# Comment on wcd-2022-10

Anonymous Referee #1

---

Referee comment on "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

---

**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+SSPT 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+SSPT 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+SSPT 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**

[1] I have a major concern about the comparison between DFI and DFI+SSPT. 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+SSPT is an ensemble of 10 members. The reduction of error in DFI+SSPT could result from a better sampling of possible outcomes. A possible evidence of that is that the "best member" of DFI+SSPT has a comparable score to DFI for the intensity metrics and is often in lesser agreement with the ensemble mean of DFI+SSPT, although it is one of its member!

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

[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.

[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.

[4] In their analysis of Fig 11 and Fig. 12, the authors analyse the fact that the precipitation intensity is reduced in DFI+SSPT 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+SSPT has more intense precipitation than DFI+SSPT ensemble mean.

[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.

[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.

**Minor revisions**

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"?

**Technical issues**

line 52 : influence -> influences

line 125 : Why were two different convection schemes used?

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

line 384 : which was occured -> which occured