Response to Interactive comment on "The improvements to the regional South China Sea Operational Oceanography Forecasting System" by Xueming Zhu et al.

Reply on Anonymous Referee #2

Zhu, et al, present a look at changes to an existing operational South China Sea regional forecasting system, SCSOFSv1. They examine a number of changes to the system including: i) revision of the model grid and discretization; ii) change to the atmospheric forcing from direct forcing to utilizing the COARE bulk algorithm; iii) changing the advection scheme; and, finally, iv) making improvements to the ensemble optimal interpolation assimilation scheme.

The results show various improvements for the changes that have been made to the operational system. The manuscript is more of a technical report than a research article. While the results are not general, or improve our understanding of the science of predictability in the region (dynamical analyses, etc.), there may be some useful information that would be valid to a specialized reader. So, in the interests of providing information, it is a paper that when in final form could be published in Ocean Science.

The authors thank the reviewer's confirmation to our manuscript. We completely agree with the reviewer that this manuscript is more of a technical report than a research article. But we think that most of these changes to the SCSOFS are scientific based technology and can be supported by basic physical oceanography theories or numerical simulation technic. It would provide useful information and references to numerical modelers, the communities of ROMS or operational oceanography how to set their numerical models in later.

That said, the paper is fundamentally lacking in two primary ways. First, to be of any use to that select reader, it must provide details and explanation for the changes and how they would dynamically improve the representation of the region. Secondly, the paper lacks any rigorous analysis of the "improvements." Simply stating that one thing was changed and the summertime surface temperatures are cooler is not particularly useful.

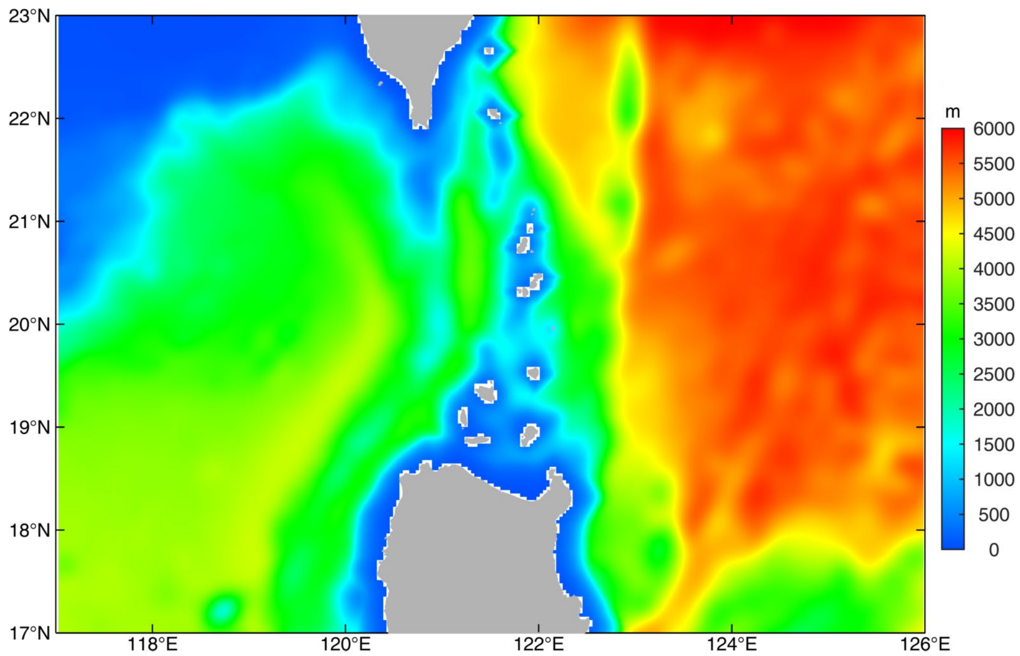
Thanks for pointing out our manuscript's deficiencies. We have revised this manuscript by adding some significant and substantial information and details which was not given in the original manuscript, such as the difference between Akima scheme and fourth-order centered scheme, the configurations and many details of EnOI scheme. We also supplement some "rigorous analysis" lacked in the original manuscript, such as why the impact of moving eastern lateral boundary on the SCS, COARE 3.0 bulk algorithm does not work better than direct forcing at all time, the reason for results of Fig. 11d better than that of Fig.11c, and adding a new Figure 12.

First, I will go through the changes discussed by the authors as topics, then paper comments.

