Response to Reviewers

Reviewer #2

We thank the reviewer for the careful read of the manuscript.

Before addressing Reviewer #2's comments we note that we have reorganized the manuscript subtantially in order to incorporate Reviewer #1's detailed suggestions. We have also informed the editorial team and wish to state here that we have added Timothy A. Smith, NOAA Physical Sciences Lab, Boulder, CO, USA, to the team of authors.

In the following we address the reviewer's comments (reviewer's comments in red, our response in black).

This is an interesting review of the current status of Machine Learning for Ocean Forecasting, especially for people from outside the subject domain.
We thank the Reviewer for their comment.

The main concern I have is that level of detail of the discussion is rather uneven. For example, the discussion of Sec. 3 "Enhancing data assimilation with ML algorithms" seems just a placeholder for further development. The alternative is that not a lot of activity has been going in the field, in which case this should be stated.
The impression is mostly correct. Most efforts have been dedicated to surrogate modeling, either of the full ocean GCM or of components – parameterization schemes in particular – of the model. Whereas hybrid DA/ML methods have not been as widespread yet, they are an important application of ML and we disuss them here.

Following the reviewer's suggestion, we are now stating the relative paucity of related activities in ocean modeling. Among others, the revised manuscript now contains the following statement:

> "The use of hybrid DA/ML approaches, be it in the context of ensemble DA or adjoint-based methods (e.g., 4DVar) presents substantial algorithmic hurdles (e.g., availability of a differentiable dynamical core in the context of adjoint-based DA), which explains the relative paucity of such studies to date compared to purely data-driven methods.

Other comments are posted in the attached annotated version of the manuscript.
We have addressed all comments in the annotated PDF. Below, we take up those comments in need of a response.

Line 86:
Maybe some high level explanation of how this approach "tries to solve differential equations using NNs" would useful here.
We have now removed this generic statement in favor of a more detailed list of ML approaches that have been explored in the context of numerical weather prediction.

Line 115:
That is a very fashionable definition! In practice DTs are effectively high resolution versions of standard NWP or Earth System numerical models and their ensemble implementation. So I would take this definition off, as it does not bring additional information to the discussion.
We disagree with this statement, in particular with the notion that DTs "are effectively high-resolution versions of standard NWP or Earth System numerical models and their ensemble implementation". The US National Academies' (2023) which we are citing makes the specific point that this is a misguided concept of DTs, although it is indeed used by various groups. We choose to stick to the definition (and vision) of DTs as laid out in the National Academies' (2023) report.

Line 137:
Would be interesting to provide examples of where this ability to incorporate physical laws and constraints has been demonstrated.
In the revised and restructured version, we now clarify how ML approaches are combined with physical laws in new section 2.1 "Hybrid physics-ML models: enhancing forecast models and data assimilation with ML algorithms".