

Comments Reviewer 1

I like the method presented here as producing (cheap, and therefore more flexible) forecasts that are complimentary to the traditional seasonal to decadal forecasts using an initialised climate model. I don't have many other comments and I think the initial reviewer comments have been well addressed by the authors.

We thank the reviewer for taking the time to review the manuscript once again and for providing constructive comments.

Small typos:

L 37: "associated with phenomena"
It has been changed.

L 64: "lack of a more"
It has been changed.

L 65: Remove the full stop between "shocks" and "and"
It has been changed.

L 298: "Africa and Australia"
It has been changed.

L 417: as a representative model
It has been changed.

L 431: seamless manner. Crucially" (add the missing full stop)
It has been changed.

Main comments:

1. "A key strength of the analog-based method is its ability to leverage a large-sized ensemble at minimal computational cost as opposed to the significant cost it requires to generate such large ensembles with initialized prediction systems." (L 169)
I mostly agree with this statement, however is there a slight nuance here? The method you use for analog selection is a simple ranking of the MAE, so as you increase the analog ensemble size, are you also degrading the ensemble selection by selecting ensemble members that are further away from the initial state? You demonstrate in the comments to the original reviewer that 149 members are generally more skilful than 25, but at some point, that will not be true.

Figure S1 in the supplementary material shows the fraction of land area with statistically significant positive correlation as a function of ensemble size for all variables studied (SPI-

3, 12, 24 and 48 and TAS 3, 12, 24 and 48 months). It is quite clear from the figure that the skill of the forecasts increases with ensemble size with no sign of reaching a maxima with ensembles smaller than 149, which is our ensemble size maximum. We agree with the reviewer about the potential shortcomings of selecting large ensembles using the analog method. In the previous review we had written in the methods section (L62-65) the following sentences to make it clear “An incomplete representation of the climate state at initialization is likely the major disadvantage of the analog-based predictions because of the finite available states present in the multi-model catalog. These states may be less representative or “further away” from the observed target state than the initial states in an initialized climate prediction system. ”

2. The discussion about Figure 4 around line 240: Could a reference to the figure number be added in the text?

It has been added.

3. “This implies that a recalibration of the ensemble could render the analog-based forecasts 260 more valuable.” (L 259)

Could you add to this comment? I’m not sure how the recalibration could be done. Would the recalibration be done to the field before choosing the analogs? Could it be that the reduction in variability is due to the analog sampling. If many approximate matches are sampled and average, the size of the signal may be reduced?

This is a good point. We think that the only possible calibration would have to be done a posteriori as it is usually done for other predictions. We have added before line 239 the following sentences to address the comment. “This underestimation of the local variability may be a result of the sampling and averaging of hundreds of analogs with some of them being less representative of the observed conditions in the particular area. Despite this, at the global scale, the larger the ensemble of analogs the higher the skill (Fig. S1). As commonly done for other types of climate predictions, recalibrating the ensemble could make the analog-based forecasts more valuable.”