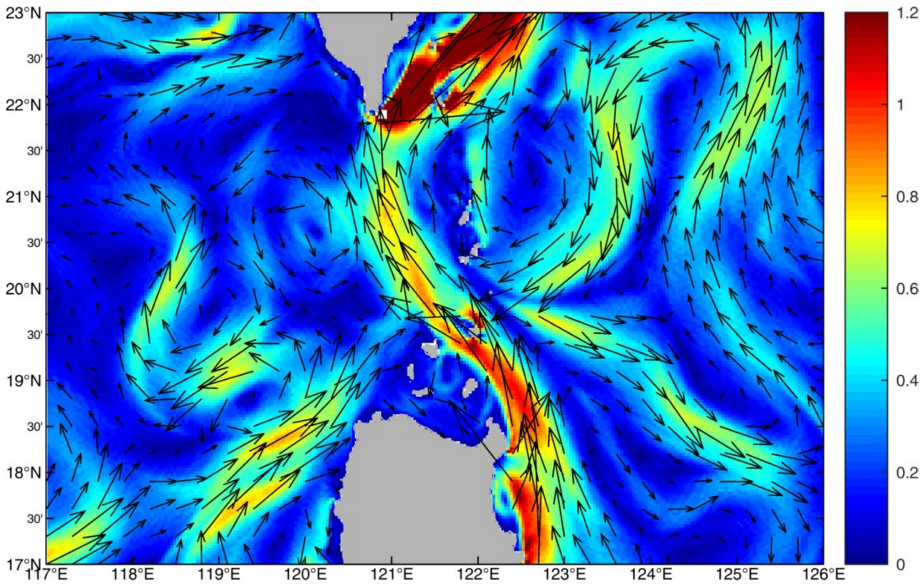
The first change made to the model is changing the model grid representation of the region. The bathymetry is re-generated, smoothed, and the number of layers increase. The authors state that they smooth the bathymetry to reduce the bathymetric slope to an r-factor under 0.2. However, for an s-level model with various stretching function, the so-called "Haney" *number is more valid and it increases with increasing the number of layers. There is no analysis as to whether this grid has improved spurious pressure-gradient flows or can better resolve the complicated bathymetric derived flows of the region. Figure 1 seems to show that the Luzon strait may be far too deep, which would alter the early Kuroshio flow with a more predominant intrusion through the strait.*

We changed the model grid representation of the region to make the model fit the actual coastline better, since it is an operational forecasting system and will be used to provide real forecasting. We derived and re-generated bathymetry from GEBCO, which has higher horizontal resolution (30 arc-second) than ETOPO1 (1 minute) and should be more reasonable and closed to the real bathymetry according to the introduction in its website. We smooth the bathymetry to keep the slope parameter (r) under 0.2 by referencing Marchesiello et al. (2009) and Beckmann and Haidvogel (1993). ROMS has a new pressure-gradient scheme associated to a modified equation of state limits computational errors of the pressure-gradient, which can improve spurious pressure-gradient flows and better resolve the complicated bathymetric derived flows of the region, especially with the higher vertical resolution. In SCSOFSv2, we have increased the vertical layers from 36 to 50, and employed an improved double stretching function in order to preserve a sufficient resolution in the upper layer. The two grid stiffness ratios parameters, slope parameter (*r*) and Haney number, range from 0.22 and 9.78 in SCSOFSv1 to 0.17 and 13.80 in SCSOFSv2, respectively. We have added some statements in the revised manuscript between lines 132 and 139, and Lines 144-146. According to our experience, if the model run with significant errors of pressure-gradient, it would blow out very quickly, but our system has run more than 40 years stably. It indicates that our system is stable enough.

For the bathymetry of Luzon strait in Figure 1, we add a figure by zooming in the Luzon strait (FigR1). It shows that the maximum depth is about 3000-3500m in Luzon strait, it is normal. We also show a daily averaged surface velocity around Luzon strait on July 15, 2018 (FigR2). It does not show predominant intrusion through the strait.



FigR 1: Bathymetry around Luzon strait



FigR 2: Surface velocity around Luzon strait on July 15, 2018

Modification:

L132-139: revised to "Then it is smoothed by applying a selective filter 8 times to reduce the isolated seamounts on the deep ocean, so that the "slope parameter" $r=\Delta h/2h$ is lower than a maximum value r_0=0.2 for each grid (Beckmann and Haidvogel, 1993; Marchesiello et al., 2009), in order to supress the computational errors of the pressure-gradient (Shchepetkin and McWilliams, 2003). Then the two grid stiffness ratios parameters, slope parameter (r) and Haney number, change from 0.22 and 9.78 in SCSOFSv1 to 0.17 and 13.80 in SCSOFSv2, respectively. The maximum depth is set to be 6000m still, but the minimum depth changed from 10m in SCSOFSv1 to 5m in SCSOFSv2 (Wang, 1996). The final smoothed bathymetry is shown in Fig.1."

L142: "the" revised to "an"

L144: "to make it thinner on the upper 300m" revised to "to make it preserve a sufficient resolution in the upper 300m";

L145: added "In this case, the thinnest layer changes from 0.16m in SCSOFSv1 to 0.09m in SCSOFSv2 near the surface."

Secondly, the authors remove the island of Guam because it disrupted the flow in the v1 of the system. Guam lies as the southern island in the Mariana island arc that extends from the Mariana trench to the surface to impinge on the westward flow. Guam (as part of the archipelago) does disrupt the flow. If v1 of the system had an incorrect disruption, this implies that the boundary conditions used are not necessarily appropriate or account for the archipelago. More analysis and discussion of these items is needed to show that the new grid better represents the dynamics of the region.

Thanks for taking care about this point. We agree with the reviewer that the boundary conditions used are not necessarily appropriate in SCSOFSv1. Just like we have answered to the first reviewer, we want to get a reasonable large scale NEC background circulation in the interior of the SCS. Since we cannot get a real open condition to force the model, we have to use Radiation and Nudging condition to resolve the open condition problem. The reason to set the eastern lateral boundary far away from the SCS, is that we try to make the interior of SCS do not encounter the open boundary problem. Since kinds of boundary conditions, Radiation and nudging conditions and sponge layers, are described in our model. The results of model next to the lateral boundary within 10 grids will be unbelievable. So the solution to remove the island of Guam is the easiest way for us to improve the system. Of course, more analysis and discussion on the Guam is needed, but we do not think it is a good idea to discuss this problem with our SCSOFS model, the island of Guam is too close to the open boundary. A large model domain, perhaps with the whole Pacific Ocean, may be needed to address this question, in order to avoid the effects of open boundary.

