Response for the manuscript "Stochastically perturbed physics-tendencies based ensemble mean approach in the WRF model: a study for the North Indian Ocean tropical cyclones" by Gaurav Tiwari et al., Weather Clim. Dynam. Discuss., https://doi.org/10.5194/wcd-2022-10-RC1, 2022

We thank the editor and the Anonymous Referee #1 for their valuable comments. The suggestions have certainly improved the manuscript quality. A pointwise reply to the reviewer's comments is provided below.

Anonymous Referee's Comments

General assessment

The authors analyse the ensemble forecast performance of tropical cyclones when a stochastic disturbance is added to the physics trends (SPPT) in the WRF model. They also add a numerical filter initialization (DFI) to the initial state of the forecast. The performance of the DFI+SPPT ensemble forecast is evaluated against a control simulation and against a simulation that uses only DFI. These various model set ups are use to forcast two tropical cyclones which developed in the North Indian Ocean. The authors show that for a number of parameters evaluating the transverse and longitudinal position error, intensity, precipitation associated with CT, DFI+SPPT performs better. Therefore, they recommend that it be used in future forecasts made with WRF.

The paper and especially the result section describe the figures (which are of good quality) in great detail: there are many numbers and acronyms which sometimes make the reading difficult. The text could sometimes be simplified. While the conclusion merely summarises the main results, what is striking is the absence of any discussion of the results in the paper, which I think is necessary.

From a scientific point of view, the study of the impact of the SPPT parameterisation is relevant, since its usefulness for forecasting is debated. On the other hand, the evaluation made here is questionable: the DFI+SSPT ensemble is compared to a deterministic forecast of the CNTL and DFI, which should not allow to conclude (I develop this concern below). The difference between DFI+SPPT and the other experiments is always interpreted as due to the introduction of SPPT but could be equally due the size of the ensemble. Similarly, the difference between CNTL and the other experiments is always interpreted as the introduction of DFI, but could be equally due to the retuning that has been performed.

I therefore recommend that the paper be reconsidered for publication after a major revision.

Major concerns

Comment 1: I have a major concern about the comparison between DFI and DFI+SPPT. This comparison is essential because it supports the only major conclusion of the paper : "The SPPT based ensemble mean approach with digital filter initialization in the WRF model has shown considerable improvements in detecting the cyclone characteristics compared to other experiments."

The issue is that DFI has one single member while DFI+SPPT is an ensemble of 10 members. The reduction of error in DFI+SPPT could result from a better sampling of possible outcomes. A possible evidence of that is that the "best member" of DFI+SPPT has a comparable score to

DFI for the intensity metrics and is often in lesser agreement with the ensemble mean of DFI+SPPT, although it is one of its member!

A more rigorous assessment should compare two ensembles of similar sizes for DFI and DFI+SPPT.

Reply: Thank you for the comment. The essential objective of the study is to elaborate on the impact of a digital filter initialization (DFI) process within the WRF model on the historical prediction (simulation) of tropical cyclone activities over the North Indian Ocean. With this approach, we found remarkable improvements in the cyclone's track simulation (compared to the CNTL) but very less impact of DFI on the cyclone's intensity simulation. Thus, this prototype setup can be used for further future studies.

Moreover, to improve the skill of the WRF model (with DFI) for cyclone's intensity simulation, we generated the ensemble members using SPPT. We agree with the reviewer, that the reduction of error in DFI+SPPT could result from a better sampling of possible outcomes. Various studies have shown the improved precipitation in the ensemble mean (even further improved performance with increasing ensemble members) by comparing the control simulation with the ensemble mean of the perturbed simulation.

This study is not focused on the source of improvement but rather on obtaining the possible best setup and the results indicate the reduction of error in DFI+SPPT.

It is possible that in the large ensemble, some of the ensembles might be worse than in control which might lead to worsening the performance of the ensemble mean. In line with this, the purpose of keeping the Best Member was to illustrate the behaviour of a single member also, including the mean of all the members. Since the selection of the Best Member was based on cyclone's track simulation only, it exhibited a comparable score to DFI for the intensity metrics.

Comment 2: There is no mention of an ocean model in the model setup, so I assume that all experiments are atmosphere-only. It seems necessary to describe the SST product. In particular, because the surface latent heat flux is argued to be the cause of the errors in track of CNTL.

