Comment on gmd-2022-159

Anonymous Referee #2

Referee comment on "Monthly-scale extended predictions using the atmospheric model coupled with a slab ocean" by Zhenming Wang et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2022-159-RC2, 2022

General Comments

In this very interesting study, the author has compare the performance of extended-range forecasts in WRF-ROMS against a similar system but replaced the 3D ocean with a slab ocean model. Comparing SST and TC forecasts between these two systems suggests that a coupled forecasting system with a rather simple slab ocean model can do a decent, or sometimes even better job than that with a full 3D dynamic ocean. These results are scientifically interesting. However, the presentation of this work needs some improvements Therefore, I suggest acceptance after major revision. Please see detailed comments below.

In this study, the authors argue that coupled (extended-range) forecasts with a slab ocean model performed better than a similar system but using a 3D dynamical ocean model. This argument may be true for the setup used by this study, but need to be proven for other system, and shouldn't be presented as a general statement for all coupled forecasting systems. There is no double that using a slab ocean has some benefits w.r.t 3D dynamic ocean modelling (e.g. reduced computing cost, or sometimes, even reduce SST bias if the 3D ocean model is not well configured). However, it is generally believed that you do need a 3D ocean model and a fully coupled system to achieve better forecasts. A good example is the tropical cyclone forecasts studies carried out by Mogensen et al (2018). TCs interact with the SST in three ways: the heat transport to the atmosphere e.g., the vertical mixing with deeper water e.g., and the upwelling generated by Ekman pumping e.g.. While the first (resp. second) process could be easily represented by using a slab ocean model (with the help from a 1D mixed-layer model), a three-dimensional model is nonetheless required to properly represent all three processes.

Therefore, I suggest to play down this strong statement (that a slab-ocean model is overall superior than a 3D ocean model in the sense of making better extended-range forecasts) in this manuscript. Instead, this work can be presented as a study to prove that even a simplified coupled forecasting system with a slab ocean model can work effectively,

or in some cases, better than a fully coupled forecasting system using a 3D ocean model.

Mogensen, K.S.; Magnusson, L.; Bidlot, J.R. Tropical cyclone sensitivity to ocean coupling in the ECMWF coupled model. J. Geophys. Res.-Ocean. 2017, 122, 4392–4412.

Other general comments

- Design of experiments: it is suggested to extend your experiment period (6 months is a bit short) and switch to a slightly different forecasting interval (e.g. 1-wk instead of everyday) when launching forecasts. I understand that it requires significantly amount of work so I leave it to the authors to decide whether or not to carry out this work. But it is worth to point out that results in this study are subject to a relative short testing period that only occupies the second half of the year (no boreal spring or first half of summer season).
- Methodology: Section 2.3 need more details, particularly about your DA method (WCDA) used in this study and your experiment setup. E.g. do you use the same WCDA for both WRF-ROMS and WRF-SOM systems? How do you ensure that initial and boundary conditions in both systems are the same?
- Verification: If possible, one should always try to verify forecasts against observation. In the sense of SST forecasts, there are plenty of observation based SST product available, e.g. L4 products like GHRSST, OIv2 SST, ESA CCI SST. Verification against these products should be encourage instead of against HYCOM reanalysis. If the authors chose to use other reanalysis product (e.g. HYCOM) as verification data set, then at least the ration behind this choice should be added.

Specific Comments

L41: In the extended period prediction, SST is the most important …

This is rather strong statement. SST is no double an important part of air-sea interaction, but other element, e.g. sea-ice condition, is also crucial for accurate extended-range forecast.

L52-56: When discussion potential issues related with coupling to 3D ocean model, it is suggested to point out that any issue (as suggested in Wu et al., 1997 and Ren and Qian 2010) can be specifically related with the ocean model used in their studies. These issues are normally specific to their system configurations – e.g. model resolution, boundary

conditions, model parameterisation, numerical stepping, etc), and can be improved by, e.g. using a better bathymetry input if the model bias is directly related with a sub-optimal bathymetry file.

L56-58: Again I believe that it is not a general issue in dynamic 3D ocean modelling, but rather an issue specific to the set up in that study carried out by Hu et al., 2017.

L61-63: I suggest to restrain from saying that SST is the "**most**" important factor provided by the ocean.

L91-94: Please rephase this sentence to clarify what you mean.

L100: SOM was referring to Slab Ocean Model, here was referred to as "a simple model"

Eq1: how do you compute Q_ocn in SOM? Does SOM treat subsurface as a bottom boundary condition from CFSv2 reanalysis ?

L130: typos with two "from"

L136: Do you mean you use CFSv2 forecasts as boundary condition for WRF-ROMS spun-up? Why not use CFSv2 reanalysis?

L138-140: Could you elaborate a bit on the WCDA approach taken in your experiment? What's obs are assimilated and how do you set up your WCDA framework? This is a rather important component for any forecasting system.

L158: .. is evaluated by the root mean square error (RMSE) "and" the anomaly correlation coefficient (ACC), …

Eq5 and 6, f_ij can be obs or analysis data, but not the truth value (which we never know).

Section 3.1 first paragraph can be moved to Introduction

L193-194: in Figure 3a and 3b, cold bias in WRF-ROMS in the green box can be an issue related with, e.g. inappropriate boundary conditions and/or lack of sea-ice model in your system, particularly during the winter season in the north hemisphere at this latitude. Reason that SOM is doing better may simply because it lacks 3D ocean advection to propagate this cold bias from boundary to other regions. Or, may be the upper ocean mixing is over-estimated in the WRF-ROMS experiment, leading to systematically cold bias almost everywhere. These possibilities should be checked/explained.

L205: suggest to replace predictability by "mean biases"

L206-2015: I understood that definition of mixed layer depth is quite different between WRF-SOM (wind stress and surface heat determined) WRF-ROMS (0.2C from SST) and HYCOM reanalysis (not sure, need add this information). This makes it difficult to interpolate results at Fig. 5, as various MLD definitions can diff as much as 50-100m.

L210: What do you mean by reanalysis data from ECMWF?

L211: Please add information to define the so called "subsurface" in Fig 5-e (spatially averaged temperature below mixed layer depth in WRF-ROMS?). SST in WRF-ROMS is systematically colder than WRF-SOM is a results directly related with their MLD differences. Atmospheric heat fluxes warm the surface ocean in winter season (the experiment period), a relatively shallower MLD in WRF-SOM means more heat resides in the mixed layer, leading to higher SST. Lack of vertical convection and Ekman pumping effect in the slab-ocean model can enhance this surface heat residence effect even more in WRF-SOM forecasts.

L213-215: … the data assimilation can accelerate the heat loss and intensify the cooling in this area… Could you please elaborate a bit here? In general, data assimilation adds constrain to the ocean state variables (temperature/salinity, e.g.) towards observations. Performance of analysis (which is used to initiate your forecasts as I understood) is subject to your model biases, boundary conditions, efficiency and effectiveness of DA method, as well as the quality and quantity of your input observation streams. It is unusual to claim that DA has degraded your analysis performance w.r.t your free run (without DA), which basically means that your DA system is not working as intended.

The statement "by eliminating the influence of initial conditions and ocean heat transport …" sounds like you are suggesting that a forecasting system initialized from a free run (without DA) works better than that initialized from a analysis (with DA), which is not consistent with the modern NWP system.