Earth Syst. Dynam. Discuss.,
https://doi.org/10.5194/esd-2019-67-RC2, 2020

Interactive comment on "Comparing internal variabilities in three regional single model initial-condition large ensembles (SMILE) over Europe" *by* Fabian von Trentini et al.

Anonymous Referee #2

Received and published: 13 February 2020

The authors use large ensembles to compare the representation of internal variability in three regional climate models forced by historical and a future scenario forcing. They use observation-based data as a benchmark for the historical period. Large ensemble simulations of single climate models are an essential tool for estimating uncertainty of climate change projections due to internal unforced variability, for detection and attribution studies and so forth. The present study is therefore useful as a validation of such tools and to enhance our understanding of unforced internal climate variability at regional scales. My main concerns with this work however are its presentation, which is confusing at times, the implementation and interpretation of the methodology, and the interpretation of results.

Major:

1. The authors recognize that there is confusion in the literature on what is meant by “internal variability" of the climate system (Lines 44-46). I agree. I also agree with the authors’ definition of internal variability (Lines 45-51, although it can be shortened). However, in many occasions the authors seem to equate internal unforced variability with inter-annual variability, which add to the confusion (e.g., Lines 10-12, Lines 53-55). I would recommend to clearly define the two from the onset noting that internal variability is unforced whereas inter-annual variability can be externally forced by natural and anthropogenic aerosols, GHGs, solar radiation, land change use and so forth. If inter-annual variability is understood as derived from detrended time series, then explicitly say so from the onset, and clearly indicate how they are detrended.

2. The methodology presented in lines 140–150 is used to asses the interannual variability in the models against that in observations (results in Fig. 5-7). It should be clearly stated that this methodology is not an assessment of model internal (unforced) variability alone, since the time series are affected by the forced signal. Therefore, if there is no agreement between model and observations, we should not conclude that the model representation of internal variability is incorrect, as it may be consequence of the externally forced signal (e.g., the model may have a perfect representation of internal variability, but a too strong response to volcanic eruptions leading to disagreement in the anomaly distributions of Figs. 5-7). On the other hand, I would agree that if the observed and modelled distributions are coincident, this would suggest that both the model response to external forcing and its internal unforced variability are well represented. I don’t think this point is clearly made in the methods section and the discussions of sections 4.2 and 5. The way the methodology is presented and the results discussed seem as if the model response to external forcing and that from the observations are in perfect agreement.

3. Based on what is expected from the methodology introduced in lines 140–152, the distributions for the ensemble members in, say, Fig. 7 should largely coincide. They don't. In some cases they are quite different as noted by the authors. It is unclear then how to assess the agreement between model and observations based on these distributions. Are these differences because of a small sample size, or because the ensembles are not large enough? Could they be consequence of poorly sampled (multi)-decadal variability? Can the authors comment on this? I didn't quite follow the rationale of the last sentence of section 4.2, particularly the bit about added "information".

Minor:

In the title: Consider changing "variabilities" to "variability"

Line 11: "... (here: inter-annual variability) ...". Do you mean "on inter-annual timescales"? Inter-annual variability is affected by both externally forced and internally unforced variability. See comments above.

Lines 53-57: I don't think this is accurate and should be reworded. The ensemble spread about the mean can be used to measure the internal unforced variability of the model, but may not be representative of inter-annual variations in the presence of, say, a strong volcanic eruption. Therefore, using IMV to assess IAV may be a good approximation in some cases, but may also be in error. This should be clearly stated.

Line 105: I believe the work by Fyfe et al., 2017 uses the regional climate model CanESM2-CanRCM4 which is different from CanESM2-CRCM5.

Line 115: Although the authors provide a reference, I would find useful a brief comment on the weakness of the E-OBS dataset.

Line 120-124: Consider moving the text "These indicators (...) transport of rivers and many more", to the introduction and leave this section only for the methods.

Lines 129-139: The discussion on whether the indicators should be computed on the original grid to evaluate the averaged quantities, instead of regridding first and then evaluate the averages over a common grid, seems too long. The authors claim that both approaches give similar results and chose the former, which I believe is the recommended approach (Diaconescu et al., J. Hydrometeorol. 16, 2301–2310). The discussion could be shortened, and this reference cited.

Line 140–150: It would be helpful to have explicit references to Figs. 5-7 when discussing the method. Expressions like "we then plot" or "are plotted" are used without showing an actual plot.

Line 143: Change "neglected" by "avoided" or the like.

Line 165: The authors claim that the spread between members is a well suited metric to examine projected changes of inter-annual variability. I would not agree, unless it refers to detrended inter-annual variability (i.e., after removing the ensemble mean). If so, please clearly specify, although note that the previous sentence (line 163) would then be confusing because both IMV and IAV should be insensitive to external forcing, as no temporal autocorrelation is assumed (line 58).

Figure 3 and 4: I'm not sure how to interpret the changes in precipitation. It seems that the plots are intended to show changes in the climatological mean for each ensemble member over two different time periods. Why not showing just that? Precipitation is given here in %. Are these differences of relative values, or relative values of differences? Line 171 defines relative indices by dividing the ensemble anomalies about the mean by the ensemble mean. I can see how this may be useful to assess changes in ensemble spread, but this is not what is shown in Figs. 3 and 4, right? Please define clearly the quantities in these figures and clearly describe their behaviour.

Line 202: Remove “years 1957–2015 in" from the parentheses.

Line 207: “(slightly) too low" is odd.

Line 209: What does the authors mean by density functions “somewhat inflated"? Please clarify.

Line 305: Change to “an statistical model" instead?

Interactive comment on Earth Syst. Dynam. Discuss., https://doi.org/10.5194/esd-2019-67, 2019.