Earth Syst. Dynam. Discuss., author comment AC2
https://doi.org/10.5194/esd-2021-105-AC2, 2022

Reply on RC2

Nicola Maher et al.

Author comment on "Combining machine learning and SMILEs to classify, better understand, and project changes in ENSO events" by Nicola Maher et al., Earth Syst. Dynam. Discuss., https://doi.org/10.5194/esd-2021-105-AC2, 2022

This paper applied the integrated observational dataset to train the classification of the EP El Niño, CP type El Niño, and La Niña with supervised learning and to investigate the ENSO diversity/complexity changes in multi-model large ensembles. Specifically, they found the supervised machine learning can reasonably classify ENSO events/types and the observed increase of CP El Niño events is within the range of internal variability, so does the ENSO amplitude and frequency changes. The research topic is interesting and necessary; however, there are issues in the machine learning setup and the goal/finding is not unique for machine learning. Therefore, this paper should not be accepted in Earth System Dynamics before major revisions.

A few major comments are followings:

We thank the reviewer for their constructive comments.

ML related

- The setup of the supervised learning uses the combination of 18 observational datasets However, the combination of 18 observational datasets may overweight a few events and have limited difference. For instance, the events after 1980 are covered for most datasets but the events before are only covered by half of them. The authors should discuss this issue and provide additional analyses in the supplementary. One suggestion is to test with subgroup of the datasets. Another issue for the integrated observational datasets is the lack of differences for the dataset. Even though the reconstructions are all slightly different, the SSTs are still representing the same events. That is, the actual events consider in this study is only 14 CP, 20 EP, and 26 LN. This issue should be mentioned in the manuscript and needs to be tested with a small subgroup (or even extremely just one dataset) of datasets.

We will retest the algorithm with only HadISST and add the results to Table 3. However,

we will continue to use this method as shown by Pascolini-Campbell et al 2015, events are classified quite differently when using different data products. As such we believe the use of many products is justified in this study.

- The setup of the supervised learning uses the features from 5 regions from October to March. However, limited dynamical reasons are provided and other regions and times should be mentioned (or even tested). For example, the authors show results from the smaller regions and times in the supplementary, but not larger regions and times. For instance, the north subtropical region is known to be important for the onset of CP El Niño and recent papers have found an improvement from including it (Tseng et al., 2021). And the summer is related to how specific ENSO type is onset (Yu & Fang 2018). The authors should provide dynamical reasons for the choice of the regions and times, otherwise, the study should examine more regions and times for showing the current choice is an optimal one.

We believe that keeping this is important because it provides context by comparing to other studies and it shows there is no clear relationship between amplitude change and the zonal gradient change when evaluating many SMILEs. While we don't explain the dynamics, our approach is a purely statistical exercise that makes use of more data than other studies and is therefore more robust and a valuable contribution.

- The comparison of the increased CP El Niño frequency to SMILEs should be more precise. The authors use the ensemble spreads in each year to consider as the range of change for the internal variability; however, this is different with the increased CP El Niño frequency over a certain period. The authors should check how large the CP El Niño frequency can change in each ensemble and discuss the spread of the changes for all SMILEs.

This is why we added the ensemble member 1. We will assess the max trend and min trend of the ensembles to better evaluate this in the revised manuscript.

Minor comments are provided below:

All minor comments will be address in the revision. For 3. We will revise the Figure to make it clearer and easier to interpret.

-

Does the training and classification use the original SST or SST anomalies? Please clearly describe in the text.

- The calculation of frequency should also be mentioned in the method section, not only in the caption of Figure 3.

- The Figure 6 is a bit difficult to read as there are many colors and lines.

- Line 205, ‘to far’

- Line 48, ‘niños’