (see Discussion, lines 612-625). These unique conditions mean that a larger body of literature specifically focusing on the Bay of Bengal is necessary for guiding the development of specialised regional models for this region.
- Rainfall evaluation (as highlighted by the reviewer); the accurate representation of TC dynamics in the convection-permitting model is an important result and encouraging for future coupled model developments.
- Large-domain convective-scale simulations with and without ocean coupling for 6 different case studies at many different initialisation times; many previous studies use a single initialisation time, but here we prove the necessity of considering multiple initialisation times, given the sensitivity of the forecasts to initialisation time.
- Investigation of the lead time limitations of the model simulations (considering the domain size)
- Comparison of storm cases with different intensities and with different orientations of tracks along the coast and across India

Significant results from this study include:

- Both regional models produce similar track errors comparable to track errors in other studies when comparing similar lead times.
- The coupled configuration produces weaker TCs than the atmosphere-only configuration, consistent with lower SSTs in the coupled model and an overestimated cooling response in TC wakes.
- The convection-permitting system accurately captures TC dynamics, shown by its accurate representation of the relationship between rain rate asymmetry and wind shear.
- The wind-MSLP bias is consistent with other environmental prediction systems including the MetUM global model, an important observation for guiding future model developments.
- The use of time-lagged ensemble simulations highlighted the sensitivity of TC simulations to initialisation time; an important observation given that many TC studies in the literature use only a single initialisation time.

We will address the reviewer's point by highlighting the novel aspects of the manuscript better in the abstract, introduction and conclusions.

Considering both reviews, we intend to make major revisions to the manuscript to ensure it more clearly describes the aims of the study, in terms of understanding the impact of coupling, and make stronger statements on what the manuscript is not aiming to do (specifically around show-casing a 'forecast-ready' simulation capability). We note the positive review of Reviewer 1, recommending minor revisions would suffice, and believe many of the concerns of Reviewer 2 would be addressed through a revised framing of the aims of the manuscript, along with some additional analysis as discussed above (e.g., a description of the ocean model analysis carried out by colleagues, and a description of the track errors in the MetUM global model for comparison). Having made these changes and given the novelty of the manuscript as outlined above, we think it will merit publication in WCD; we thank the editor and reviewers for their time and consideration of this manuscript.

References:

Chowdary, J., Srinivas, G., Fousiya, T. S., Parekh, A., Gnanaseelan, C., Seo, H., & MacKinnon, J. (2016). Representation of Bay of Bengal Upper-Ocean Salinity in General Circulation Models. Oceanography, 29(2), 38–49. https://doi.org/10.5670/oceanog.2016.37

Jayakumar, A, Abel, SJ, Turner, AG, et al.: Performance of the NCMRWF convection□permitting model during contrasting monsoon phases of the 2016 INCOMPASS field campaign. Q J R Meteorol Soc., 1– 21. doi:10.1002/qj.3689, 2019.

Jayakumar, A., Sethunadh, J., Rakhi, R., Arulalan, T., Mohandas, S., Iyengar, G. R., and Rajagopal, E. N.: Behaviour of predicted convective clouds and precipitation in the high□resolution Unified Model over the Indian summer monsoon region, Earth and Space Science, 4, 303– 313, doi:10.1002/2016EA000242, 2017..

Krishna, V. V. G., Murty, V. S. N., Sarma, M. S. S., & Sastry, J. S. (1993). Thermal response of upper layers of Bay of Bengal to forcing of a severe cyclonic storm: A case study. IJMS, 22(1). http://nopr.niscair.res.in/handle/123456789/37744

Kumar, B. P., D'Asaro, E., Suresh Kumar, N., & Ravichandran, M. (2019). Widespread cooling of the Bay of Bengal by tropical storm Roanu. Deep Sea Research Part II: Topical Studies in Oceanography, 168, 104652. https://doi.org/10.1016/j.dsr2.2019.104652

Mamgain, A., Rajagopal, E. N., Mitra, A. K., and Webster, S.: Short-Range Prediction of Monsoon Precipitation by NCMRWF Regional Unified Model with Explicit Convection, Pure Appl. Geophys., 175, 1197-1218, doi:10.1007/s00024-017-1754-0, 2018.

Nadimpalli, R., Osuri, K. K., Mohanty, U. C., Das, A. K., Kumar, A., Sil, S., & Niyogi, D. (2020). Forecasting tropical cyclones in the Bay of Bengal using quasi-operational WRF and HWRF modeling systems: An assessment study. Meteorology and Atmospheric Physics, 132(1), 1–17. https://doi.org/10.1007/s00703-019-00669-6

Neetu, S., Lengaigne, M., Vincent, E. M., Vialard, J., Madec, G., Samson, G., Kumar, M. R. R., & Durand, F. (2012). Influence of upper-ocean stratification on tropical cyclone-induced surface cooling in the Bay of Bengal. Journal of Geophysical Research: Oceans, 117(C12). https://doi.org/10.1029/2012JC008433

Singh, V., Konduru, R. T., Srivastava, A. K., Momin, I. M., Kumar, S., Singh, A. K., Bisht, D. S., Tiwari, S., & Sinha, A. K. (2021). Predicting the rapid intensification and dynamics of pre-monsoon extremely severe cyclonic storm 'Fani' (2019) over the Bay of Bengal in a 12-km global model. Atmospheric Research, 247, 105222. https://doi.org/10.1016/j.atmosres.2020.105222

Valdivieso, M., N. Klingaman and C. Holloway (2021): Added Value of Atmospheric and Oceanic Resolutions for Predictions of Air-Sea Interactions in the Bay of Bengal. WCSSP CASPER Deliverable Report 2.2, 19 pp, March 2021.