Modification:

L195: added "Also, it is shown that the Kuroshio of eastern Philippine Island and ocean circulations of northeastern SCS get stronger while the island of Guam removed."

L189: replaced Figure2 with a new one to enlarge the domain to include SCS.

The second change is the switch from direct forcing (momentum, heat, and salt) to utilizing the COARE bulk fluxes algorithm. Of course, the COARE algorithm has been in wide use for two decades in modeling, so the authors don't have anything new to present on the topic; however, the changes and differences seem to suggest not an improvement by the COARE algorithm, but rather a deficiency in the original direct forcing used by the original system. The COARE algorithm produces direct forcing using the atmospheric conditions and the SSS and SST of the model at the time of the forcing. This time dependency is obviously advantageous; however, to see such strong changes in the means shows that there was/is a forcing issue. The direct forcing should be nearly comparable, just lacking the ability to adjust to local intra-seasonal variance.

Thanks for mentioning this. We agree with the reviewer that the COARE algorithm has an obviously advantage of time dependency in theory by producing direct forcing using the atmospheric conditions and the SSS and SST of the model at the time of the forcing. But it looks like that it is not advantageous at all time in reality. We can find that the results from BulkFormula are worse than that from SCSOFSv1 in January and February in Figure 4. Actually, for this problem, we have published other paper (Li et al., 2019) in Chinese with English abstract. It concludes that 'the accurate dataset of surface forcing field is more effective than changing the choice of forcing mode'. We have also added some discussion about this in the revised manuscript between lines 273 and 277.

In fact, even though there is a deficiency in the direct forcing method, it has been coded in many popular ocean models, such as ROMS, FVCOM, NEMO, POM, because it is easy for coding and running model. And many modelers, especially for the beginners, like to use this method to force their model for convenience. So we would like to remind them to select the sea surface forcing method with caution.

Modification:

L273: added "However, domain averaged RMSE of monthly mean SST differences from SCSOFSv1 is lower than that from BulkFormula in January and February. It indicates that COARE 3.0 bulk algorithm does not work better than direct forcing at all time, even with effective negative feedback mechanism. This may be surface forcing field data dependent, and the accurate dataset of sea surface atmospheric forcing is more effective than the forcing methodology selection (Li et al., 2019)."

The third change was to switch advection schemes from a third-order, upstream in the horizontal and a central difference in the vertical to using the Akima scheme for both *horizontal and vertical advection. The advection schemes, by design, represent the wavenumber spectrum differently, and in reading this manuscript, it is unclear why one would be preferred over another. Some cursory comparisons are provided, but how is this a result that would be useful to the community? There needs to be some explanation and analysis on how the flows are represented by each scheme. In the third-order scheme, it is far less sensitive to the prescription of explicit horizontal viscosity. What about in the Akima scheme? Did you vary your viscosity for momentum and diffusion for tracers? Is the sub-grid scheme chosen (authors did not mention which scheme, e.g., Mellor-Yamada 2.5, LMD, etc.) sensitive to the advection scheme?*

Thanks again for pointing out this. We have revised the manuscript to show the differences between the Akima and four-order centered schemes in the calculation of curvature term. Then the Akima scheme represents its advantage of reducing spurious oscillations, which arised with nonsmooth advected fields, since the time stepping is done independently from spatial discretization in ROMS (Shchepetkin and McWilliams, 2003, 2005). We have not changed the explicit horizontal viscosity for momentum and diffusion for tracers schemes in SCSOFSv2, so we did not mention it in the original manuscript. We have added the description of the horizontal viscosity and sub-grid scheme chosen in revised manuscript between lines 197 and 201. Right now, we have not tested different diffusion schemes for tracers yet, and will test it in other paper about model settings effect on the spurious diapycnal mixing. Thanks for your suggestion. We want to propose AAG schemes combination in this paper to the ROMS community, especially for the beginners. Although the UCI schemes combination is the default and popular settings, instead of the best. Many ROMS' users did not like to change some basic settings of model except for the experts as reviewers, thus they would like to ask some help for data assimilation.

Modification:

L197-201: added "For the advection schemes of momentum, third-order upstream and fourth-order centered schemes are used in both horizontal and vertical, respectively. Harmonic mixing scheme is used for both viscosity for momentum and diffusion for tracers in horizontal. Mellor-Yamada Level-2.5 vertical turbulent mixing closure scheme is used for both momentum and tracers. All of them in SCSOFSv2 are set to be same as in SCSOFSv1."

L297: added "(Shchepetkin and McWilliams, 2005)"

L309-313: added "The fourth-order Akima scheme is a little different from the fourth-order centered scheme by replacing the simple mid-point average with harmonic averaging in the calculation of curvature term. Since the time stepping is done independently from spatial discretization in ROMS, the Akima scheme represents its advantage of reducing spurious oscillations, which arises with nonsmoothed advected fields, with respect to the fourth-order centered (Shchepetkin and McWilliams, 2003, 2005)."

