<u>Response to Reviewer 3's comments</u>

We thank Reviewer 3 for the time to go through our manuscript in details. This manuscript describes a new and efficient method to produce a physical TC event set in the western North Pacific basin. In general, reviewers think after careful revision, the results of this study is of great interest and relevance, and it will be a nice contribution to the field of TC risk assessment. Here is our point-to-point response to Reviewer 3's comments.

<u>General Comment</u>

I think that the topic of this study is of great interest and relevance, and that it is suitable to NHESS. Besides, the paper is generally well written, the methodology clearly illustrated and the results well presented and discussed. However, there are also a few (minor) corrections and some improvements of the text that could be made to further improve the manuscript before to proceed with its publication.

Therefore, my recommendation is to accept the manuscript for publication after ***minor revisions****.*

<u>Specific Remarks</u>

1. *Page 1, line 14: "… characteristics of the new event set is consistent to the…" should read "… characteristics of the new event set are consistent to the…"*

We thank Reviewer 3 for pointing this out, **we have corrected this in the revised manuscript**.

2. *Page 1, line 24: "… 67.1 billion RMB …" Many readers could be helped to understand the economic significance of this figure by accompanying it with the corresponding value in US Dollars or Euros.*

We thank Reviewer 3's suggestion, **we have added the corresponding value in Euros in the revised manuscript**.

3. Page 2, line 45: "… (ii) the storms in the typhoon event set might not be physically consistent." Please, clarify what do you exactly mean here with "physically consistent"?

It means event sets created by stochastic perturbations will create TC events that (with respect to their inner dynamical structure) are not necessarily physically consistent anymore. As just surface footprints are stochastically modelled from existing tracks, there is no check whether those events (in the stochastically modelled from) are physically possible and how they could be realised in a fully dynamical consistent view, thus fulfilling all known physical relations and derived constraints by the means of physical laws. Consequently, the amount of unrealistic physical properties due to the oversimplified stochastic simulation is unknown and laws of physical interactions are potentially ignored. **We have modified the sentence in the revised manuscript to clarify this point [see lines 46-52]**.

4. *Page 2, line 63–64: "In this study, we show the TPEPS event set has much higher information content: more TC events and more extremely high impact TC events." Higher and more than what?*

We thank Reviewer 3 for pointing out this point. **This sentence should read "*In this study, we show the TPEPS event set has much higher information content: more TC events and more extremely high impact TC events than historical or reanalysis-based TC event set.*" (Lines 70-71)**

5. *Page 4. Line 126: "WiTRACK identifies windstorm events of all kind, including MEPS TCs, PEPS TCs, MEPS extratropical cyclones." I suppose it identifies also PEPS extratropical cyclones.*

Yes, it does. **We have added PEPS extratropical cyclones to the list for clarification**.

6. *Page 4, line 175: "The removal of these events ensures the TPEPS event set is independent of any pre-existing weather patterns." The goal here is to build a large set of typhoon events in order to provide a solid statistical evaluation of their characteristics, so why is it so important that the considered TPEPS events are independent of any pre-existing weather patterns?*

To use this as an extension of event numbers and thus as an alternative reality, the inclusion of real existing events will incorporate some bias towards observed events as all of them will create a multiple realisation in the ensemble members started at the time such a real event occurred. By not considering those ensemble members, which are closely related to observed events, will secure that indeed new events are used to build the pure EPS event set. It has to be noted though that the inclusion of those events should not change the overall track distribution, or in other words, the track distribution from pure EPS and real EPS events is fairly similar.

7. *Page 6, line 193, Figure 1: please add the units to the colour bar.*

We have added the unit to colour bar.

8. *Page 6, line 197, Table 5: Why there is such a large difference in the number of simulated TC wind storms between the TIGGE models? Is this due to the different number of ensemble members of the EPSs? The large majority of the considered TPEPS are from two EPSs: the ECMWF and the NCEP. What consequences could this fact have on the analysis results?*

The main reasons for differences in the number of detected TC windstorms between TIGGE models are they have (1) different numbers of ensemble members of the EPSs, (2) different number of runs per day, and (3) different maximum forecast lead time (c.f. Table 1). Given the spatial and temporal distributions of the individual PEPS event sets are similar to each other, the analysis on the overall TPEPS event set is reliable.

9. Page 6, line 202: Fig. 1d, should read Fig. 2d.

We thank Reviewer 3 for pointing this out, **we have corrected this in the revised manuscript**.

10. Page 6, line 203: Fig. 2 should read Fig. 3.

We thank Reviewer 3 for pointing this out, **we have corrected this in the revised manuscript**.

11. Page 6–7, line 212–220: I'm not sure I fully understand the explanation the authors provide for the discrepancy between the spatial distribution of the TPEPS event set and JRA–55 events as shown in Figure 2 (panels c and f). The fact that the JRA-55 event set can be considered as a subset of the TIGGE event set does not explain the difference in spatial distribution. According to this view, in fact, the JRA-55 events can be seen as randomly selected from a larger set (the TIGGE set), and thus they should also be spatially distributed as this event set. Also, why the higher level of the 98th percentile values of the JRA-55 wind should explain the lower number of typhoons in this area?

