Dear authors,

After reading the revised manuscript, I have to agree with critical comments of reviewers. Since the manuscript is mainly based on tuning parameters of established inflation methods RTPP and RTPS and their combinations with the IAU, from the scientific point of view, it already lacks innovation to some degree. Then it is hoped that authors could have made thorough discussion of results, unfortunately, it is also missing in the revised manuscript. I strongly recommend authors to improve the manuscript in three aspects as follows:

We thank the editor for reviewing our paper and providing useful comments. The dynamical balance is important for EnKF-based ocean data assimilation systems in which satellite and in-situ observations are assimilated at frequent interval. This study reveals that the combination of the IAU and RTPP only improves both dynamical balance and accuracy, by conducting the sensitivity experiments on various settings for IAU and covariance inflation methods. As indicated by Reviewer 2, the RTPP and IAU now might be standard methods, especially for the atmospheric community. However, to the best of our knowledge, most of EnKF-based ocean data assimilation systems have not adopted the IAU nor RTPP/RTPS (Table 1), and there are no studies to indicate that the combination of the IAU and RTPP is the best for dynamical balance and accuracy and to adopt it to EnKF-based ocean data assimilation systems.

Although this study does not include an object to quantitatively investigate the detailed mechanisms how the IAU and each covariance inflations affect the dynamical balance and accuracy, we reply as much as possible following the editor's comments.

1)Please make the research question more clear. As reviewers pointed out, the best combination could vary, for instance, if the ensemble size or localization change. Why do author think that tuning RTPP and RTPS in combination of the IAU is most worth exploring.

In the third paragraph in Section 1, we have discussed that dynamical imbalance might degrade the accuracy in frequent data assimilation. *To provide accurate analyses in EnKF-based ocean data assimilation system in which satellite and in-situ observations are assimilated at a frequent interval, it is necessary to investigate an optimal setting for both dynamical balance and accuracy.* We have added this description at the end of this paragraph.

In the fourth paragraph in Section 1, we have described that the IAU and covariance inflation methods might be important schemes for dynamical balance and accuracy. To indicate the importance of the choice the IAU and covariance inflation, we have added the description of "*In EnKF-based ocean data assimilation system, how to apply the analysis update to the model evolution and how to inflate the ensemble spread could make significant differences for the dynamical balance and accuracy. However, the IAU and RTPP/RTPS has not been widely used in an EnKF-based ocean data assimilation systems (Table 1).*" Thus, following the editor's comment, we have attempted to specify the research question: to investigate the optimal schemes for IAU and covariance inflation methods to obtain accurate and balanced analyses.

As described in the second paragraph in Section 6, the suitable setting, especially for the RTPP parameter, depends on the other tuning parameters and experimental settings. However, the results in this study would be helpful for readers who newly establish and develop EnKF-based ocean data assimilation systems.

2)It is disappointing that authors attempt to bypass reviewers' critical comments either by saying that it is out of the scope of this study or by that it is in line with published articles. By doing so, the study loses its own values and becomes irrelevant for the broader community which may have to deal with the same issue. In my opinion, two things here can be accomplished. 1) Currently, authors only present the results but quite few efforts have been made to explain them. The places worth explaining have been pointed out by reviewers. Please avoid to use the expressions like "out the scope" or "in line with...". The best tuned parameters can be different in different systems or in different settings of the same system, therefore, it is necessary to understand why those values turn out to be the best. 2) As known, imbalance has great impacts on the forecast skills, it is inappropriate to conclude which combination is the best without showing their forecast skills.

We have used "out of the scope" and "future studies" only when the huge amounts of computational resources are required to conduct additional experiments for long periods (Reviewer 1 Major comment #4 and Reviewer 3 Major comment #5) and to investigate tuning parameters such as localization scale (Reviewer 1 Major comment #3 and Reviewer 2 Major comment #7). To show the correspondence to previous studies following the fourth major comment from Reviewer 2, we have additionally cited only Yan et al. (2014) who investigated impacts of the IAU on vertical velocity and accuracy in twin experiments using a relatively idealized EnKF-based ocean data assimilation

system. We hope your understanding that we have attempted to reply to reviewers' and editor's comments to the best of our ability.

Regarding the first point, we have included the following descriptions to provide the insight into how the IAU and covariance inflation methods contribute to dynamical balance and accuracy. In the first paragraph in subsection 4.1, the IAU reduces the noise from high-frequency gravity waves associated with initial shocks and improves the balance. In the first paragraph in subsection 4.2.1, in contrast, the IAU degrades the accuracy because the ensemble spread is relatively small and the IAU does not use the SSH analysis increments. The small analysis increments might result in better dynamical balance. In the first paragraph in subsection 4.2.1, RTPP and RTPS improves accuracy by inflating the ensemble spread but the resulting large analysis increments might reduce dynamical balance. In Section 5, we have discussed why the MULT does not work well. We have also added the discussion why the combination of the IAU and RTPP with relaxation parameter of $\alpha_{RTPP} = 0.8-0.9$ is the most suitable at the end of the first paragraph in Section 6.

Regarding the second point, we have to conduct all experiments again from the beginning if we investigate the accuracy of the prediction similar to Lei and Whitaker (2016) and He et al. (2020) as replied to Reviewer 3, and currently we do not have the huge amounts of computational resources for this purpose. Daily-mean outputs in all sensitivity experiments rather than instantaneous analyses are used to evaluate the accuracy, and the ΔNBE and RMSDs are relatively stable throughout the experimental periods in all experiments except for the MULT and MULT+IAU experiments. Consequently, the prediction accuracy during the short period of a few days would be the qualitatively same as the accuracy estimated from the daily-mean outputs.

3)I have the same feeling as the reviewer, it is difficult to understand the symbols for statistical significance. I am sure there is a better way to present this.

Following the comments from Reviewer 1 and the editor, we have modified the symbols in Figs. 1, 3, 4, 7, and 8. Throughout the manuscript, black stars, open circles and triangles, and closed circles and triangles denote the NOINFL+IAU experiment, significant improvement and degradation, and improvement and degradation with no significant differences, respectively.

Last but the least, I am interested in the definition of imbalance. Does it include the ageostrophic component? To see clean impacts on imbalance caused by data assimilation,

is it better to calculate the imbalance difference between background and analysis?

ΔNBE defined in this study include only the geostrophic velocity except for the ageostrophic velocity. Here, we assume that the ageostrophic velocity is caused by the wind stress only except for geostrophic shear following the classical Ekman theory (Cronin and Tozuka 2016). The wind field is not modified by data assimilation in this system, and therefore the ageostrophic velocity would not be changed. Therefore, it is not necessary to include the ageostrophic velocity in the formulation of the non-linear balance equation (NBE).

$\boldsymbol{u}^{\boldsymbol{a}} = \boldsymbol{u}^{\boldsymbol{f}} + \delta\boldsymbol{u}$ and $\eta^{a} = \eta^{f} + \delta\eta$ is substituted into the geostrophic equation, where superscripts a and f indicate analysis and forecast, respectively. Since the geostrophic balance is satisfied in the forecast field, the increment terms only remain in Eq. (6). The solution is theoretically same if the analyses itself are used.

■ Reference

Cronin MF, Tozuka T (2016) Steady state ocean response to wind forcing in extratropical frontal regions. Sci Rep 6:28842. https://doi.org/10.1038/srep28842