The final change was to modify the ensemble optimal interpolation (EnOI). The authors reference an earlier paper, Zhu, et al., 2016, that does not contain details about their EnOI implementation. Oke, 2002, does a good job of explaining the EnOI methodology as a simplification of the ensemble Kalman Filter; however, it has a number of choices that must be made in its implementation. The authors of this paper do not describe any of their choices, their impacts, or reasoning. First, the choice of ensemble members. It is not well described in the manuscript. I will explain my understanding; however, I am unsure how correct this is due *to the lack of any detail. The authors break a 7 day assimilation window into 3 hourly periods. All observations are assigned to their nearest 3 hour slot, providing a total of 57 ensemble members. For each member, the innovations are calculated, and used for the EnOI setup. There is no discussion of localization, which is absolutely required in the ensemble scheme increments.*

We are sorry for the poor English and making the reviewer misunderstand our meaning. For the ensemble members, we select about 590 members from the 10-year (2008-2017) free run with daily output in total, which include 59 members from before and after 30 days of target assimilating date in each year. We divide 7 day assimilation window into 3 hourly periods for calculating the innovations more accurately, not for the ensemble members. This is the main objective of introducing FGAT method. We can get the same number of innovations with the assimilated number of observations, instead of 57 ensemble members by employing FGAT. We do have used localization with the radius 150 km. We have revised the manuscript in between lines 375 and 410.

Modification:

L375-410: revised to "Secondly, we have introduced the method of computing the anomalies of ensemble numbers used for constructing the background error covariance following Lellouche et al. (2013). In SCSOFSv1, the anomalies are computed by subtracting a 10-year average from a long-term (typically 10 years) model free run snapshots with 5-day interval for the ocean state, i.e. sea surface height and three-dimensional temperature, salinity, zonal velocity, and meridional velocity. And the ensemble is selected within a 60 d window around the target assimilation date from each year, adding up to about 130 members in total (Ji et al., 2015; Zhu et al., 2016). However, in SCSOFSv2, a Hanning low-pass filter is employed to create running mean according to Lellouche et al. (2013) in order to get intra-seasonal variability in the ocean state. Thus the anomalies are computed by subtracting the running mean with 20-day time window from a 10-year (2008-2017) free run daily averaged results. Especially, it is pointed out that the daily averaged free run results are selected within 60 d window around the target assimilation date from each year of 2008-2017 and used to compose ensemble members, thus about 590 members totally in SCSOFSv2. It means that the background error covariances rely on a fixed basis and intra-seasonally variable ensemble of anomalies, which improves the dynamic dependency.

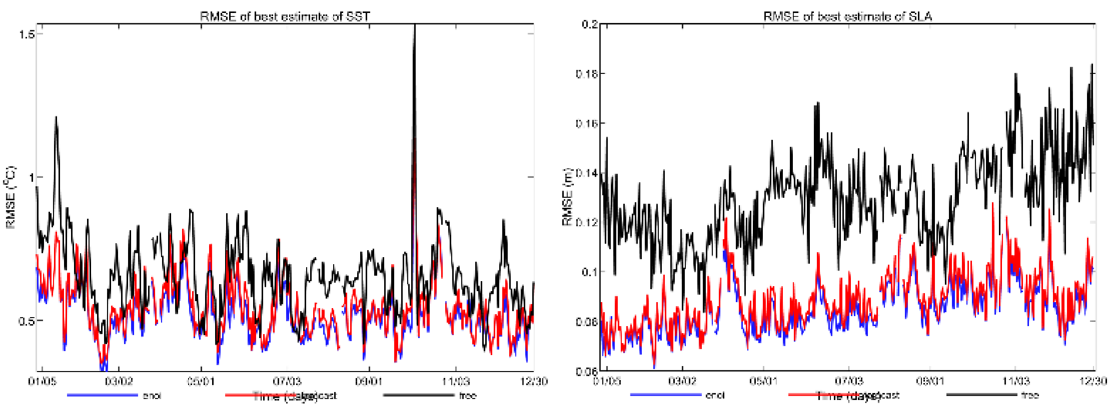
Thirdly, for each analysis step with a 7-day assimilation cycle, all observations of SLA within the 7-day time window before the analysis time are treated as observed at the analysis time in SCSOFSv1, with assumption of all observations were still valid at the analysis time. The time misfit between the observation and model forecast would cause non-negligible biases for the calculation of innovations. Actually, it is inconvenient to calculate the synchronous innovations between the observation and model forecast entirely, since the spatio and temporal distributions of along-track SLA and Argo data are irregular and variable at each analysis step. In order to alleviate this deficiency, the First Guess at Appropriate Time (FGAT) method (Lee and Barker, 2005;Cummings, 2005;Lee et al., 22-25 June 2004;Sandery, 2018) is used in SCSOFSv2. Considering the intense computing and storage cost, we have divided the 7-day time window into 56 3-hour time slots (Fig.8), and archived 57 snapshots with a 3-hour interval while model forecast run following the previous analysis run. Then the innovations can be calculated within each 3-hour time slot by using the observations subtracts the nearest model forecast. It means that the maximum temporal misfit of the innovations between the observation and model forecast would be decreased from 7 days to 1.5 hours by using FGAT. Meanwhile, the localization is still used with the radius set to be 150 km as in SCSOFSv1."

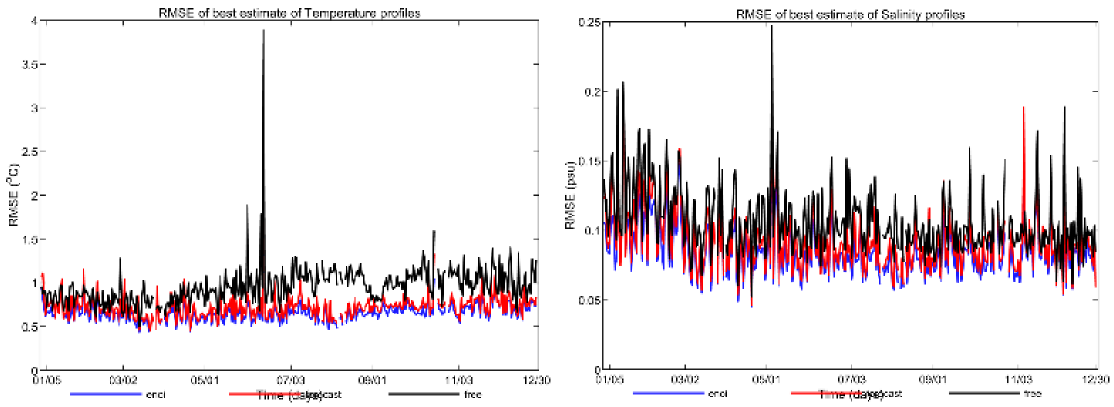
The authors take the increments and turn them into temporal nudging by dividing the EnOI derived increments by 57 and applying them weighted through time. The way it is presented would not seem to make sense. There are 57 innovations relative to the background. However, once an increment is made, the subsequent increments are now invalid because they are now being applied to a field with different innovations. This is why the original EnOI scheme applied the increments at the beginning of the window. Furthermore, localization is required to make this somewhat reasonable.

We are sorry again for the poor English and making the reviewer misunderstand our meaning. From the MOOAS, we just can get one increment for each ocean states variable, and the number of innovations is large and equal to the number of assimilated observations instead of number of ensemble members of 57. The number of innovations is variate in each data assimilation cycle, since the number of along-track SLA and Argo T/S profiles is variate. We do have used localization. We have revised it in the new manuscript between lines 375 and 410.

The final issue with this is that for a forecasting system, the authors never examine the forecasts. Adding forcing through time should reduce your forecasting skill because as soon as the forecast begins, the artificial forcing term disappears and the system is "shocked" back to its original state with the forcing that was "holding" it in place removed. How did these improvements change the forecast ability of the system? Does the EnOI system provide significant improvement over a model with the "best" atmospheric forcing?

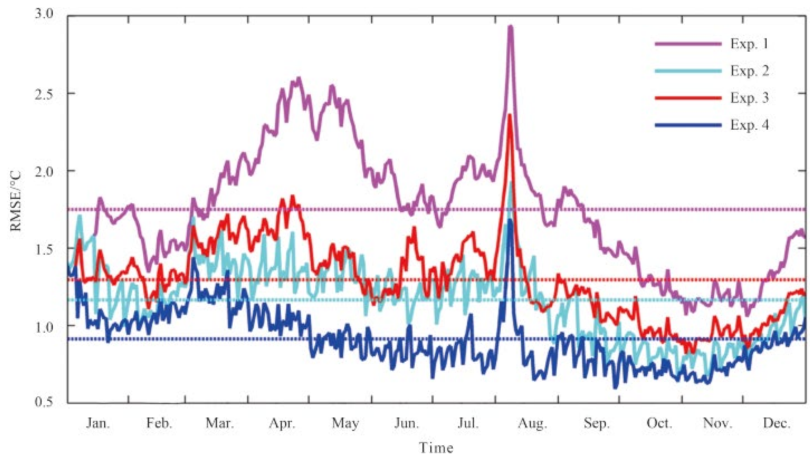
Thanks for mentioning about this. Yes, we have not examined the forecasts. It is because that the new system has not been operated formally for some reasons. So all the results examined in this paper are hindcast results. But in our opinion, we have changed the basic physical model settings, all the improvements mentioned in this paper could work for the forecast, if they work for the hindcast. For the artificial forcing term from MOOAS, it just works in "analysis stage" not in "forecast stage", even in real operational forecasting. It is used to get more reasonable model initial fields for the 7-day forecasting. Here we provide one more figure (FigR 3) to show the results comparison among the free run (black), forecast stage (red), and analysis stage (blue) in 2018. We can find that those RMSEs from forecast stage are a little larger than those from analysis stage, but smaller than those from free run.





FigR 3: Comparison of the whole domain averaged RMSE timeseries for SST, SLA, Argo T&S in 2018, among free run (black), forecast stage (red), and analysis stage (blue)

Actually, we have not tested whether the EnOI system provide significant improvement over a model with the "best" atmospheric forcing. For those major improvements mentioned in this paper, we all implement them with CFSR atmospheric forcing. We think that CFSR data is the "best" atmospheric forcing in our system for now. We plan to rerun the model with ERA5 atmospheric forcing in next step. But it is hard to say which one is the "best", such as CFSR, ERA5, NAR and so on. Here we would like to refer the Fig. 7 from Ji et al. (2015) as follow FigR4,



FigR4: The daily mean (solid lines) and yearly mean (dash lines) time series of domain averaged SST RMSE of the four experiments compared with the OSTIA SST in 2011. The purple, cyan, red and blue are for Exp. 1, Exp. 2, Exp. 3 and Exp. 4, respectively.

