A design of the optimal setup for coupled data assimilation for decadal climate predictions

Yulia Polkova

in collaboration with
Armin Köhl, Guokun Lyu, Detlef Stammer, Silke Schubert, Frank Lunkeit

Center for Earth System Research and Sustainability (CEN)
Institute of Oceanography, Universität Hamburg

iuliia.polkova@uni-hamburg.de

Data assimilation for decadal climate predictions

→ **From potential predictability studies:**

… room for improvement from advancing initialization/data assimilation (Boer et al 2013)

… but we are not sure if potential predictability really indicates the upper-limit

→ **From decadal prediction studies:**

… native and coupled data assimilation systems might lead to better predictions (Polkova et 2019a)

… initialization shocks, drifts and biases are still an elephant in the room (Meehl et al 2021)

… dynamically-consistent initialization mitigates shocks (but effect is only seen on short timescales: zonal momentum balance; Liu et al 2016, Polkova et al 2019b)

… effect of dynamical-consistency on long timescales is not clear … unless we do smth very wrong during initialization, the effect is not that obvious (Pohlmann et al 2017, Kröger et al 2018)

… no clear evidence from investing in sophisticated data assimilation for decadal predictions … old-timer nudging is still a runner (Meehl et al 2014, Meehl et al 2021)

… a matter of resources available (Marotzke et al 2016)

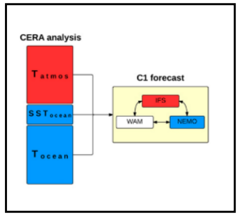
Origin of initialization shocks

Mulholland et al 2015: an effect of using non-native un-coupled reanalyses on initialization shocks (in the form of spurious waves, transports, flux exchange) **in weather forecasts:**

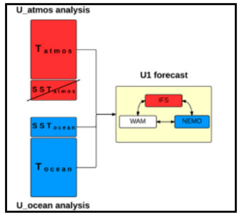
- Dynamical imbalances between **initial state and prediction system**
- Dynamical imbalances between initial state (ocean) and initial state (atmosphere)

can lead to initial shocks and eventually to loss of prediction skill

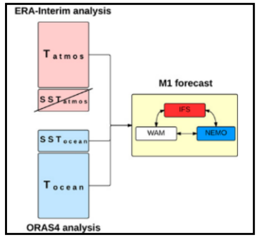
I. coupled **native**



II. un-coupled **native**



III. un-coupled **non-native**



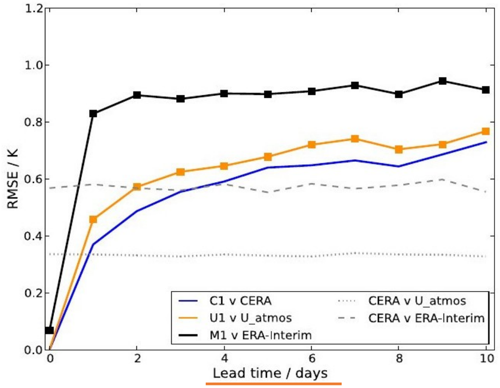


FIG. 5. 1000-hPa temperature forecast RMSE averaged over the Niño-3 region for C1 (blue), U1 (orange), and M1 (black) each evaluated against their own corresponding analysis, as labeled. RMSD between CERA and the other two analyses are shown for comparison (gray dashed and gray dotted). Squares mark where points in the U1 and M1 series are different from C1 at the 90% significance level, using confidence intervals calculated via the bootstrapping method.

Figures from Mulholland et al 2015
https://doi.org/10.1175/MWR-D-15-0076.1

Objective of the project

→ Investigate an effect of using **native strongly-coupled data assimilation** on mitigating initialization shocks **in decadal climate predictions**

→ … given that more evidence is needed before investing in expensive and resources-demanding data assimilation systems

Earth System Assimilation Model

4D-VAR for ocean and atmosphere (strongly coupled: Stammer et al 2018)

Long assimilation window (regularization: Lyu et al 2018, Abarbanel et al 2010)

Intermediate complexity

low resolution, simplified processes for sea ice and
land processes, hydrostatic approx., radiative transfer

Components of CESAM

Ocean general circulation model MITgcm

Atmosphere spectral general circulation model PlaSim

Land surface and soil parameterizations

Thermodynamic **sea ice** model

Resolution

T21(5.6°)L10 (T42(2.8°)L10) in the atmosphere

4°L15 (1°L23) in the ocean

Computing time 2 minutes per 1 model-year (monthly output)

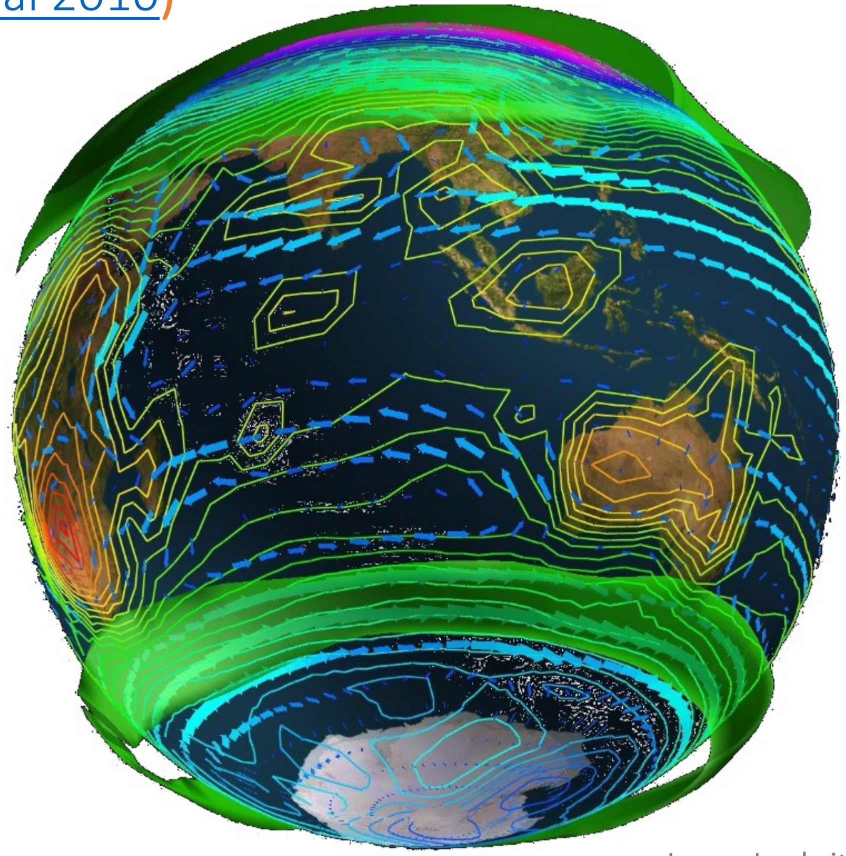


Image Lunkeit

Source code www.cen.uni-hamburg.de/research/cen-models/cesam.html

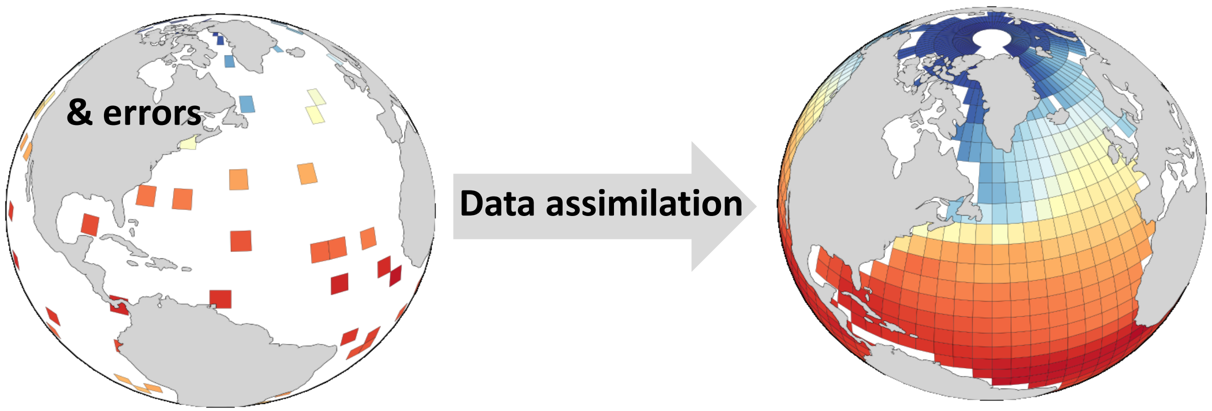
Experiments

Simulation	Time	Ocean/Atmos Data	Forcing Data
piControl	1 member, 300 years	WOA climatology	*const* CO2 & solar forcing from 1850
Historical ensemble	10 members, 1850-2018		time-varying CO2 & solar forcing
Atmospheric nudging	1 member 1980-2018	full field daily ERA5	
Coupled nudging **& prediction ensemble*	1 member 1980-2018	full field daily ERA5 & monthly GECCO3	
	**10 members 10-year long start every year*		
4DVAR assimilation **& prediction ensemble*	1 member 1980-**2018*	full field daily and 10-day averages ERA5 & monthly EN4, HadISST and WOA SSS	
	**10 members 10-year long start every year*		

** in-progress and planned*

iuliia.polkova@uni-hamburg.de

With the help of 4D-VAR assimilation obtain full picture about initial state of the climate system



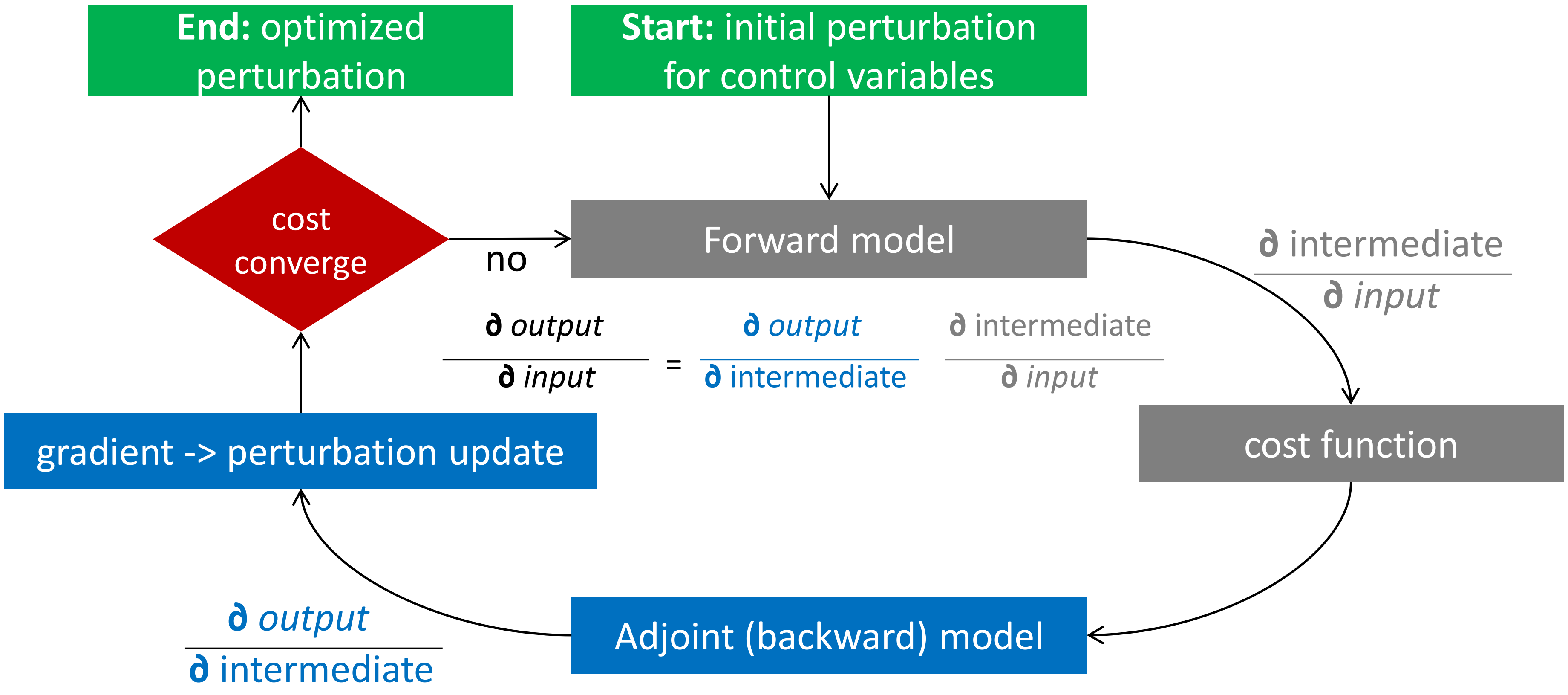
M(input)=output

1. M(observations)=optimized climate state
2. M(optimized climate state)=coupled reanalysis

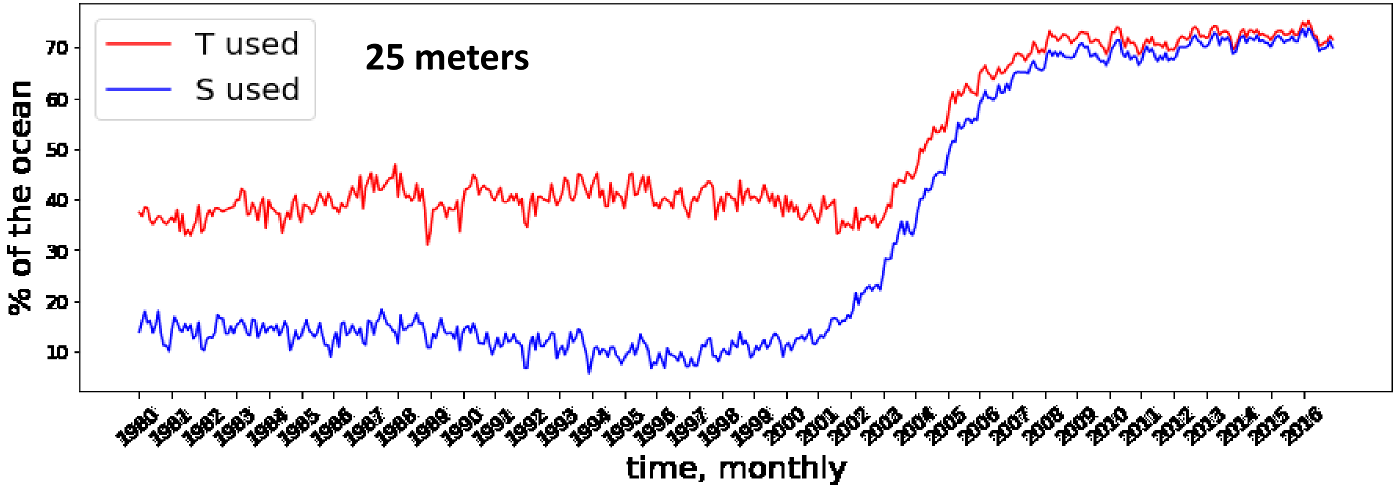
4D-VAR assimilation for obtaining optimal initial state

