Method - EOF analysis

Annual average climatology for surface air and ocean temperatures were available for 9 PMIP2 models. In PMIP3, we have the same variables for 8 models. Here we combined the two ensembles to make a total of 17 models. We perform an uncentred analysis in order that the first EOF represents the ensemble mean signal (not precisely, but quite closely).

The first 6 EOFs for this combined PMIP2+PMIP3 ensemble are shown in Figure 1. Are they just statistics, or do they have any physical meaning? Let me know what you think! 🙏

The method is tested using a leave-one-out analysis. This entails using the EOFs from 16 models; to reconstruct the climate of the 17th model, using pseudo data taken from the 17th model at the locations of the real data. The improvement of the fit as more EOFs are included in the ensemble is illustrated in Figure 2. It is clear that the goodness of fit levels off fairly rapidly, suggesting that there is little benefit from including the higher EOFs if observational error is set to a high level, the fit actually degrades with more than about 4 EOFs indicating over-fitting, but this did not happen with realistic error levels.

Some of the PMIP3 models are later versions of the models used in PMIP2, so may have similarities, which could result in some redundancy and a poor set of EOFs. This will be explored more rigorously in later work, but if there was clear redundancy we might expect to see very low error achieved when all 16 EOFs are used in the leave-one-out analysis and this does not appear to be the case.

Figure 3: Thin lines show how the RMS errors change as more EOFs are used in the reconstruction of each model (using the other 16). Thick black line is the average.

Results - Reconstruction of Last Glacial Maximum

The method used in AH13 combined all 9 models available at that time. Methodologically this would be equivalent to using all 17 EOFs in this work, but it seems that would be trespassing well into the area of over-fitting due to internal variability in the model outputs. In particular with the somewhat sparse datasets there is the danger that a slightly better fit at some of the data points, can be obtained by including a large scaling of some of the later EOFs which have unrealistically large changes in data voids.

It is not clear exactly how many EOFs to include in our analyses. Figure 3 suggests that little may be gained by using more than about 4 EOFs. In Figure 4 we show results using 4 and 9 EOFs. Which do you think is more realistic? See also Figure 1 which shows the result from AH13s and the simple ensemble mean, for comparison.

Reconstruction of the mid-Pliocene warm period

In principle, a similar approach may be taken for the mid-Pliocene warm period, modelled in PMIP1 and PMIP2. We have TAS and TOS for 8 model from PMIP1 and 4 are presently available for PMIP2 on EOF. Data are less plentiful and less reliable than for the LGM, so it is inevitable that a less confident reconstruction will be obtained. This work is underway. The approach of AH13 does yield heuristic uncertainty measurements (not shown here). Future work will introduce an explicitly Bayesian framework for the analysis. This may also be helpful for generating a reconstruction that realistically represents our understanding of the climatology of the mid-Pliocene.

Figure 4: Reconstructions of annual mean TAS and TOS using 1, 4 or 9 EOFs of the combined PMIP2+PMIP3 ensemble, also with differences compared to the PMIP2+PMIP3 ensemble mean (ensemble mean shown in Figure 1 c,d).

Conclusion

We use multiple linear regression of the EOFs of the ensemble rather than the ensemble members themselves to reconstruct the LGM climate. Around 4 EOFs adds considerable spatial information to the ensemble mean, while hopefully not overfitting noise. The previous result of AH13 is possibly slightly noisy, but perhaps not as bad as we had feared it might be!

References

https://www.theguardian.com/lifeandstyle/wordofmouth/2014/may/28/how-to-make-the-perfect-wiener-schnitzel