Project: **888**

Project title: **Assimilation of paleoclimate proxy data into MPI-ESM using ensemble member selection**

Project lead: **Johann Jungclaus**

Report period: **1.1.2014 - 31.12.2015**

Data assimilation (DA) is an emerging research area and can be considered as one of the key challenges in palaeoclimatology. Different ensemble-based DA approaches for palaeoclimate reconstructions have been recently undertaken, but no systematic comparison among them has been attempted and large understanding gaps of the ideal characteristics and input data that would give the best state estimates remain. In this project, we are among the first to perform DA using a GCM, specifically the MPI Earth System Model (MPI-ESM), based on ensemble member selection for specific periods of the last millennium. ECHAM6 was run at T31 horizontal resolution with 31 vertical levels, and MPIOM at a GR30 horizontal resolution and 40 vertical levels. The land surface model was JSBACH and no ocean biogeochemistry model was employed. The model is a low-resolution version of the model used for the Coupled Model Intercomparison Project (CMIP) Phase 5 simulations.

As a first step in the study, we compared an off-line and an online ensemble-based method, with the testing period being the 17th century, which led into the Maunder Minimum. For this purpose, we employed 10 ensemble members for each DA scheme and assimilated the PAGES 2K Northern Hemisphere (NH) continental temperature reconstructions. In the off-line approach, the ensemble for the entire simulation period was generated first and then the ensemble was used in combination with the empirical information to produce the analysis. In contrast, in the on-line approach, the ensembles were generated sequentially for sub-periods based on the analysis of previous sub-periods. We found that both schemes performed better than the simulations without DA. The on-line method would be expected to perform better if the assimilation led to states of the slow components of the climate system that are close to reality and the system had sufficient memory to propagate this information forward in time. In our comparison, which was based on analysing correlations and differences between the analysis and the proxy-based reconstructions, we found similar skill for both methods on the continental and hemispheric scales (e.g. figure 1). This indicates either a lack of control of the slow components in our setup or a lack of skill in the information propagation on decadal timescales. Although the performance of the two schemes is similar and the on-line method is more difficult to implement, the temporal consistency of the analysis in the on-line method makes it in general preferable.

After choosing the on-line DA scheme as the most appropriate, 20 ensemble members were run employing this setup for the reconstruction of the climate for 1750–1850 AD. The performance of the assimilation was evaluated on large and small spatial scales. We found good skill for large-scale temperatures, but there was no agreement between the DA analysis and the proxy-based reconstructions for small-scale temperature patterns within Europe or with reconstructions for the North Atlantic Oscillation (NAO) index. To explain the lack of added value in small spatial scales, a maximum covariance analysis (MCA) of links between NH temperature and sea level pressure was performed based on a control simulation with MPI-ESM. For annual values, winter and spring the Northern Annular Mode (NAM) is the pattern that is most closely linked to the NH continental temperatures, while for summer and autumn it is a wave-like pattern. This link was reproduced in the DA for winter, spring and annual means, providing potential for constraining the NAM/NAO phase and in turn regional temperature variability. It was shown that the lack of actual small-scale skill is likely due to the fact that the link might be too weak, as the NH continental mean temperatures are not the best predictors for large-scale circulation anomalies, or that the PAGES 2K temperatures include noise. Both factors can lead to circulation anomalies in the DA analysis that are substantially different from reality, leading to unrealistic representation of small-scale temperature variability. Moreover, we show that even if the true amplitudes of the leading MCA circulation patterns were known, there is still a large amount of unexplained local temperature variance. Based on these results, we argue that assimilating temperature reconstructions with a higher spatial resolution might improve the DA performance.

Currently, we analyse simulations performed using the on-line DA scheme for the period 1850-1949 AD. Two sets of DA simulations are conducted, one selecting the best among 20 ensemble members based upon comparison with the PAGES 2K proxy-based reconstructions, and the second by selecting ensemble members after comparison with the HadCRUT3v instrumental record. The simulations for the instrumental period allow validating the assimilation method against instrumental records, which would be a substantial contribution to the research area of palaeoclimatic modelling and thus an essential part of the study.

The above work has led to two recent publications:

Matsikaris, A., Widmann, M., and Jungclaus, J.: On-line and off-line data assimilation in palaeoclimatology: a case study, Climate of the Past, 11, 81, doi:10.5194/cp-11-81-2015, 2015a.

Matsikaris, A., Widmann, M., and Jungclaus, J.: Assimilating continental mean temperatures to reconstruct the climate of the late pre-industrial period, Climate Dynamics, published online, doi:10.1007/s00382-015-2785-9, 2015b.

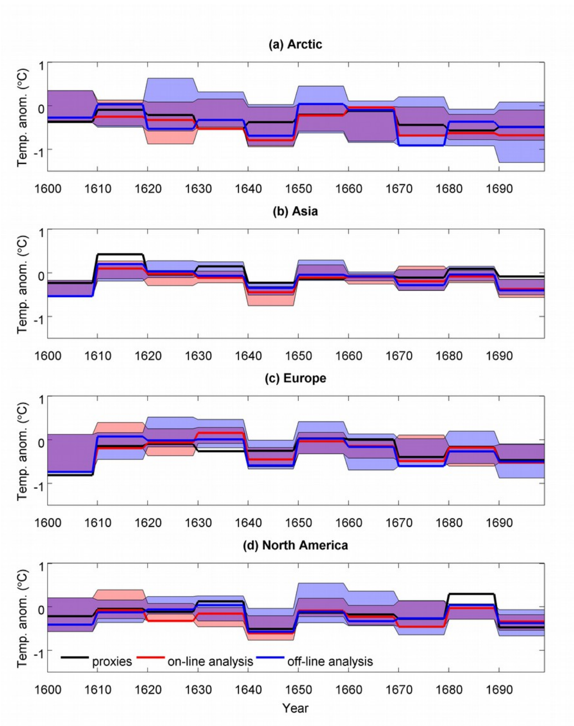


**Figure 1.** Continental decadal mean temperature anomalies w.r.t. the 850–1850 AD mean in the Northern Hemisphere for the 17th century, for the on-line (red shading) and off-line (blue shading) ensemble members, the on-line (red line) and off-line DA analysis (blue line), and the proxy-based reconstructions (black line).