Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2017-162-AC2, 2017

Geoscientific Model Development Discussions

Interactive comment on "Parametric Decadal Climate Forecast Recalibration (DeFoReSt 1.0)" *by* Alexander Pasternack et al.

Alexander Pasternack et al.

alexander.pasternack@met.fu-berlin.de

Received and published: 13 October 2017

Answer to referee 2

Thank you very much for your informative and detailed comments.

General comments

"This paper presents what may be a very important work in the study of decadal climate prediction. The authors presented the methodologically oriented post-processing

model "Parametric Decadal Climate Forecast Recalibration (DeFoReSt)," to correct decadal climate prediction. The method uses earlier published approaches, and extend it to the long-term prediction by allowing the forecast errors to depend on forecast lead time. The performance of the presented approach is well established using different assessment measures."

Specific comments

1. "It is mentioned in section 3.3 on lines 1-2 that the parameters are estimated by minimizing the average CRPS over the training period. Does this mean only the portion (the training) of the data was used for estimating the parameters? if yes, why the whole data was not used? it is expected that a training-validation grouping of data for checking the performance of DeFoReSt. But a parameter within DeFoReSt has to be primarily estimated from the whole data."

   Answer: In case of a validation/comparison with a reference data set (e.g., climatology or raw model) the training data set is only a portion of the whole available data set, while the remaining data is used for validation. We aim at estimating a forecast error for a setting comparable to the operational forecast situation where no observations for the forecast period is available. DeFoReSt parameters can only be estimated from the available observataion period but the re-calibration is carried out on the forecast period, i.e. outside the period used for parameter estimation. Hence using the full hindcast period for estimating parameters and obtain a "forecast error" for hindcast from the same period would lead to overestimation of skill. Parameter estimation using the full available data set could be used once we use DeFoReSt for re-calibration decadal forecasts, for e.g., 2018-2027.

2. "DeFoReSt was defined based on ensemble mean and variance functions (by my reading on the paper), where 15 ensemble members were selected. However,

sys- tematic errors vary widely between ensemble members, a simple ensemble averaging limits the relevance of DeFoReSt for long-term prediction. Arisido et al (2017) and Tebaldi et al (2005), see below, demonstrated that the common ensemble averaging method where each ensemble member has the same weight poses serious issues. I encourage the authors to discuss this issue supporting their argument with these pa- pers and take into account the advisories in their revision.

- Arisido, M.W., Gaetan, C., Zanchettin, D. et al. Stoch Environ Res Risk Assess (2017). https://doi.org/10.1007/s00477-017-1383-2 - Tebaldi C, Smith RL, Nychka D, Mearns LO (2005) Quantifying uncertainty in projec- tions of regional climate change: A Bayesian approach to the analysis of multimodel ensembles. Journal of Climate 18:1524-1540"

Answer: It is true that the ensemble members of a multi-model ensemble cannot be treated equally because every corresponding model has different systematic errors. However, in this study we apply DeFoReSt to a single model ensemble with 15 members generated by lagged-day-initialization from MPI-ESM-LR; i.e., we do not expect that the single ensemble members have different systematic errors (due to the model). Nonetheless, for a recalibration of a multimodel ensemble DeFoReSt needs to be adapted. Which would be a topic of further research.

3. "How are the orders chosen for the polynomials used in equations such as (9) and (10). Without some cross-validation study it is not clear how a third order polynomial suffices for the drift along lead time. A sensitivity analysis for different order scenar- ios should guide for closer to the optimal choice needed for capturing the underlying features in a data."

   Answer: We agree that there is need for a transparent model selection strategy! As already mentioned in section 6, this will be topic for future studies. For the first version of DeFoReSt we follow the suggestion of Gangstø et al. (2013) and use

a third order polynomial addressing the unconditional and conditional bias and a second order addressing the ensemble dispersion. With an increasing order of the polynomial the flexibility of the fitted curve increases, while the parameter uncertainty also increases. Here, Gangstø et al. (2013) suggested that a third order polynomial is a good compromise between flexibility and parameter uncertainty. For the correction of the ensemble dispersion we assumed that a higher flexibility may not be necessary, because the MSE -which influences the dispersion- is addressed by a third order polynomial of unconditional and conditional bias.

-Gangstø, R., A.P. Weigel, M.A. Liniger, C. Appenzeller, 2013: Methodological aspects of the validation of decadal predictions. – Climate Res. 55, 181–200, DOI: 10.3354/cr01135.

Technical corrections

1. "line 19 page 2, the acronym "PDF" should be defined on the first use"

   Answer: Will be corrected

2. "line 1 Page 15, "..to change polynomially.." Remove "polynomially", then it is clear that the conditional bias and the ensemble dispersion change with lead time, while they change linearly with start time."

   Answer: Will be corrected

3. "υ_5 in (A9) page 17, the coefficient of τ^2, is typos error?"

   Answer: Indeed, it should be b_5 instead of υ_5.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-162, 2017.