Reply: We thank the reviewer for the suggestion. This study uses a standalone atmospheric model (WRF) for all the experiments. We used a fixed SST value for the simulations, which is a necessary criterion for the WRF simulations with active DFI mode. Therefore, the SST product is not described in the manuscript. Having been an atmosphere model, the feedback of the oceanic processes is not attributed to the atmospheric processes; however, with a fixed SST also, the WRF model provides the value of surface latent heat flux as a variable at the user-specific time interval.

Comment 3: It is likely that there is no or a strong underestimation of the cold wake feedback, and as such, it is not surprising that the experiments tend to overestimate the intensity of the two TCs. The fact that DFI+SSPT captures the peak intensity of Nivar, while the other experiments overestimate it, is probably not a good thing, as the cyclone would have been weaker with a SST cooling. The authors should discuss that.

Reply: Thank you. We agree with the reviewer that underestimating the cold wake feedback in the model simulations can overestimate the intensity of the simulated cyclones. However, better sampling of the ensemble members in the DFI+SPPT experiments is the plausible reason for capturing the peak intensity of the cyclones regarding the observation.

In this study, the SST was kept constant throughout the simulation length (because of the requirement for the DFI scheme activation). Due to this, other mechanisms could probably be responsible for a such overestimation in the cyclone's intensity, especially from the CNTL, DFI, and Best Member experiments, which might be a great motivation for further research.

Comment 4: In their analysis of Fig. 11 and Fig. 12, the authors analyse the fact that the precipitation intensity is reduced in DFI+SPPT and in closer agreement with the observations as an improvement due the SPPT scheme. But it is most likely the result of averaging the ensemble. Again, an evidence of that is that the best member of DFI+SPPT has more intense precipitation than DFI+SPPT ensemble mean.

Reply: Thank you for the comment. We agree with the reviewer that improvement in the SPPT scheme is likely the result of averaging the ensemble. We included this discussion in the revised manuscript in Line No. 306-310 as follows:

"Thus, In Fig. 11 and Fig. 12, the cyclone-induced rainfall was better simulated by the DFI+SPPT experiment than by CNTL and DFI, mainly due to the better sampling of the ensemble members generated by the SPPT scheme. It illustrates that a set of small ensembles can enhance the WRF model framework skill for rainfall prediction/simulation."

Comment 5: A retuning of DFI and DFI+SPPT has been performed. Which parameters have been retuned? This retuning is as likely to explain the differences in track between CNTL and the other experiments, as the introduction of a DFI in the initial state. It should be described.

Reply: Thank you for this important comment. The setting for SPPT parameters is provided in Table 1. The standard deviation value of the random perturbation field at each grid point was set to 0.2. Random perturbation length-scale and temporal decorrelation of the random field were set to 150,000 meters and 21,600 seconds, respectively.

The (tuning or the) setting of the DFI parameters was based on the literature survey.

The tuning details and related reference is included in the revised manuscript as follows:

"A filter based on the Dolph-Chebyshev window (Lynch, 1997) was used in this study. The filter's backstop was set to 60 minutes, while the forward stop was set to 30 minutes".

Lynch, P.: The Dolph–Chebyshev window: A simple optimal filter. Monthly weather review 125, 655-660, 1997.

Comment 6: The authors suggest on the contrary that the lesser surface turbulent heat flux is the cause of the difference in CNTL track for Tauktae: could they test their hypothesis? I believe that the different sets of parameters would cause CNTL to track more west than the retuned experiments, which would cause lesser surface turbulent heat flux, rather than the contrary.

Reply: Thank you for the comment. Regarding mentioned point, it was found from the analysis that the higher (lesser) accumulated upward latent heat flux in the left (right) of the simulated track was one of the primary causes of the difference in the CNTL track for Tauktae. As suggested, further we tested our argument by evaluating the simulated spatial distribution of lower tropospheric (950 hPa) absolute vorticity for cyclone Tauktae from all the experiments (shown in Fig. S12). We selected two successive time steps (4^{th} and 5^{th}) to capture the behaviour of the cyclonic vortex toward its movement.

Particularly from the CNTL simulation, the contour of vorticity maxima at the 4^{th} time step (Fig. S12a) was in the west of the cyclone's track which got shifted to the south of the track in the 5th time step (Fig. S12b); whereas, the storm still travelled north-westward. More or less, a similar sort of behaviour was seen in other experiments too. Thus, it can be said that upward latent heat flux played a pivotal role in directing the cyclone Tauktae from CNTL simulation more west than the retuned experiments.