Exp.1 is the control run with realistic initial conditions and surface atmospheric forcing, but without data assimilation. Exp.2 is run with OSTIA SST assimilated by employing EnOI only. Exp.3 is run with OSTIA SST assimilated by correcting the surface atmospheric forcing (Qcorrection method). Exp.4 is run with OSTIA SST assimilated by employing both EnOI and Qcorrection method. It shows that EnOI in Exp.2 reduces RMSE more effectively than Qcorrection in Exp.3.

* Minor Comments

1. The authors are non-native English speakers, so I give some leeway, but the manuscript is difficult to read and confusing in many places that are too many to list here. Example lines: 68, 120, 123, 189, paragraph at 195, sentence at 210, 229, most of the EnOI description, 426,467, *473, 540, 545)*

The manuscript has been revised and proofread thoroughly according to the referees' comments by the co-author Ms. Miaoyin Zhang again, who is a professional of same academic background with proficient English written skills, wrote her master thesis in English and received her master degree in England.

Modification:

L69: "or improving" revised to "changing"; "in order to assimilate" revised to ", assimilating"; "according to the benefits" revised to "benefiting".

L122: "important" revised to "significant"; "and transport" revised to "transportation"; "then" revised to "thus"; "simulating more accurate" revised to "better simulation of".

L125: "The bathymetry of ETOPO1 data set, with 1 arc-minute grid resolution from U.S. National Geophysical Data Center (NGDC) is substituted by the General Bathymetric Chart of the Oceans (GEBCO_08, Grid version 20091120, http://www.gebco.net) global continuous terrain model for ocean and land with 30 arc-second spatial resolution in SCSOFSv2." revised to "The bathymetry is replaced by the General Bathymetric Chart of the Oceans (GEBCO_2014 Grid) global continuous terrain model for ocean and land, which is with 30 arc-second spatial resolution in SCSOFSv2, from ETOPO1 data set in SCSOFSv1, which is with 1 arc-minute grid resolution from U.S. National Geophysical Data Center (NGDC)."

L209: "Most of bias or errors in the operational systems mainly induced by initial errors and model deficiencies, which can be attributed to some major recurring problems like sea surface atmospheric forcing, intrinsic deficiencies of numerical model (e.g. discrete schemes, parameterization schemes for sub-grid scale), and the assimilation schemes. In this section, we elaborate solutions, that are not mentioned in Sect.2, to such problems applying in SCSOFSv2." revised to "Most of bias or errors in the operational systems are mainly induced by some major recurring problems, for example sea surface atmospheric forcing, intrinsic deficiencies of numerical model (e.g., discrete schemes, parameterization schemes for sub-grid scale), initial errors, and the assimilation schemes. In this section, we elaborate solutions, that are not mentioned in Sect.2, to such problems applied in SCSOFSv2."

L215: "The air-sea interaction is one of essential physical processes that affect vertical mixing and thermal structure of the upper-ocean. The air-sea fluxes mainly include momentum flux, fresh water flux and heat flux. SST is an important indicator of ocean circulation, ocean front, upwelling and sea water mixing, whose variation mainly depending on the air-sea interaction and the ocean thermal and dynamical factors (Bao et al., 2002). Thus, for OOFS and ocean numerical modelling, accurate simulation and forecasting of SST is one of the most important metric to evaluate the modelling and forecasting skill." revised to "The air-sea interaction is one of the most essential physical processes that affect vertical mixing and thermal structure of the upper-ocean. The air-sea fluxes mainly include momentum flux, fresh water flux and heat flux. SST is an important indicator of ocean circulation, ocean front, upwelling and sea water mixing, whose variation mainly depending on the air-sea interaction, the ocean thermal and dynamical factors (Bao et al., 2002). Thus, for OOFS and ocean numerical modelling, simulation and forecast accuracy of SST is one important metric to evaluate the modelling and forecasting performance."

L230: "Since the SST using in the calculation of those three air-sea fluxes is extracted from ocean model, the increasing of SST thus induces the variations of sensible heat flux, latent heat flux, and longwave radiation then increasing loss of ocean heat, and inhibiting the further increasing of SST, and vice versa." revised to "Since the SST using in the calculation of those three air-sea fluxes is extracted from ocean model, the increase of SST induces the variations of sensible heat flux, latent heat flux, and longwave radiation as a result, which then lead to increasing loss of ocean heat, and inhibiting further increase of SST, and vice versa."

L252: "The differences are pronouncedly higher for the results from SCSOFSv1 than the results from BulkFormula and SCSOFSv2." revised to "The difference from SCSOFSv1 is pronouncedly higher than the differences from BulkFormula and SCSOFSv2."

L367-410: most of the EnOI description has been revised.

L471: "For the whole improving process, the accuracy" revised to "The accuracy"

L472: "AC increasing" revised to "AC increased"; "RMSE decreasing" revised to "RMSE is decreasing"

L512: "slightly" revised to "slight".

L513: "may because there is" revised to "maybe because".

L515: "the data sets' " deleted; "temporal resolution" revised to " temporal resolution of the datasets".

L519: "good represents" revised to " well representing".

L594: "literature" revised to "study"; "all the" revised to "major".

L595: " based on" revised to " following"

L597: "horizontal" revised to "spatial"

L600: "sensitive" revised to "significant";

L601: "include" revised to "acquire".

L602: "Upgrade leads" revised to "Upgrades lead".

L603: "disappearing" revised to "eliminating"

2. Figure 1: This map ratio is strange. There are roughly 25 deg in the vertical and 45 deg in the horizontal. The Philippine Sea is presented as smaller than the SCS. Likewise Figure 2 doesn't show the impact of changes on the SCS, the entire point of the paper.