We agree with Reviewer 3 that the JRA-55 event set can be seen as a subset randomly selected from a larger set (i.e. the TIGGE event set). This means if we randomly sample the TPEPS event set, we can obtain a subset highly similar to the JRA-55 event set. For demonstration, we have conducted bootstrap resampling on the TPEPS event set to obtain 10,000 sets of subsample. Each set of subsamples has 668 events to mimic the number of events in the JRA-55 event set. For each set of subsamples, the track density is calculated, and used to calculate uncentred pattern correlation between the resampling set of subsamples and the JRA-55 event set. In order to focus on relevant entries, for a particular grid box, if the values of track density for a resampling set and the JRA-55 event set are both less than one, such grid box is not used in the pattern correlation calculation. The mean, standard deviation, minimum, and maximum of the uncentred pattern correlation of the 10,000 set of subsamples are 0.9380, 0.0107, 0.8961, and 0.9697, respectively. This suggests the spatial pattern of the JRA-55 event set is highly similar to a small random subset of the TPEPS event set. Consequently, the JRA-55 event set can be seen as a subset randomly selected from the TPEPS event set. On the other hand, it is **not** be possible to deduce the basic population (e.g. the TPEPS event set) from a small sample set (e.g. the JRA-55 event set). Although the spatial distribution of the small set sample is similar to the subsamples of the basic population and thus usable as one possible realisation of the basic population, the small sample set does not contain all of the information of the underlying population. Furthermore, the statistical estimate of extremes would also be different for the small sample set (e.g. JRA-55 event set) and the basic population (e.g. TPEPS event set). **We have included the above explanation in the revised manuscript (Lines 259-280).**

Upon further investigation, we found that the 98th percentile is not the reason that leads to the differences in spatial distribution. The major difference between the track density of TPEPS and JRA-55 is that there is an eastward bias in the TPEPS. There are several reasons that could contribute to this. The eastward bias in the track density appears to be a common feature in many GCMs (e.g. Camargo et al., 2005; Bell et al., 2013; Roberts et al., 2020), this has also been observed in seasonal forecast output (Camp et al., 2015). Finite simulation time has also contributed to this bias as TC that forms in the region east of 150 °E would not have the time to move into the western part of WNP before the end of simulation time. Differences in the amount of tracks could also contribute to the differences as more diverse tracks would be captured. **We have added a respective explanatory comment at lines 252-258**.

12. *Page 7, line 248–249: As formulated here, this sentence seems to imply that TCs with weaker winds are also less spatially extended, which is not true.*

The impact area of a TC in this study refers to the total area which has experienced TC-associated extreme wind (i.e. larger than local climatological 98th percentile wind speed). Given the fact that the wind speed of TC wind field decays radially outward, TC with weaker winds would have a smaller impact area because the outer wind speed would be below the 98th local climatological wind percentile value. **We have added more descriptions about impact area in the revised manuscript to clarify this point [Lines 305-306]**.

13. *Page 7, 252–255: "… impact (Befort et al., 2020). Many of the low impact TCs …" should probably read "… impact (Befort et al., 2020), many of the low impact TCs …".*

We thank Reviewer 3 for pointing this out, **we have corrected this in the revised manuscript**.

14. *Figure 8: In the text of the manuscript, there are references to panels labelled with letters (a, b, … f), but the panels in Figure 8 are not labelled.*

We thank Reviewer 3 for spotting this error. **We have corrected this in the revised manuscript**.

15. *Page 9, line 317: "… based on minimisation of the root-mean-square-error (RMSE) of …". Of what?*

We thank Reviewer 3 for pointing this out, **we have corrected this in the revised manuscript**. This is the root-mean-square-error of the quantile mapping output.

References

Bell, R., Strachan, J., Vidale, P. L., Hodges, K., and Roberts, M.: Response of Tropical Cyclones to Idealized Climate Change Experiments in a Global High-Resolution Coupled General Circulation Model, J Climate, 26, 7966-7980, 10.1175/JCLI-D-12-00749.1, 2013.

Camargo, S. J., Barnston, A. G., and Zebiak, S. E.: A statistical assessment of tropical cyclone activity in atmospheric general circulation models, Tellus A, 57, 589-604, 10.1111/j.1600-0870.2005.00117.x, 2005.

Camp, J., Roberts, M., MacLachlan, C., Wallace, E., Hermanson, L., Brookshaw, A., Arribas, A., and Scaife, A. A.: Seasonal forecasting of tropical storms using the Met Office GloSea5 seasonal forecast system, Q J Roy Meteor Soc, 141, 2206-2219, 10.1002/qj.2516, 2015.

Roberts, M. J., Camp, J., Seddon, J., Vidale, P. L., Hodges, K., Vanniere, B., Mecking, J., Haarsma, R., Bellucci, A., Scoccimarro, E., Caron, L.-P., Chauvin, F., Terray, L., Valcke, S., Moine, M.-P., Putrasahan, D., Roberts, C., Senan, R., Zarzycki, C., and Ullrich, P.: Impact of Model Resolution on Tropical Cyclone Simulation Using the HighResMIP–PRIMAVERA Multimodel Ensemble, J Climate, 33, 2557-2583, 10.1175/JCLI-D-19-0639.1, 2020.