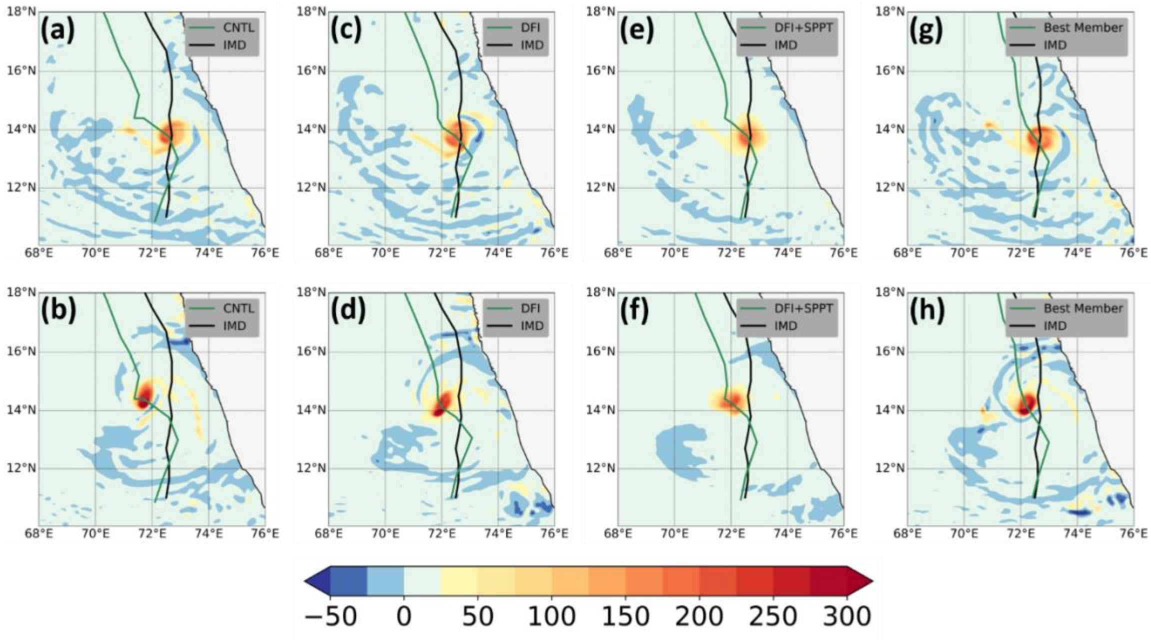

Figure S12: Spatial distribution of absolute vorticity of cyclone Tauktae in the lower troposphere (950 hPa) plotted for the 4th (first row) and 5th time steps (second row) for CNTL (first column), DFI (second column), DFI+SPPT (third column), and Best Member (fourth column) experiments.

Minor revisions

Comment 7: line 33 : “frequent and intense TCs” is a confusing statement. Does it mean “more frequent and more intense TCs” : there is certainly no consensus on an increase of TC frequency! Or do they mean “more frequent intense TCs”?

Reply: Thank you for the comment. In this statement, we want to convey that the frequency and intensity of TCs in the ARB is expected to increase compared to the historical database. The sentence has been rephrased in the revised manuscript as follows:

In the NIO, particularly in the ARB, the frequency of TCs is expected to increase compared to the historical database. At the same time, their intensity will also increase as the earth continues to warm (Deshpande et al., 2021).

Technical issues

Comment 8: line 52 : influence -> influences

Reply: Thank you. The suggestion has implemented in the revised manuscript, Line no. 58.

Comment 9: line 125 : Why were two different convection schemes used?

Reply: Thank you for the comment. The convection scheme is one of the key constraints that control the model performance. Recent studies found that the same scheme is performing best for one region while worse for other regions and suggested using the mixed convection scheme (different scheme over the different areas) (Mishra and Dwivedi et al., 2019). With these lines, We performed some sensitivity runs and selected the best-performing convection scheme. The Kain-Fritsch cumulus convection scheme over the ARB and Grell 3D Ensemble scheme over the BoB domain to get the best-simulated tracks of TCs. Rest other schemes were similar for both domains.

Comment 10: line 198 : from the Best Member -> for the Best Member.

Reply: Thank you. The suggestion has implemented in the revised manuscript, Line no. 205-206.

Comment 11: line 384 : which was occured -> which occurred

Reply: Thank you. The suggestion has implemented in the revised manuscript, Line no. 414.