We think reviewer makes a misunderstanding that there is roughly 33 deg, from -4.5 deg to 28.4 deg, but not only 25 deg in the vertical. We have changed the Figure 2 to include the SCS. It is shown that once removing the island of Guam, the Kuroshio of eastern of Philippine Island and the currents interior of SCS get stronger. We have revised the manuscript between lines 195 and 196, and replot the Figure 2.

3. Paragraph near line 140: Why would you use a different set of initial conditions from the boundary conditions you are using? What is the persistence of the initial condition information? Over the 16 years or so that you spinup experiments, wouldn't the boundary conditions replace the initial conditions? There is no explanation for why you would use one product to initialize the model and SODA as your lateral boundaries.

We use a different set of initial conditions from the boundary conditions, is because we want to set more realistic and accurate initial conditions to the model. GDEM has a finer horizontal (0.25 deg) and vertical (78 layers until 6600m) resolutions than those (0.5 deg in horizontal and 40 layers until 5375m in vertical) in SODA 2.2.4. And GDEM mainly comes from observation, but SODA comes from numerical model reanalysis. The persistence of initial condition information should be very long, especially for the deep layer over 3000m depth in the middle of SCS and western Pacific Ocean. We do not think 16 years or so is enough for the boundary conditions to replace the initial conditions, since our model domain is large and the open boundaries far away from the interior. Since there is only monthly climatological temperature and salinity in GDEM, it is not suitable to provide lateral boundary conditions with lack of velocity, water level variables and history timeseries for all variables.

4. Line 233 states that coastal sea surface forcing was "heating up the ocean". The region is massive with a deep ocean basin. How would some slight imbalances along a coast heat up the entire ocean of like 124 km2?

Thanks for pointing out this. We have revised the text in the manuscript to clarify this confusion. We just want to present that the sea surface atmospheric forcing heats up the ocean “near the coast”, but not the entire ocean. Since the anormal higher SST mainly occurs in the coastal shallow area or a few small bays embedded into the mainland which is hard to be resolved well with 2-3 horizontal grids at 1/30° resolution.

Modification:

L257: “ocean” revised to “coastal water”.

5. Line 241: Figure 4 is not a histogram. It is a bar chart.

L264: “histograms” revised to “bars”.

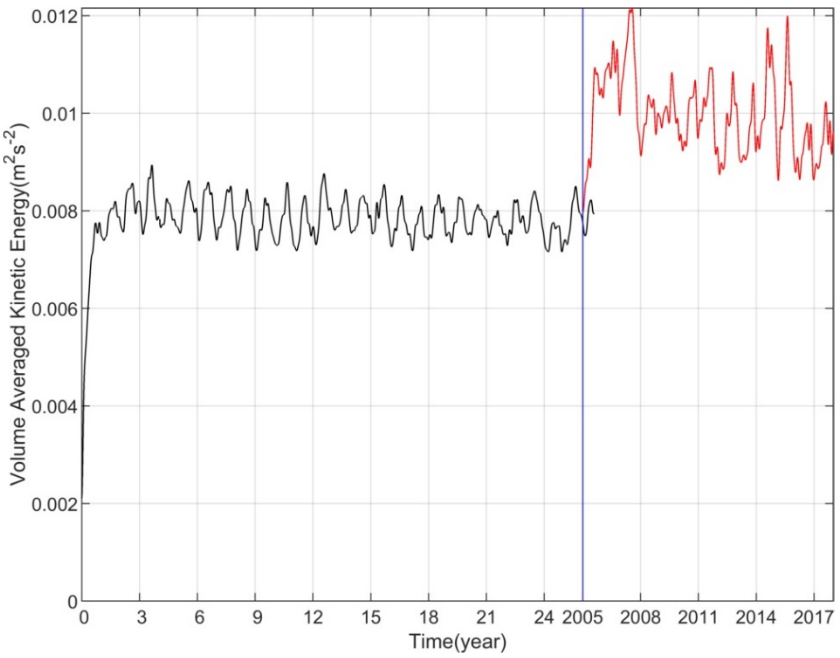
6. Figure 5 caption, should this be SCSOFSv1 in (c)?

L294: “SCSOFSv2” revised to “SCSOFSv1”.

7. Paragraph around 275: You spin up the model until it is "in stable status". What does this mean? How do you determine "stable status"? During the spinup, you say that the temperature increases, but how does SODA compare? As per (3) above, are you not just converging towards the SODA state?

Please see the FigR4 as follow, showing the volume averaged Kinetic Energy (KE) for the whole model domain. We usually determine “stable status” according to the variability of KE. We can find that the model begins to be stable since the sixth model year from FigR4 according to KE. But from Figure 7, it looks like the domain averaged monthly mean temperature and salinity at 1000m layer still increasing in the sixth model year. Until the sixteenth year, the ratio of increasing for both temperature and salinity can be stable, not show significant increase any more.

We do not need to compare the model temperature with SODA during the spin-up, since we only cyclic use monthly climatology SODA open boundary from January to December with 360 days period in this model stage. Temperature and salinity in the open boundary are not continue increasing, and the model is not able to converge towards the SODA state.



FigR4: Volume averaged Kinetic Energy for the whole model domain. Black line is from climatological running, red line is from hindcast running.

