Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2020-275-RC1, 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 #1

Received and published: 6 January 2021

This article is a novel and very interesting piece of work, with potential for applications in the field of climate science, some of which are presented in the article. The authors describe a latent adjustment autoencoder modified with the addition of a linear component between the input sea level pressure Empirical Orthogonal Function timeseries and the latent space of the autocoder. In a first application, the authors show that this allows to remove the internal variability of winter precipitation over Europe and extract the thermodynamical forced signal from only a few members of simulations instead of averaging a very large number of simulations. A second application they show is a weather generator, based on bootstrapping the SLP EOFs and then decoding the precipitation fields. All applications are limited to generating present-day like winter

precipitation patterns since the only input variable is sea level pressure. It may limit the application of the method to other seasons where precipitation may be less tied to SLP patterns. Although I am not a specialist in machine learning, I found the method well-explained and I had a look at the code which also seems well-explained and portable for the use by others. However, I would like some additional discussion in the text and a few changes to the figures before it can be accepted.

1 Major comments:

On the method itself:

1. I don't understand why you train your autocoder on 1955-2070 data. There is a chance that you include some thermodynamical signal in the precipitation field when you minimise $Y - \hat{Y}_X$. I understand that you detrend the SLP EOF time series, but you don't detrend precipitation. Why not training on 1955-1995 and potentially use more members to have the same amount of data?
2. It would be good to know the minimal amount of data needed to train the algorithm. Indeed, if 1955-2070 daily data from a 9 member ensemble is needed to train the algorithm, then it would be cheaper to directly calculate the forced response from this 9-member ensemble (see comments below on Fig. 8) without dynamical adjustment. Ideally, one would like to dynamically adjust expensive simulations which cannot be run for long periods of time (e.g. a few decades).

2 On the examples of application:

1. I am convinced by the use of the new tool for dynamical adjustment on a large domain and seasonal scale (Fig. 6), this seems to be very successful, even

with only 1 member. This is quite impressive. For more detailed spatial scales however, it is less successful and I guess from extrapolating Fig. 8 that using 7 or 8 members for the “traditional runs” (out of 50) outperforms the dynamical adjustment. I would like to see more discussion on this in the text and I think that Fig. 8 could be improved with a few changes:

- extend the x axis to at least 10 members, to see when a “traditional averaging” outperforms the dynamical adjustment (this implies performing dynamical adjustement on more holdout members).
- you plot only one value and it does not correspond to the one in the text (line 255), I presume for 1 member you can have 41 different values (excluding the training set), so you can add median + inter-quartile range / sqrt(number of samples), so that one knows if the difference is statistically significant, but I presume so.
- Add details to the caption. I presume that it shows the RMSE of 50y trend maps calculated by averaging n members compared to 50y trends using 50 member average. It is not very clear from the caption.

2. The tool is successful for seasonal means. Can you comment on the potential use of this tool for assessing trends in extreme precipitation, for which regional models are more trustworthy than global models? The prediction in precipitation fields seems smoothed out compared to original fields. And not taking into account thermodynamical fields as predictors may be limiting the representation of extremes, even in a present-day context.

3. Regarding the weather generator, I do struggle to exactly understand the novelty of your method. If I understand correctly, you are bootstrapping the time series of EOFS, but keeping each daily EOF set as it is, so you are not “creating” new pressure patterns, just shuffling them. One could do this directly by shuffling daily precipitation maps in the same way. I agree that one would need 150 years of

present-day data instead of simulations with evolving greenhouse gases, but this is easily achieved these days. It is interesting that you show that shuffling 150 years of data seems as good as running several members, at least for the bulk of precipitation distribution. I wouldn’t think this is true for extremes. I think the use for dynamical adjustment has much more potential than the weather generator.

   I would suggest to reduce this section to have more space in the article for a figure to reply to my point 2 about the method.

3 Minor comments:

Fig. 7, 10, 11: the scatter plots are saturated, it may be better to plot a gaussian kernel density estimation https://seaborn.pydata.org/generated/seaborn.kdeplot.html

Most figures with blue shading only: I find the continuous colour shading difficult, it may be best to reduce the number of colour levels used. One could also potentially use a sequential colour map like terrain_r for precipitation fields. It will make figures more readable and may reduce the need to show square root precipitation fields, which are less intuitive.

Fig. 9: remove the numbers on it, you are not using them in the article.

"As is to be expected, the emulated predictions based on the individual spatial fields are not visually distinguishable from the original predictions." Do you mean that they look “physical” with no artefacts? They are not meant to be similar to the original predictions. This is just like Fig. 3, I don’t really see the point of this figure.

Fig. 12: caption could be a bit more wordy to be self-explanatory if readers only partially read the article.

Typos: Line 17: are expected to remains in the Line 176: Fig. 7a -> Fig. 9a

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