Comment on npg-2022-15
Anonymous Referee #2


REVIEW FOR "The Role of Internal Variability in Regional Climate Change", by Clara Deser and Adam S. Phillips, submitted to Nonlinear Processes in Geophysics

Summary:
The authors analyse historical and future European temperature and precipitation trends in the CESM2 Large Ensemble, observations, and an "Observational Large Ensemble". The analysis is based on an analogue-based dynamical adjustment method, which is used to disentangle dynamical (based on SLP analogues) and thermodynamical (residual) trends in the climate model and observational ensemble. As a result, the authors show that internal climate variability is a crucial source of uncertainty in future European climate, the level of which is broadly comparable in magnitude between the Obs-LE and CESM2-LE, and that the thermodynamical component of observations agrees well with the forced CESM2-LE component for temperature, and less well for precipitation.

Overall, the paper provides a very useful illustration of internal variability in present and future European climate, and a constructive discussion of current and outstanding issues in dynamical adjustment. The paper is also very well written and logically structured. I still have a few concerns that are outlined below, however, and I would therefore recommend moderate revisions.

Major issues:

(1) Abstract
The Abstract is well-written, however it is somewhat disconnected from the actual analysis conducted in the paper. At present the Abstract reads a bit like that from a Perspective paper, while the (by far) largest part of the paper presents actually a specific analysis of European climate. Hence, I would recommend to adjust the Abstract such that it (also) reflects the analysis conducted in the paper.
(2) Implications of high climate sensitivity in CESM2 for interpretation of thermodynamical trends
The authors interpret "the good agreement between the observed thermodynamic-residual trend component and the model's forced thermodynamic trend" (l. 604) as "further underscoring the realism of CESM2" (l. 605), and that "the model's forced temperature trend is realistic" as a powerful conclusion (l. 521). This conclusion is based on the temperature dynamical adjustment discussed on p. 30, where the authors argue that "observed thermodynamic trend is much closer in amplitude (and arguably pattern) to the model's forced response".
While I agree that these results are in general really encouraging, I do think that some caution is warranted: CESM2 is known for high climate sensitivity, so (I believe) we *should* expect some discrepancy in the amplitude of the pattern, and -contrariwise- a higher similarity in the pattern itself. Hence, why is the observed thermodynamical pattern's amplitude over Europe so high as to even match that of a high climate sensitivity model?

Moreover, for precipitation more careful conclusions would be warranted, as the residual component does not closely resemble the model's forced response. For example, the authors attribute the (large) pattern disagreement in Central Europe to "lower signal-to-noise" found in this region compared to other areas (l. 544-546), and further pattern disagreement over large areas in South Europe, such as the Balkans, Turkey, and Italy is only briefly mentioned. Here, I believe it would benefit the discussions if the authors would discuss this a bit more in-depth, and explain where the "lower signal-to-noise" explanation in this region comes from (because this is a transitional region between southern drying and northern wettening?).

(3) The authors assume that the forced CESM2 trend (i.e., ensemble average) reflects the thermodynamical response to climate change. This is consistent with literature, but reflects some simplification, which the authors acknowledge in their discussion. But, the authors also say that "future trends in SLP also contain a modest forced component indicative of enhanced westerlies over the continent" (l. 479), and I believe there may be in addition nuanced forced dynamical components with only a modest SLP signature. In earlier literature (Deser et al. 2016), the authors actually use their dynamical adjustment method to show in their Fig. 7 that the average across the dynamical contribution is rather small. I believe it may possibly benefit the present paper and argument to include and discuss a similar figure for Europe?

Minor issues:
 l. 150. McKinnon et al. 2017 is missing in the references section
 l. 631. The connection to predictability studies and the "signal-to-noise paradox" is interesting, but the short discussion is hard to follow. Maybe the implications could be made a bit more explicit here.
References: