

CoEditor Comments:

Many thanks the revisions of your paper. I sent the new version to the 2nd reviewer who suggested major revisions. As you can see in the response, the reviewer appreciates your responses to the issues about the sampling and the convergence toward the climatology, but strongly suggests to include them (maybe in condensed form) in the paper. I also had the impression that given the fact that the reviewer suggested "major revisions" you actually changed little in the paper. I suggest that you follow the reviewer's advice and include more insights about the two mentioned aspects.

I am looking forward to the updated version of your paper.

Authors' Response:

Thank you for the feedback and kind words. We have revised the manuscript, including discussion surrounding the distinction from climatological behavior. This can be found on line 290 and beyond. We choose, however, to supplement some discussion in our author's response in order to emphasize the difficulty of answering this question beyond what is now provided in the text:

"If the occurrence of bimodality is a result of the climatological distribution itself, one would expect occurrence rates to asymptote at very long forecast lead times as the forecasts approach the climatological distribution. However, if bimodality is occurring due to particular atmospheric states, one would expect there to be some 'peak' in bimodal frequency at some function of lead time prior to an asymptotic behavior, this 'peak' period would be representative of the second period in Figure 5.

This in itself, however, is a difficult question to answer systematically and requires a study of its own. The reason being, there may be some skill to forecasts even at very long lead times (weeks 4+) which cause them to not yet approach the climatological distribution (and thus asymptotic behavior in bimodal frequency is not yet exhibited). This would be expected to be especially prevalent in a location like the tropics, where the effect of ENSO has a signal that spans beyond the sub-seasonal timescale. Thus, ensemble forecasts that are both able to simulate particular weather events (which may cause the development of bimodality), as well as span long enough (perhaps much longer than the 6+ weeks of ECMWF), would be needed globally. This on top of the fact that large enough ensembles are needed to exhibit bimodality in the first place (refer back to Sect.2) and that a great number of forecasts are desired to achieve a sufficient sample size for the study, a constraint by computational power is created."

Furthermore, as a note, we have revised some content within the skill scoring section 2.3. Mainly in regards to being more explicit in our discussion and the expansion of Table 1 with ensembles of different sizes. We hope this offers a clearer discussion of a rather important point to the skill scoring community.