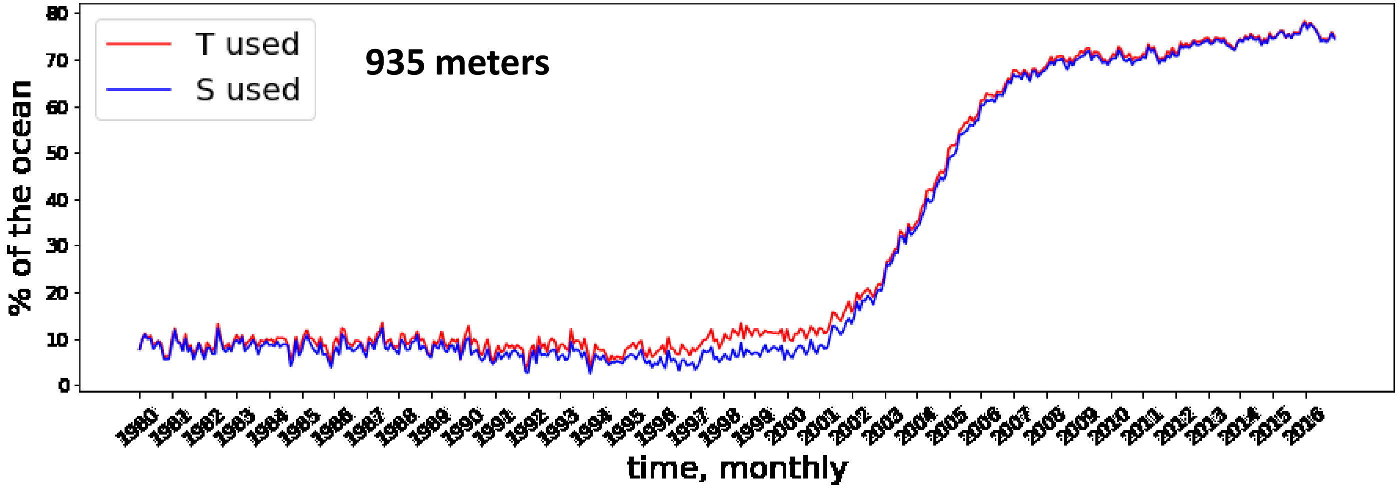
M(input)=output

1. M(observations)=optimized climate state
2. M(optimized climate state)= coupled reanalysis



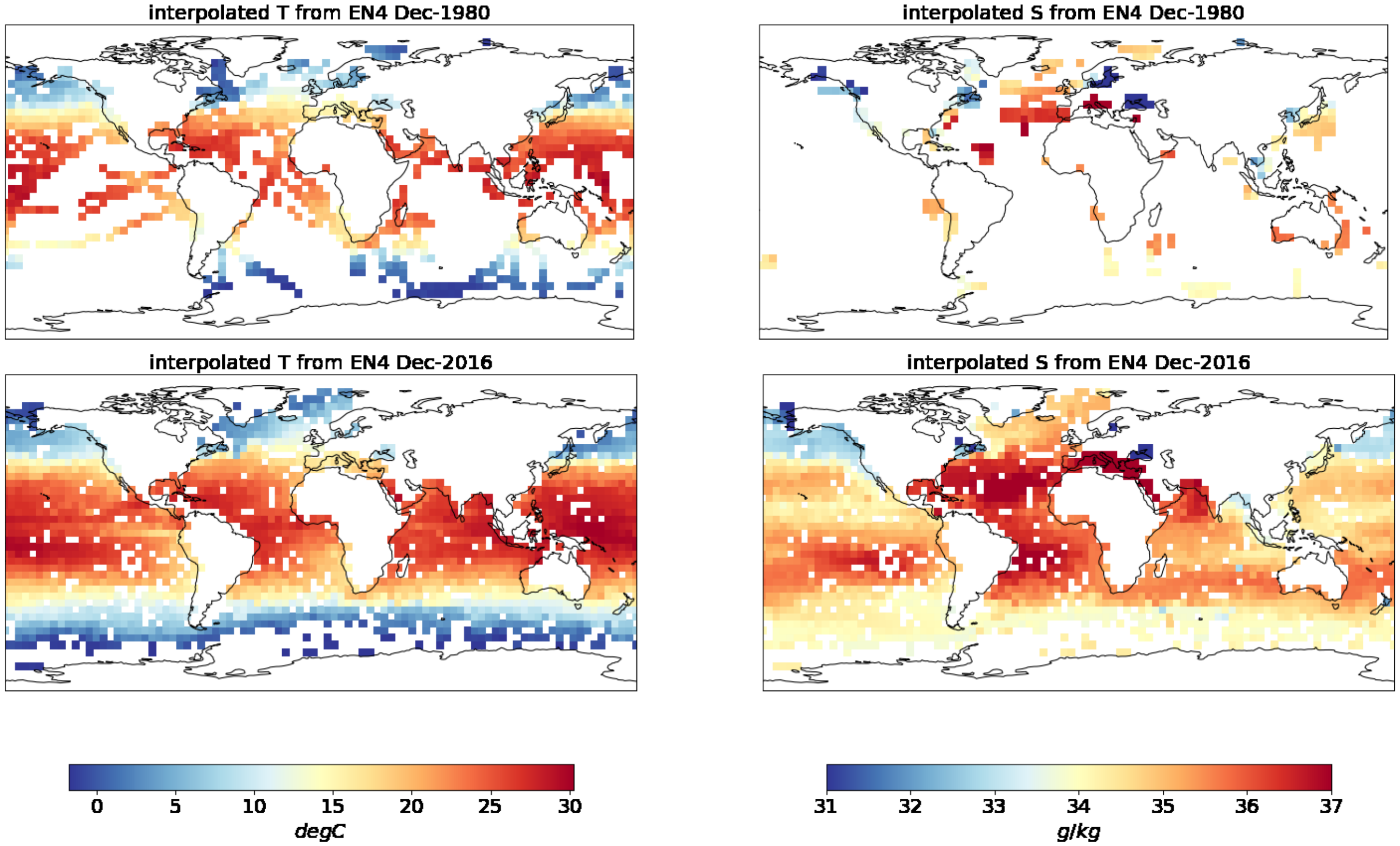
Ocean EN4 data used for assimilation





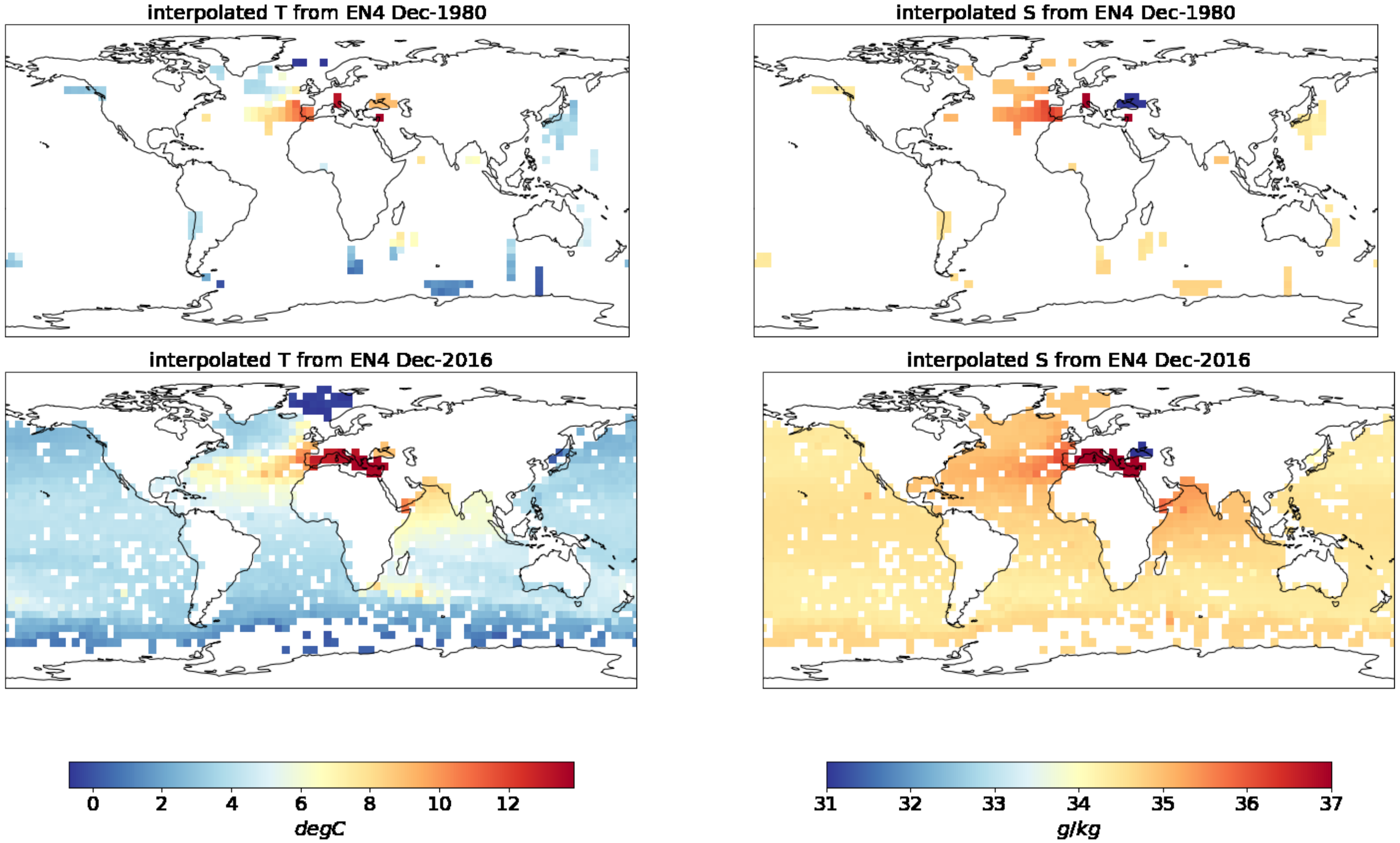
EN4 data from Met Office Hadley Centre:
https://doi.org/10.1002/2013JC009067

Ocean EN4 data used for assimilation: upper 25 meters of the ocean



EN4 data from Met Office Hadley Centre:
https://doi.org/10.1002/2013JC009067

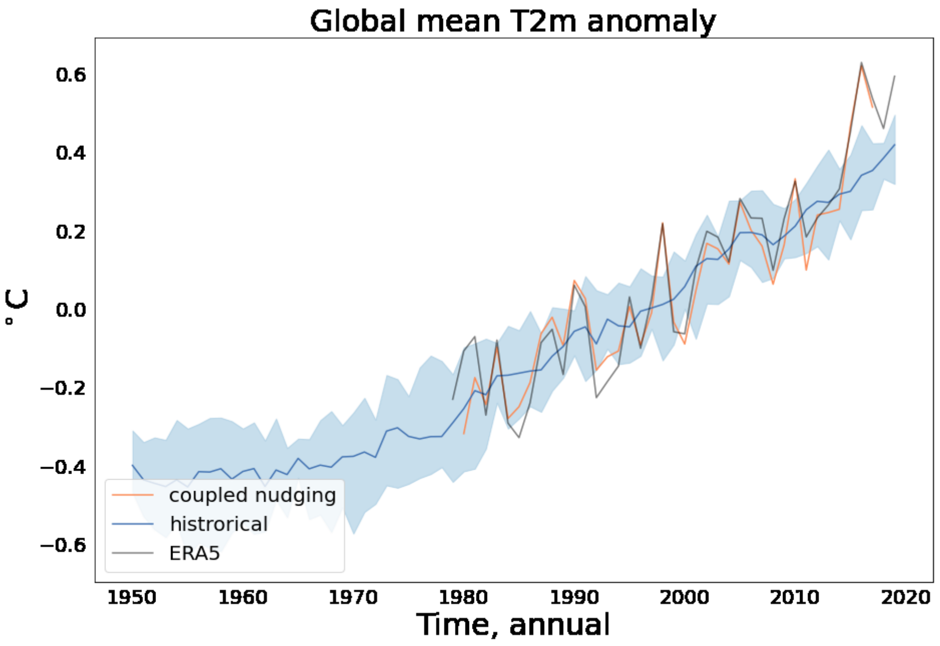
Ocean EN4 data used for assimilation: upper 935 meters of the ocean

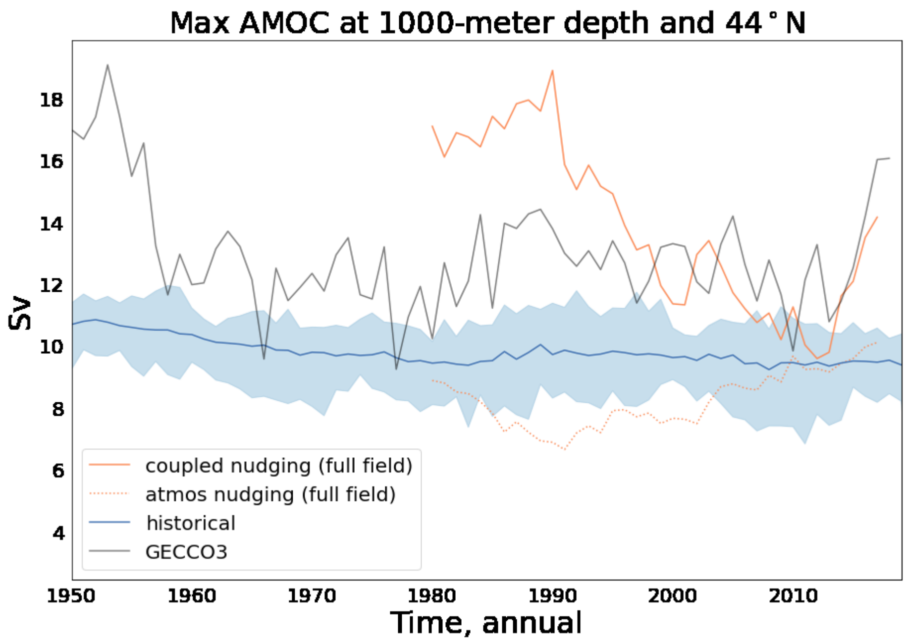


EN4 data from Met Office Hadley Centre:
https://doi.org/10.1002/2013JC009067

CESAM performance for a standard set of historical climate simulations

The externally-forced trend is well represented in CESAM historical simulations

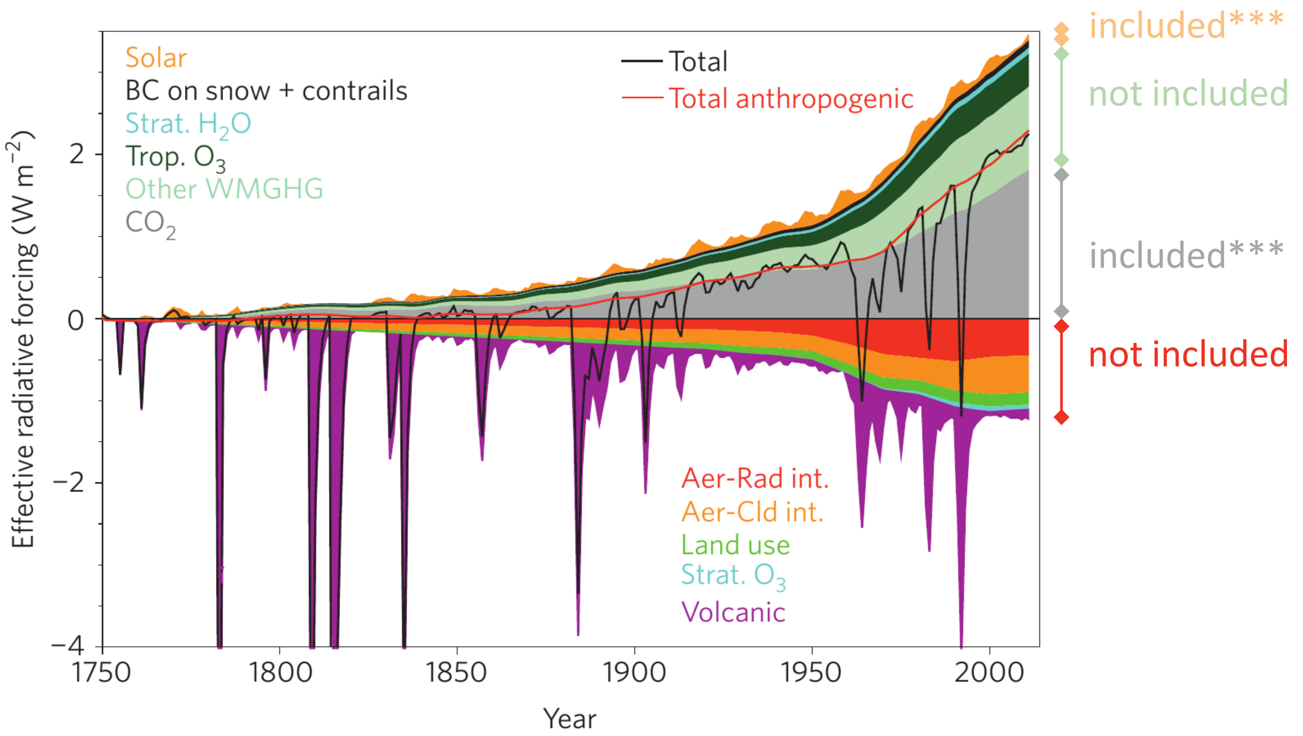




Forcing used for CESAM historical simulations

→ Forcing from the CMIP6 data set

→ Only CO2*** and solar radiation*** used for historical simulation

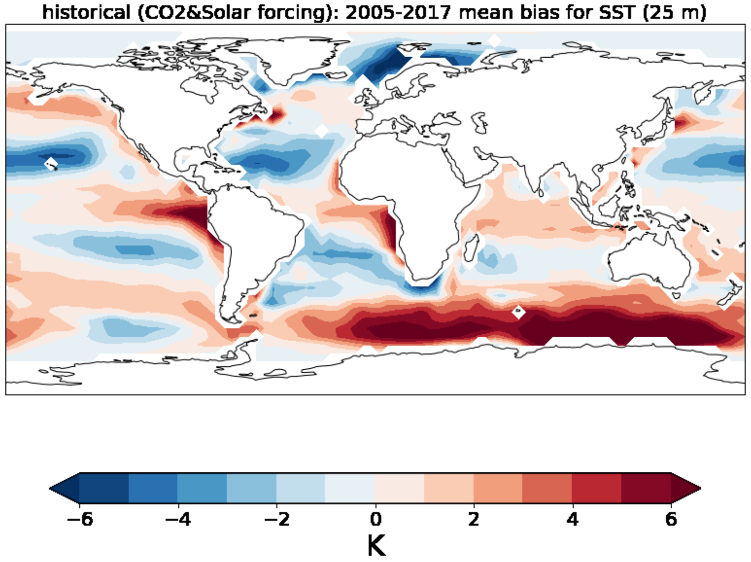


A sense of the relative importance of forcing factors

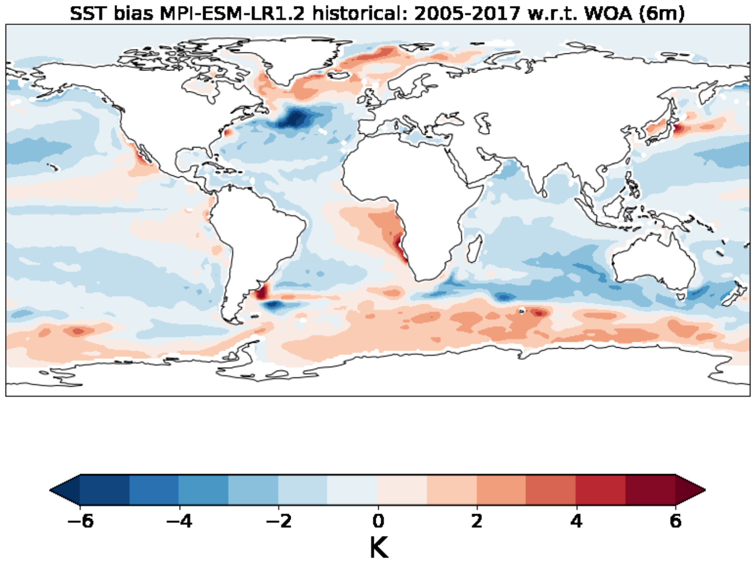
Figure from vonSchuckmann et al 2016
https://csas.earth.columbia.edu/sites/default/files/content/nclimate2876.pdf

CESAM performance w.r.t. CMIP6 model MPI-ESM-LR: SST bias (layer thickness is different!)

Historical r1-r10 CESAM

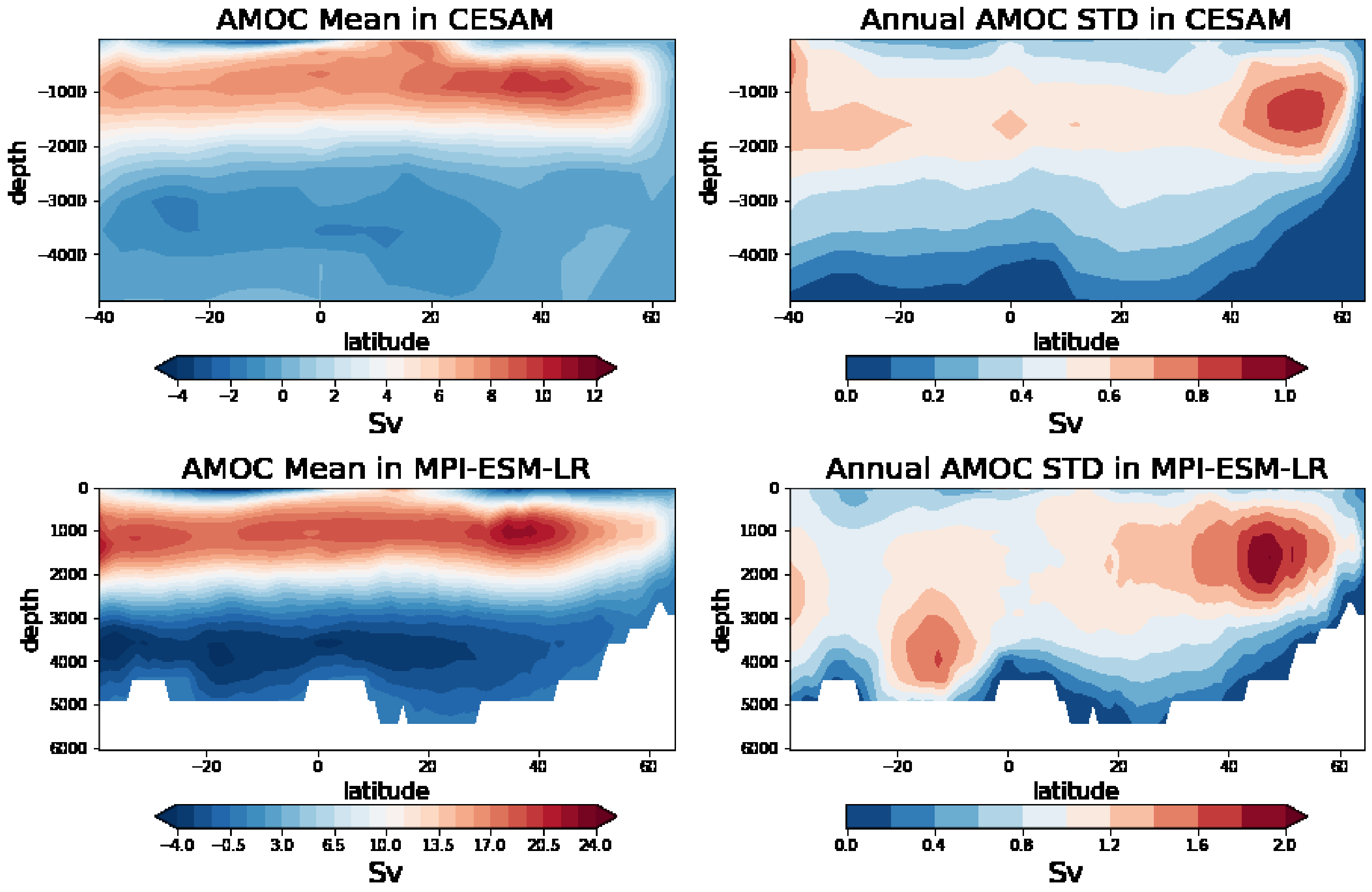


Historical r1 MPI-ESM-LR

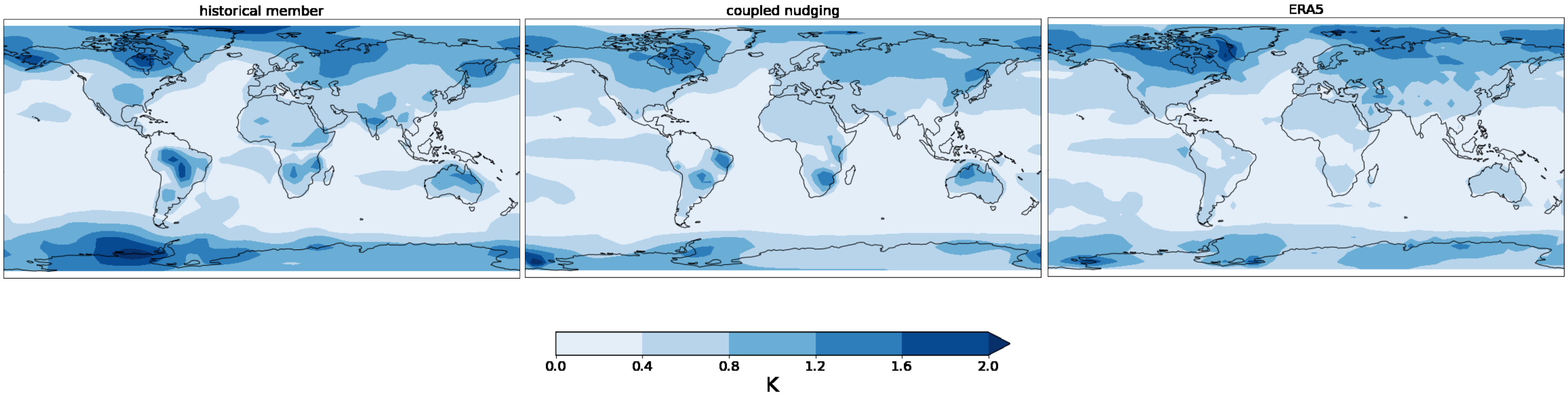


CESAM performance w.r.t. CMIP6 model MPI-ESM-LR: AMOC from historical simulations

The AMOC and its variability in CESAM is almost twice weaker than in the CMIP6 model



CESAM performance w.r.t. ERA5 reanalysis: interannual variability for surface air temperature



