Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2020-275-RC2, 2021

Geoscientific Model Development Discussions

Interactive comment on "Latent Linear Adjustment Autoencoders v1.0: A novel method for estimating and emulating dynamic precipitation at high resolution" *by* Christina Heinze-Deml et al.

Anonymous Referee #2

Received and published: 19 February 2021

This manuscript describes a novel and potentially very useful machine learning technique for separating forced signals and internal variability from climate model output (as well as other applications such as weather generation). I should say at the outset of this review that I am an expert on climate, not on machine learning, so I cannot comment on detail on the machine learning method in this study. From what I can determine though, the methodology looks mainly sound and produces sensible results. The manuscript is very well written and logically presented. I am recommending major revisions because I would like to see a sensitivity test on the time period of the training data, but other than that my comments are mainly relatively minor.

Major points:

1) The training (1955-2070) and testing (2071-2100) periods are consecutive, which I do not think is the best choice, as the training data is likely to contain a forced precipitation trend. Separating the training and testing datasets (e.g. training 1955-2020) would provide a more rigorous test of whether the dynamical adjustment method can separate internal variability from a forced signal, without much of the forced signal being present in the training dataset. The authors should test at least some of their results for sensitivity to the choice of training period.

2) I am not convinced that the forced signal that is extracted using the dynamical adjustment method is a purely thermodynamic signal of precipitation change, for two reasons. Firstly, the residual trend will include not only Clausius-Clapeyron-related increases in moisture, but also any other change in the relationship between SLP and precipitation under climate change. This could include, for example, changes in land-atmosphere interactions or weather system dynamics.

Secondly, there may be changes in the pattern of the individual SLP EOFs under climate change. Even small changes could have large consequences for regional precipitation. The authors have tried to address this point by detrending the SLP time-series, based on trends in EOF1, but I was slightly confused by the description of this detrending, and am not convinced that it would account for any (possibly subtle) changes in the shape of EOFs.

I think this is mainly a question of interpretation. The dynamical adjustment method will (as I understand it) remove any signal caused by temporal variation in the frequency of the SLP EOFs that were identified during the training period. The removed component will likely be due mainly to internal variability, though it could also include some forced signal if forcing were to drive any systematic change in the relative frequency of SLP EOFs. The residual will likely be a forced signal but I think calling it a thermodynamic precipitation change is too much of an oversimplification to be useful. Other factors could also be important.

3) Dynamical adjustment appears to have the potential to significantly reduce the size of ensembles needed to reliably extract forced trends. However, a certain number of model years are needed to train the algorithm, so it is not clear exactly what the computational cost saving would be overall. Could the authors provide an estimate of the overall fractional saving in computational cost, taking algorithm training into account?

4) Is there an alternative type of machine learning algorithm that could be used to link SLP EOFs as input directly to the 2D precipitation fields as output (e.g. some form of neural network)? What are the benefits of using the intermediate stage of the autoencoder? I am not suggesting any extra analysis here, only for the authors to justify their choice of method a bit more.

Minor points: 1) It is not clear from the objectives in section 1 that the dynamical adjustment will be used to separate forced precipitation trends from internal variability. It would be useful to the reader for this objective to be spelled out here.

2) Fig. 3: How were these examples chosen? Are they representative of the data as a whole? It might be more useful to show high, medium and low skill cases rather than a random selection.

3) Figure colour scales. It is quite difficult to get much information out of the current single shading colour scales. I appreciate this is not a simple problem, but perhaps these could be improved to show the spatial features more clearly.

4) Fig. 4 & 5: Why only use a single holdout ensemble member for this? Why not use all of them? Also, relative error might be more informative for Fig. 4, rather than absolute error which mainly picks out the regions of high precipitation.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2020-275, 2020.