Developments to the assimilation of sea surface temperature

James While, Daniel Lea, Matthew Martin

ERA-CLIM2 General Assembly, January 2017

Contents

- Introduction
- Improvements to SST bias correction
- Development of an EOF-based error covariance model in NEMOVAR.

Introduction

- The CERA system currently uses a nudging scheme to relax to a pre-existing SST analysis (from HadISST2).

- Direct assimilation of SST observations using NEMOVAR should allow more coupling between the ocean and atmosphere within the coupled data assimilation system.

- It should also allow a more consistent treatment of SST and profile temperature observations.

- At the Met Office, we currently assimilate SST observations directly in the satellite era.

- In order for the scheme to be able to replace the nudging scheme in CERA we need to deal with two issues:

  1. During the satellite era, biases in the different satellites need an improved bias correction scheme.

  2. The pre-satellite era observing network is very sparse so we need to make better use of these observations.

Introduction

SST bias correction

In any coupled system fluxes of heat and moisture between the atmosphere and ocean depend critically on the Sea Surface Temperature (SST). Biases in SST will lead to biases in these fluxes.

Assimilation of biased observations will lead to biases in SST. We therefore need to correct for these biases.

Part of our contribution to ERA-CLIM2 is a variational bias correction system for satellite SST observations.

The bias correction scheme will combine a newly implemented variational bias correction method with a correction based upon “observations-of-bias”.

Observations-of-bias are taken as the differences between standard observations and hi-quality reference data.

The bias correction system is designed to give consistent results over long periods of time; including periods where the amount of reference data is much less than it is now.

Bias correction System
theory

J:-	cost
x:-	state vector
y:-	observations
b:-	observation bias
c:-	model bias
k:-	matchups
B:-	background error covariance
S:-	model bias error covariance
O:-	observation bias error covariance
L:-	matchup error covariance
$\mathbf{H_y}$:-	observation operator for observations
$\mathbf{H_k}$:-	observation operator for matchups

Our scheme is a variational method where biases are calculated within the assimilation itself.

Specifically we aim to minimise the function:

$$
\begin{aligned}
J \quad = \quad & (\mathbf{x}-(\mathbf{x}^{\mathrm{f}}-\cancel{\mathbf{c}}))^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}-(\mathbf{x}^{\mathrm{f}}-\cancel{\mathbf{c}})) + \cancel{(\mathbf{c}-\mathbf{c}^{\mathrm{f}})^{\mathrm{T}}\mathbf{S}^{-1}(\mathbf{c}-\mathbf{c}^{\mathrm{f}})} \\
+ \quad & (\mathbf{b}-\mathbf{b}^{\mathrm{f}})^{\mathrm{T}}\mathbf{O}^{-1}(\mathbf{b}-\mathbf{b}^{\mathrm{f}}) + (\mathbf{y}-H_y(\mathbf{x}+\mathbf{b}))^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}-H_y(\mathbf{x}+\mathbf{b})) \\
+ \quad & \boxed{(\mathbf{k}-H_k(\mathbf{b}))^{\mathrm{T}}\mathbf{L}^{-1}(\mathbf{k}-H_k(\mathbf{b}))}
\end{aligned}
$$

