

The authors have made some effective revisions based on my previous review comments, such as adding FUXI in Figure 6 and including a typhoon case test. Unfortunately, the authors have not provided convincing responses to the deficiencies in the scientific rigor of the experimental setup. Therefore, I still believe that the article is not yet suitable for publication in *Remote Sensing*. The following are specific comments:

1. It needs to be reiterated that it is meaningless to drive PANGU with ERA5 when comparing it with numerical models such as GFS and EC, not solely due to the delay of ERA5. Operational numerical models obtain initial conditions through cyclic assimilation, and there is an objective quality gap between them with reanalysis datasets. In other words, if ERA5 must be used to initialize PANGU, then the comparison experiments including both NWP and AI models should also be initialized by ERA5. However, it is evidently impractical to rerun numerical models. Therefore, PANGU should only be driven by global model analysis fields such as GFS, EC, or CMA-GFS to have comparative value with numerical forecasts. As for the authors' mention of exploring the forecasting limit of PANGU, this has already been revealed in related articles and does not need repeated demonstration. Therefore, the authors need to replace the experiments and adjust the article structure accordingly.
2. I would like to discuss the main focus of the article again. The authors mention that the purpose of the article is to compare the surface meteorological elements forecasting skills of PANGU with traditional NWP. However, this is almost a consensus, and the focus of a research paper should be forward-looking. On the other hand, the China Meteorological Administration's large model comparison project has already conducted comparisons of meteorological large models in real forecasting scenarios (<http://aimfdp.nmc.cn/ai/dist/>).
3. Compared with FUXI and FENGWU, PANGU seems to have ceased updating.
4. Figure 10 should select a later initial forecast time to reveal the forecasting abilities of various AI models before and after typhoon landfall.