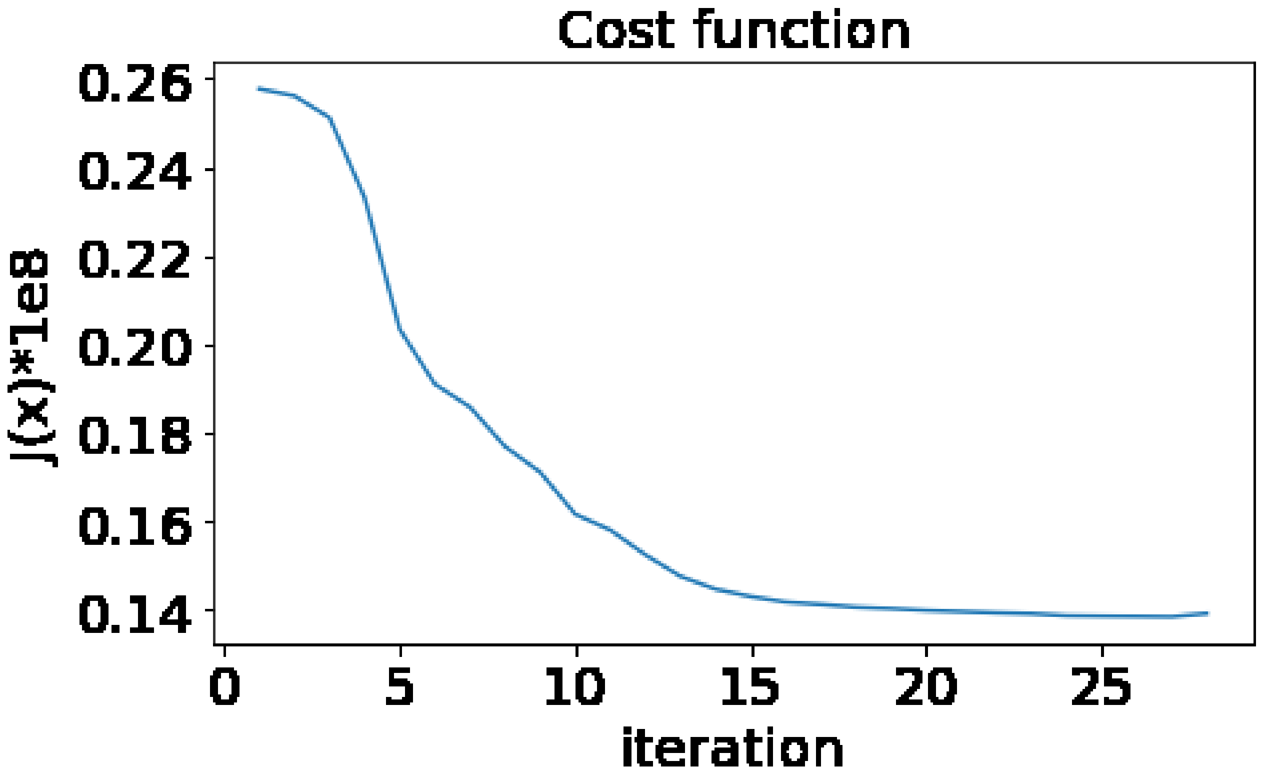
The variability of CESAM in historical runs is comparable to ERA5, except for the equatorial Pacific.
In the nudging run, this is resolved.

Details of 4D-VAR assimilation: cost function and control variables

Control variables:

→ air-sea fluxes and

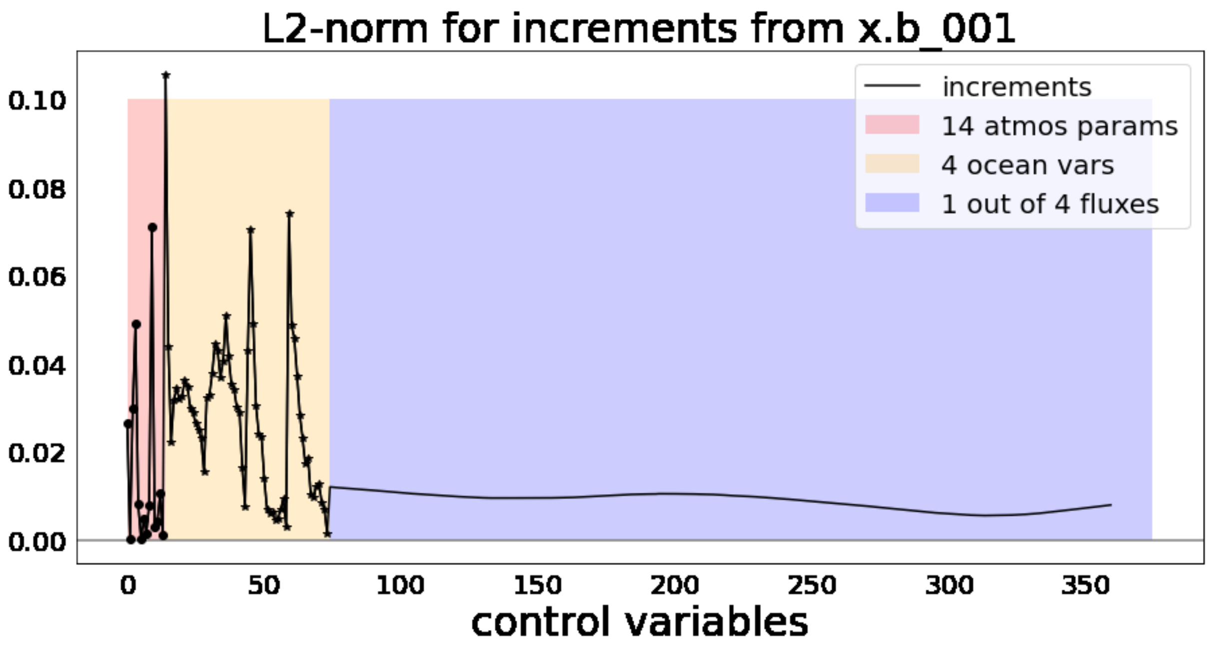
→ oceanic initial conditions



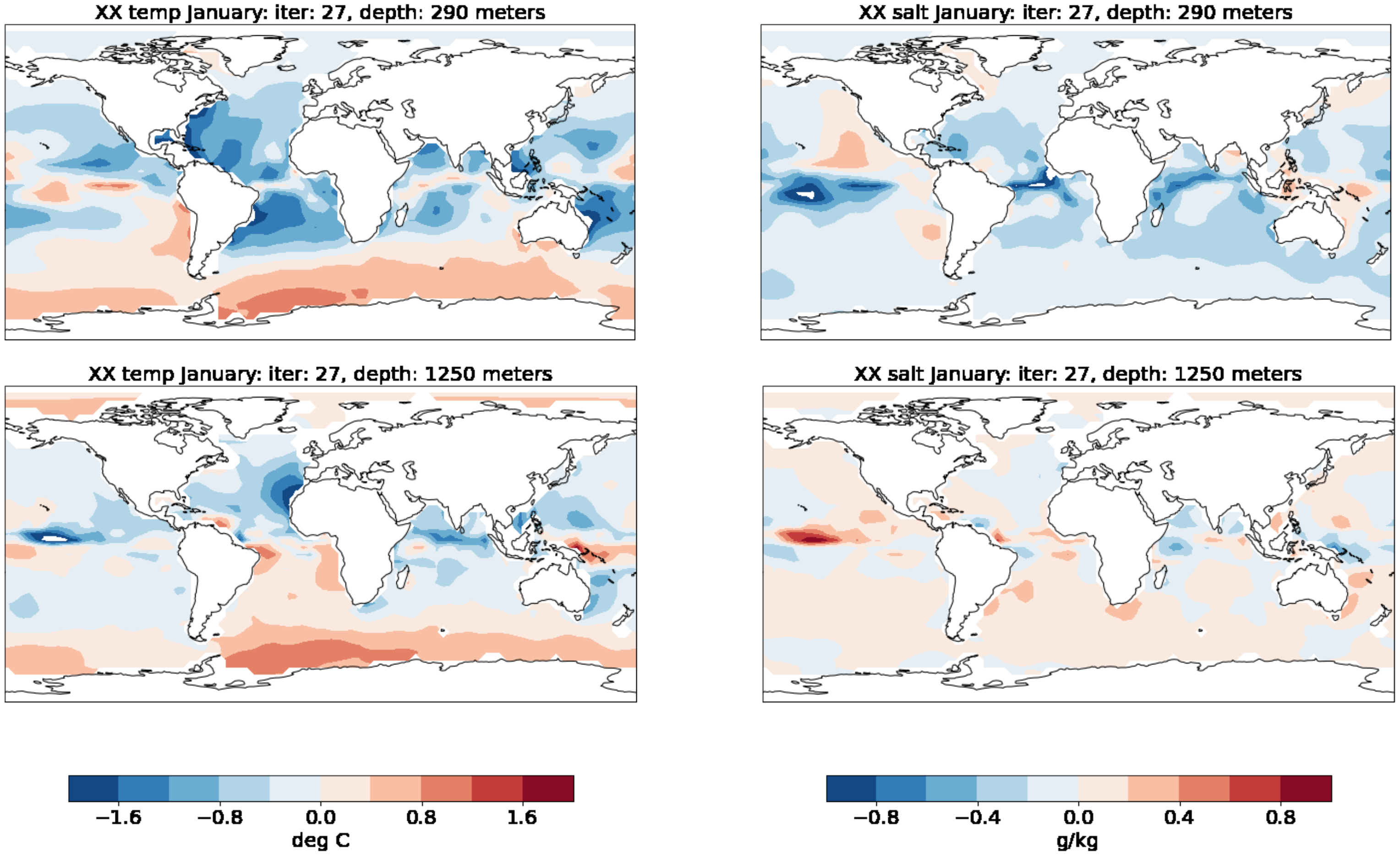
iuliia.polkova@uni-hamburg.de

Details of 4D-VAR assimilation: Scaling and contribution to the cost function

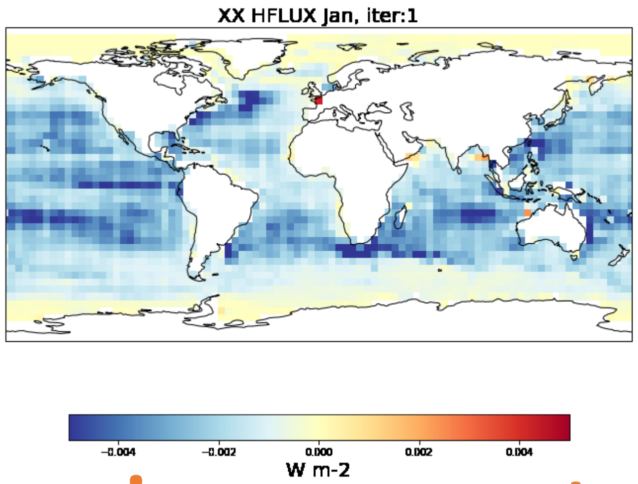
→Tuning of scaling parameters and contributions to the cost function was necessary to ensure reasonable contribution to the cost function and optimization of the selected parameters

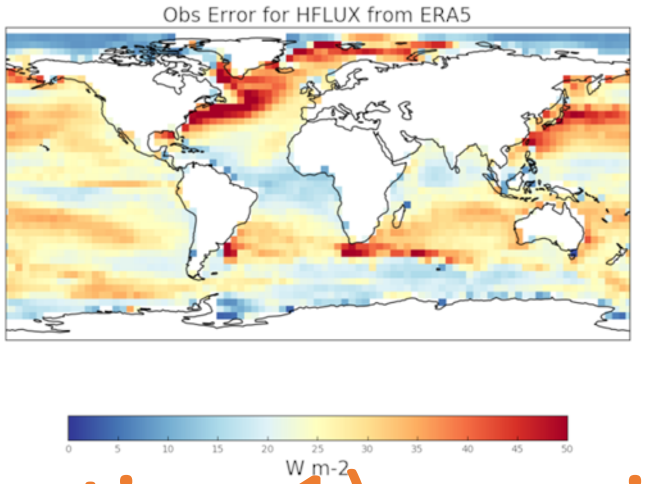


Example of optimized perturbations for ocean temperature and salinity at ~300m and 1000m depth

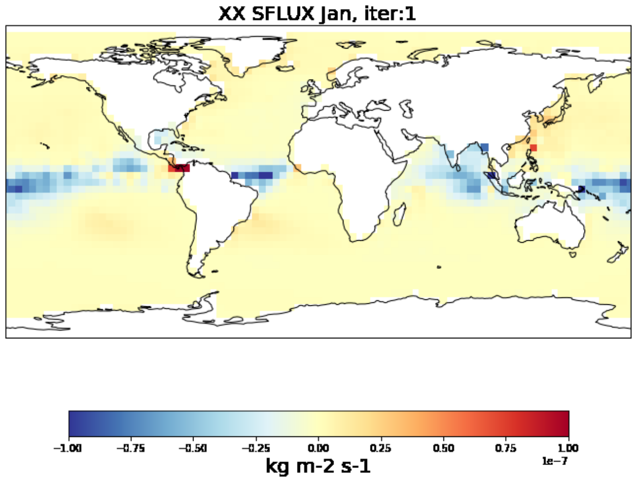


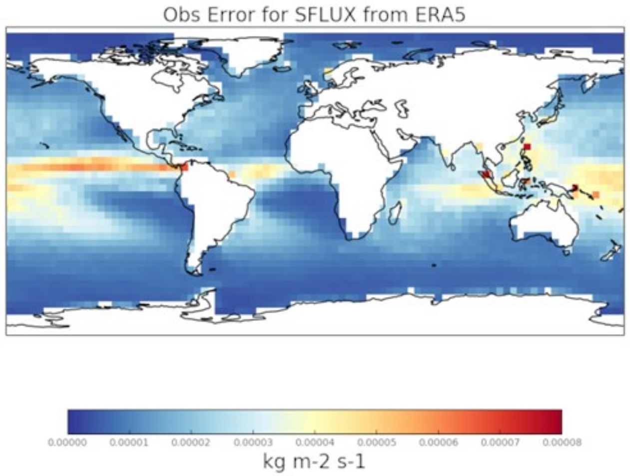
Heat flux increments (iteration 1) vs prior errors





Freshwater increments (iteration 1) vs prior errors





iuliia.polkova@uni-hamburg.de

Summary

The externally forced trend is captured in CESAM historical simulations. The model has necessary large-scale processes important for predictability at decadal timescales. The variability of CESAM is somewhat weaker in the ocean → This is to keep in mind for implications of transferability of results on CMIP6 models.