8. Line 360, for the EnOI, you generate seasonal background error covariances. Wouldn't it be more appropriate to have typical 7-day background error covariances? The covariance over a season is very large compared to the variability over 7-days. You should expect that your model is within a 7-day covariance period of the observations?

Here we generate seasonal background error covariances is referring to Xie and Zhu (2010), to compose the "running" seasonal ensemble at assimilation time and to represent the flow-dependence in the term of seasonality. What's more, we only consider the variability of each ensemble member within 20-day period, by subtracting the results of Hanning low-pass filter from the original daily averaged model outputs.

9. Section 4, "Scientific inter-comparison" I don't really see any scientific comparison, just comparing the bias, RMSE, and AC.

We have deleted "scientific" in Lines 16, 95, 443, 445, 622.

10. Table 2, what is the final row?

Thanks. It denotes that the results from the third to the sixth column are in 2013, and the results from the eighth to the nineth column are in 2018.

11. Line 460, what is "PI"?

PI means "percentage increase". It has been illustrated at line 472.

12. Figure 11c, there are strong RMSE on either side of the Luzon Strait. The observations capture the surface bounce of the strong internal M2 tides there, resulting in a 25cm "error" that is present in altimetry because the surface bounce of internal tides vary in phase and location, meaning they aren't removed in the sea surface height measurements. So, 11c looks exactly as it should. But, 11d shows significant reduction in something that your model doesn't include. You don't have tidal forcing (at least explained in the manuscript), so you don't have the process present to reduce the error. This means that your EnOI is doing something non-physical to the system by trying to force an SSH expression that is due to internal tides.

Thanks for taking care about this. It is correct that our model does not have tidal forcing right now. The SLA data assimilated in SCSOFSv2 is near real time along-track L3 product for assimilation, which is filtered but not subsampled and with DAC (Dynamic Atmospheric Correction), **ocean tide**, LWE (long wavelength error) correction applied (CMEMS-SL-QUID-008-032-051, http://marine.copernicus.eu/documents/QUID/CMEMS-SL-QUID-008-032-051.pdf). The filtering processing consists in a low-pass filtering with a cut-off wavelength of 65 km using a Lanczos filter. Residual noise and small scale signals are then removed by filtering. The surface bounce of the strong internal M2 tides can be captured by altimetry, since its frequency is very high. We do not think 25cm "error" on either side of the Luzon Strait is resulted from internal M2 tides included in altimetry data. However, we think that should be from mesoscale eddies activities. It is well known that energetic mesoscale eddy activities occur around each side of Luzon Strait. Physical model without data assimilation cannot capture the mesoscale eddies and match the observations exactly, since its intrinsic deficiencies on the numerical method and sub-grid parameterizations. This has been relied on data assimilation. The main objective of assimilating along-track SLA data is to capture mesoscale eddies better. Fig. 11d shows that significant reduction with SLA RMSE, indicating that SCSOFSv2 has represented meso-scale eddies more exactly with data assimilation than v1.3 without data assimilation in Figure 11c. It shows that EnOI plays an important role to improve the ability of representing meso-scale eddies of SCSOFS by assimilating SLA data. We have added content to discuss it in revised manuscript between lines 537 and 542, and added one more Figure 12 with title "Daily averaged SLA (color shaded) and surface velocity anomaly (vector) on January 15, 2018, from AVISO, SCSOFSv1.3, and SCSOFSv2, respectively".

Reference:

Beckmann, A., Haidvogel, D.B.: Numerical simulation of flow around a tall isolated seamount. Part I: problem formulation and model accuracy. J. Phys. Oceanogr. 23, 1737-1753, 1993.

Ji, Q., Zhu, X., Wang, H., Liu, G., Gao, S., Ji, X., and Xu, Q.: Assimilating operational SST and sea ice analysis data into an operational circulation model for the coastal seas of China. Acta Oceanol. Sin., 34, 54-64, 10.1007/s13131-015-0691-y, 2015.

Li Ang，Zhang Miaoyin，Zhu Xueming*, et al. A research on the optimal approach of CFSR surface flux data correction based on different surface forcing modes. Haiyang Xuebao，2019, 41(11)：51–63，doi:10.3969/j.issn.0253−4193.2019.11.006. (In Chinese with English abstract)

Marchesiello, P., Debreu, L., and Couvelard, X.: Spurious diapycnal mixing in terrain-following coordinate models: The problem and a solution, Ocean Model., 26, 156-169, 10.1016/j.ocemod.2008.09.004, 2009.

Shchepetkin, A. F., and McWilliams, J. C.: A method for computing horizontal pressure-gradient force in an oceanic model with a nonaligned vertical coordinate, J. Geophys. Res., 108, 3090, 10.1029/2001JC001047, 2003.

Shchepetkin, A. F., and McWilliams, J. C.: The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model, Ocean Model., 9, 347-404, 10.1016/j.ocemod.2004.08.002, 2005.

Xie, J., Zhu, J.: Ensemble optimal interpolation schemes for assimilating Argo profiles into a hybrid coordinate ocean model, Ocean Modelling, 33(3 – 4): 283 – 298, 10.1016/j.ocemod.2010.03.002.

Zhu, X., Wang, H., Liu, G., Régnier, C., Kuang, X., DakuiWang, Ren, S., Jing, Z., and Drévillon, M.: Comparison and validation of global and regional ocean forecasting systems for the South China Sea, Nat. Hazards Earth Syst. Sci., 16, 1639-1655, 10.5194/nhess-16-1639-2016, 2016.