Review of: HyPhAI v1.0: Hybrid Physics-AI architecture for cloud cover nowcasting

In this paper, the authors introduced a novel hybrid model for cloud cover nowcasting. They utilize the advection equation to advance cloud cover maps, with velocities estimated through a neural network. The network parameters are optimized end-to-end to minimize cross-entropy loss between predicted cloud type probabilities and actual ones at subsequent time steps. I find the overall concept of the paper both innovative and appealing. Furthermore, the authors have included a range of results and diagnostics to evaluate the effectiveness of their methodology. Overall, I believe the paper is suitable for the GMD audience. However, there are certain aspects, particularly concerning the paper's writing style, that could benefit from further refinement.

Major comments :

1. Introduction and positioning of the paper: I appreciate the authors' introduction and I think that it effectively sets the stage for the study. However, I believe additional refinement could be beneficial. For example, around line 65, the authors discuss residual modeling as a methodology to mitigate imperfections in physics-based models. From my understanding, residual modeling in this context means using machine learning to correct errors of physical models. I suggest to avoid using the term "residual modeling" here, as it could potentially be confused with residual networks.

   In the same context, the authors mention that model correction "does not have the ability to enforce physics-based constraint." Correcting model errors can be approached in various ways. For instance, one approach involves defining machine learning-based parameterizations of physical models, where deep learning doesn't model a residual but instead represents missing physical processes. Recent advancements in the field have explored incorporating various physical constraints into such deep learning-based corrections. I suspect that this is what the authors are referring to in lines 70 to 80. I suggest that the authors revise this section and include references to recent state-of-the-art research to provide further context and clarity.

2. The majority of section two is not necessary and should be removed. It is quite standard to write numerical schemes in PyTorch. You can just mention in section three that you implemented your numerical advection scheme in PyTorch, and you can discuss the space and time discretizations you used in the appendix. Everything else including the discussion on automatic differentiation, gradient decent and the subsection 2.1 on combining neural nets with physics should be removed.

3. I also suggest to trim down the text in the remaining sections for example: around line 330 you should remove "which rely on computing the loss gradients with respect

to the model parameters. These gradients guide the update of the model's weights during the training process"

4. The authors developed 4 different versions of their model based on different parameterizations of the source term. However, in the experiments, only two configurations, the HyPhAI-1 and 2 are tested. If you plan to only test these two configurations, you should remove HyPhAI-3 and 4.

I hope that the authors will understand my comments in a constructive way, and that I value their work and the time they invested in the preparation of the manuscript. It might be that I have misunderstood something, in this case, if something wasn't clear for me as a reviewer, it is possible that it wouldn't be clear also for the readers.