So far, the coupled reanalysis was carried out for 1 year (1980) to set up the assimilation scheme. The setup is designed with the assimilation window of 1 year, EN4 T&S, HadISST&WOASSS ocean data and ERA5 atmospheric data for wind, air temperature, humidity, SW and LW, surface pressure and cloud area fraction. Long assimilation windows are possible due to the regularization scheme in the atmosphere. It is a strongly coupled data assimilation scheme because the information flows between the model components. But since the time for information propagation in the atmosphere is controlled by the regularization, the assimilation scheme is less strong than one would want it to be.

Next steps are:
long optimization run → optimized state → coupled reanalysis → decadal predictions
comparison with the conventional nudging-based initialization strategies for decadal predictions

Contact for discussion: iuliia.polkova@uni-hamburg.de

The adjoint method with synchronization

(1) Add nudging term for synchronization

$$\frac{dx}{dt} = \sigma(y - x)$$
$$\frac{dy}{dt} = \rho x - y - xz - \boldsymbol{\alpha(y - y_o)}$$
$$\frac{dz}{dt} = xy - \beta z$$

(2) Select the nudging strength

(3) Estimate parameters using the adjoint method

Cost function:

$$J(\rho) = \frac{1}{2T}\sum_{t=1}^{T}\left[(y(t) - y_o)^2\right]$$

(4) Evaluate effects of estimated parameters on the free-running model

Abarbanel et al (2010), https://doi.org/10.1002/qj.600