Review- Bimodality in Ensemble Forecasts of 2-Meter Temperature- Identification

I have reviewed the paper "Bimodality in Ensemble Forecasts of 2-Meter Temperature- Identification" by Cameron Drew Bertossa, Peter Hitchcock, Arthur DeGaetano, and Riwal Plougonven.

In this manuscript the authors examine the non-Gaussianity (and specifically, bimodality) that arise in ensemble forecasts of T2m temperature. They show that fitting a Gaussian distribution to the ensemble PDF, which is commonly used, as well as the use of traditional skill scoring methods, performs poorly if bimodality is present. They then identify bimodality in ECMWF forecasts globally, and identify regions where bimodality appears strong.

This is a very interesting paper and I enjoyed reading it, it is well-written and presents interesting results. I think this is a very important issue that is often overlooked, and work of this sort has the potential to improve our forecast ability and understanding. I therefore recommend accepting the paper for publications in WCD and have only a few minor comments outlined below.

Comments:

- Line 35: What do you mean by "appropriate way"? not clear here. Do you mean the SVD analysis that you mention later?

- Eq.3- n is not defined here

- Lines 157 and 159: "skewed–> did you mean "shifted? that would be more appropriate here.

- Fig. 4b: I found this figure rather confusing. First, you only show the False Negative of the three distributions here, but from the text I was expecting to find also the False Positive. Also, why do you show the false positive of the Gaussian and not the false negative so we can compare it with the other distributions?

- Line 232: I'm not sure how you reached the 1.18 and 85% cutoff. Can you better explain this?

- Line 245: How do you determine the theoretical value of delta for the transition from unimodal to bimodal? From what I gather, the condition is $|\mu_2 - \mu_2| \leq 2\sigma$, regardless of p, which gives δ=2 even if p has now changed from $\frac{1}{2}$ to $\frac{3}{4}$. I assume you are not wrong, so I would love to understand what I am missing.

- Line 318 and Fig.3a: The fact that the mode ratio relies on the minimum (in the denominator), which is often close to zero, is a disadvantage as it can give high values just because it is low. Why not just take the ratio between to the peaks?

- General:

- What happens if there are actually more than two peaks?

-Can we somehow asses how much of the bimodality is originating from bimodality in the PDF of each forecast (e.g., perhaps in all the simulations the PDFs exhibit two maxima) as opposed to bimodality in the ensemble which results from different "regimes"? This is interesting because the spatial maps of delta look very much like regions of high temperature skewness.