



## **How informative are SST proxy data in paleoceanographic inverse modeling? - Insights from comprehensive uncertainty quantification**

Nora Loose (1,2), Patrick Heimbach (3,4), Kerim H. Nisancioglu (1,2)

(1) Department of Earth Science, University of Bergen, Bergen, Norway, (2) Bjerknes Centre for Climate Research, Bergen, Norway, (3) Institute for Computational Engineering and Sciences and Jackson School of Geosciences, The University of Texas at Austin, Austin, USA, (4) Department for Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, USA

Solving a large-scale nonlinear inverse problem by means of the adjoint method is a well-established data assimilation technique in modern oceanography, but has also won popularity in paleoceanography over the past few years. Uncertain model parameters, boundary and initial conditions are adjusted such as to minimize the discrepancy between an ocean general circulation model (OGCM) and observations. The obtained solution has the desirable property to be dynamically consistent without unphysical sources or sinks. However, it remains unclear whether sparse and uncertain paleo proxy data can provide useful constraints for the three-dimensional ocean circulation, i.e., whether they can inform the  $O(10^5)$  to  $O(10^7)$  unknowns to be estimated in an OGCM. Here we quantify how informative sea surface temperature (SST) proxy data are for OGCM diagnostics of interest, such as meridional heat transports at various latitudes.

To characterize and quantify uncertainties, we embed the deterministic inverse problem in a Bayesian context, where the inverse Hessian of the regularized data-model misfit cost function becomes the error covariance matrix of the Gaussianized posterior distribution. While calculating and inverting the high-dimensional Hessian matrix is computationally intractable, we exploit the fact that for sparse observations the Hessian is of very low rank, and extract the eigenvectors of the Hessian by means of the adjoint of the MIT general circulation model. Furthermore, we disentangle how the quality of observational constraints is determined by the interplay of two aspects: first, observability, which is the property of how well the ocean state or diagnostic of interest can be inferred from knowledge of the observations in the zero-noise limit; and second, the signal-to-noise ratio, which depends crucially on the magnitude of the data uncertainties. We find that SST proxy data cannot constrain diagnostics that depend on the three-dimensional ocean circulation, such as the Atlantic meridional overturning circulation. As a further step, we assess whether the issue is a matter of observability or signal-to-noise ratio, i.e., the spatial incompleteness of the data or its large uncertainties. We conclude that the main issue lies within the large proxy uncertainties, although there is a component that cannot be inferred from SST observations, even with noise-free data. While this work is limited by the Gaussian assumption and the examination of solely SST proxy data, it is a first step to exploit powerful tools of computational science for conducting rigorous uncertainty quantification in large-scale paleoceanographic inverse modeling.