

Interactive comment on “Calibrating large-ensemble European climate projections using observational data” by Christopher H. O’Reilly et al.

Anonymous Referee #2

Received and published: 18 April 2020

Review of “Calibrating large-ensemble European climate projections using observational data” by O’Reilly, Befort and Weisheimer, submitted to Earth System Dynamics.

This is an ambitious and novel study aimed at improving climate projections using calibration techniques developed for initialized seasonal prediction. The approaches are tested on two single-model Large Ensembles (LE) using out-of-sample verification methods based on CMIP5 models. The analysis focuses on temperature and precipitation over Europe and takes into account seasonality. Another novel aspect is the application of the calibration method on the dynamical and residual thermodynamic components separately using the technique of “dynamical adjustment”. This yields an

C1

improvement in the accuracy of projections of temperature but not precipitation. The study is comprehensive and the methods are scientifically sound. The paper is generally well written, although some clarification is needed in places. I have a number of comments and suggestions as detailed below, but they are mostly minor in scope.

1) P2 L24: remove “was applied to” 2) P2 L32: Perhaps reference Deser et al. (2020) which provides a broader view of the utility of Large Ensembles with multiple models, and includes a more comprehensive listing of LE experiments to date.

Deser, C., F. Lehner, K. B. Rodgers, T. Ault, T. L. Delworth, P. N. DiNezio, A. Fiore, C. Frankignoul, J. C. Fyfe, D. E. Horton, J. E. Kay, R. Knutti, N. S. Lovenduski, J. Marotzke, K. A. McKinnon, S. Minobe, J. Randerson, J. A. Screen, I. R. Simpson and M. Ting, 2020: Insights from earth system model initial-condition large ensembles and future prospects. *Nat. Clim. Change*, doi: 10.1038/s41558-020-0731-2.

3) P3 L23: Suggest using “CESM1-LE” in place of “LENS” throughout for parallel construction with “MPI-GE”. 4) P3 L33: Please do some sensitivity tests on the choice of SLP dataset. I know that HadSLP2 generally has lower amplitude variability (and maybe trends) than 20CR or ERA20C. 5) P4 L19: “lies” should be “lie” 6) P4 L20: “is further” should be “are further” 7) P5 L17: “correlation” is mis-spelled and there is some missing text after “ensemble and, “ 8) P6 L4: “time” should be plural 9) P6 L24: Add “Guo et al., 2019” to your list of references (this was an application to precipitation)

Guo, R., C. Deser, L. Terray and F. Lehner, 2019: Human influence on winter precipitation trends (1921-2015) over North America and Eurasia revealed by dynamical adjustment. *Geophys. Res. Lett.*, 46, doi: 10.1029/2018GL081316.

10) P7 L8: “is clearly has a” is not grammatical 11) P8 L19: This sentence is confusing because it sounds like you are only testing the methods on the MPI-GE, but that is not the case. I suggest first discussing the LENS results and then moving on to the MPI results. 12) P9 L2: is the lack of improvement in winter because the characteristics of the variability are not distinguishable between LENS and CMIP5? 13) P9 L 3: “are”

C2

should be “is” 14) P9 L5: “larger than is appropriate”: please explain what you mean. Does this imply that LENS has more variability than the other CMIP5 models, or a larger forced signal? Relatedly, it would be very nice to see some discussion of the relevance of the so-called “signal-to-noise paradox” in the seasonal-to-interannual prediction literature for climate change projections. 15) P9 L10: “in to” should be “is to” 16) P9 L18: Change “covarying signal in the reference/observational index” to “covarying signal between the reference and observational indices” for clarity (unless I misunderstand your approach). 17) P9 L21: “with a circulation driven signal”: do you want to specify whether this can be an “internal” circulation driven signal, or forced, or both? 18) P10 L20: “separately” is mis-spelled 19) P10 L21: “in the ensemble with a signal”: please clarify your intended meaning; the language is confusing. 20) P11 L1: “of temperature.”: I would add “in both seasons and models, but especially summer”. 21) P11 L27: “from the all of” ? 22) P11 L30-31: Can you provide a physical explanation for why the calibration method acts to increase the uncertainty in future projections? Does it have to do with differences between the level of variability between observations and the model? 23) P12 L9: Is the reduced drying mainly dynamical or thermodynamic in origin? 24) P12 L15 “far more consistent . . .”: I think this is an overstatement. 25) P12 L22-30: How do your results relate, if at all, to the trend biases in LENS compared to a synthetic observational Large Ensemble documented in McKinnon and Deser (2018)? McKinnon, K. A and C. Deser, 2018: Internal variability and regional climate trends in an Observational Large Ensemble. *J. Climate*, 31, 6783–6802, doi: 10.1175/JCLI-D-17-0901.1.

26) P13 L31: suggest adding “in the calibrated ensembles” after “generally smaller” 27) P14 L7: “For precipitation, where there is no clear signal over the reference period in the observations”: I am not sure what your evidence is. Guo et al. (2019) found a nice correspondence with dynamically-adjusted precipitation trends from observations and the ensemble-means of LENS and CMIP5 models. 28) P14 L12: add “relative to the internal variability” after “weaker” (i.e., the forced signal doesn’t weaken on smaller

C3

scales, just the signal-to-noise weakens). 29) P14 L21: “is kept” should be “are kept” 30) P14 L27: Cite Yeager et al. (2018) for the LENS DPLE.

Yeager, S. G., G. Danabasoglu, N. Rosenbloom, W. Strand, S. Bates, G. Meehl, A. Karspeck, K. Lindsay, M. C. Long, H. Teng, and N. S. Lovenduski, 2018: Predicting near-term changes in the Earth System: A large ensemble of initialized decadal prediction simulations using the Community Earth System Model, *Bull. Amer. Meteor. Soc.*, in press, doi: 10.1175/BAMS-D-17-0098.1.

31) P14 L28: “merged calibrated climate predictions”: insert “set of” before “climate predictions”? 32) Caption to Fig. 3: add “summer” before “temperature” 33) Title to Fig. 4: It is confusing. Suggest re-wording as: “LENS JJA Temperature” (analogous comment applies to Fig. S1). 34) Title to Fig. 5: omit the dash after “LENS” for clarity 35) Caption to Fig. 5: 2nd sentence: change “Shown” to “Results are shown . . .” . Also, the sentence describing what the black boxes mean is confusing. I would shorten to: “Black boxes indicate where the HGR-decomp method of calibration is significantly better than the HGR method (at the 90% level).” 36) Caption to Fig. 7, line 3: change “has a” to “is”. In the next line, change “worse that” to “worse than”. 37) Caption to Fig. 8: Please state what the various colors and linestyles mean, and what the shading means. Don’t rely on the legend. Indeed, the colors/linestyles in the legend seems to be at odds with that shown in Fig. 7, which had all blue for LENS and all red for MPI. Please make them consistent for clarity. 38) Fig. 9: Same comment as above: please use a consistent color scheme as in Fig. 7 (or change Fig. 7 to be consistent with Fig. 9). 39) Caption to Fig. 9: Please state the method of calibration in the caption. Is it HGR-decomp?

Interactive comment on Earth Syst. Dynam. Discuss., <https://doi.org/10.5194/esd-2020-6>, 2020.

C4