Observations-of-Bias **k**

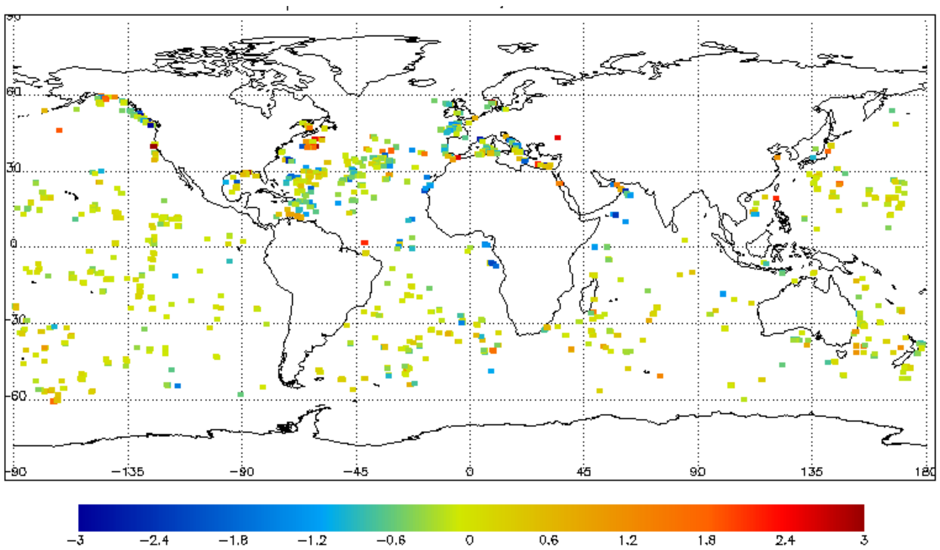
Met Office

We do not have direct observations of the bias.

Instead we use differences between co-located standard observations and assumed 'un-biased' reference data

To prevent cross correlations appearing in the cost function, all observations that are used to calculate the observations-of-bias are NOT included in the observation vector **y**.

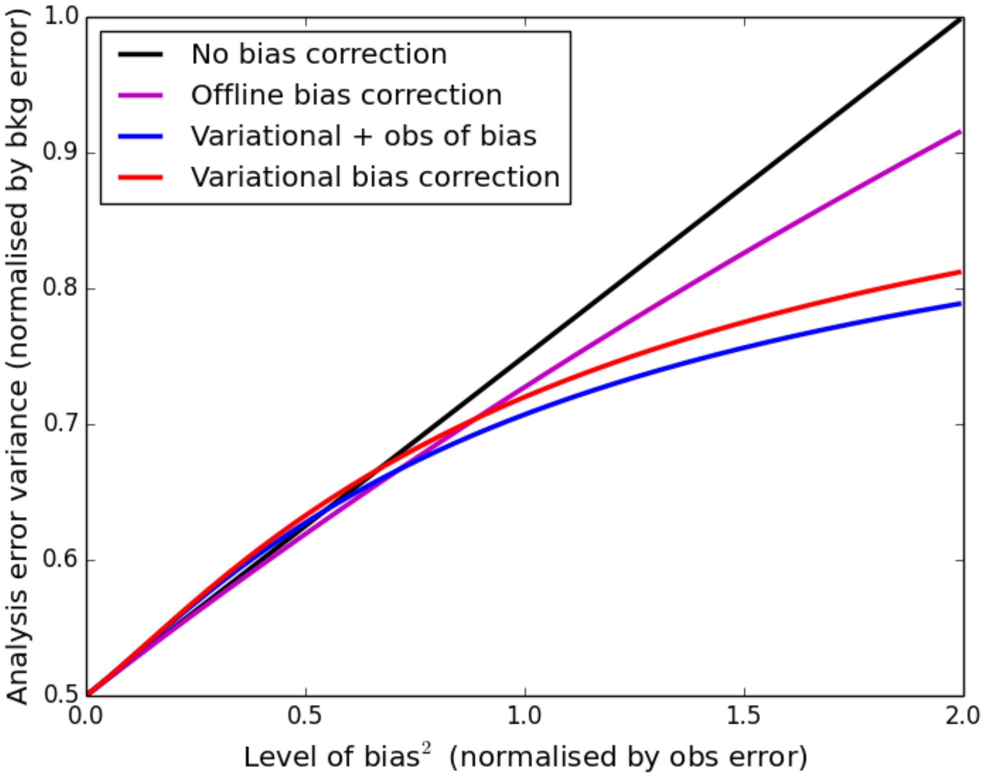
Obs-of-Bias for 14 September 2015



However, the number of co-located observations is small (~1% of the data). Consequently the impact of removing them from **y** should be small.

Conversely the benefit of having a few observations in **k** could be beneficial.

Theoretical errors for an ideal linear system



SST bias correction.

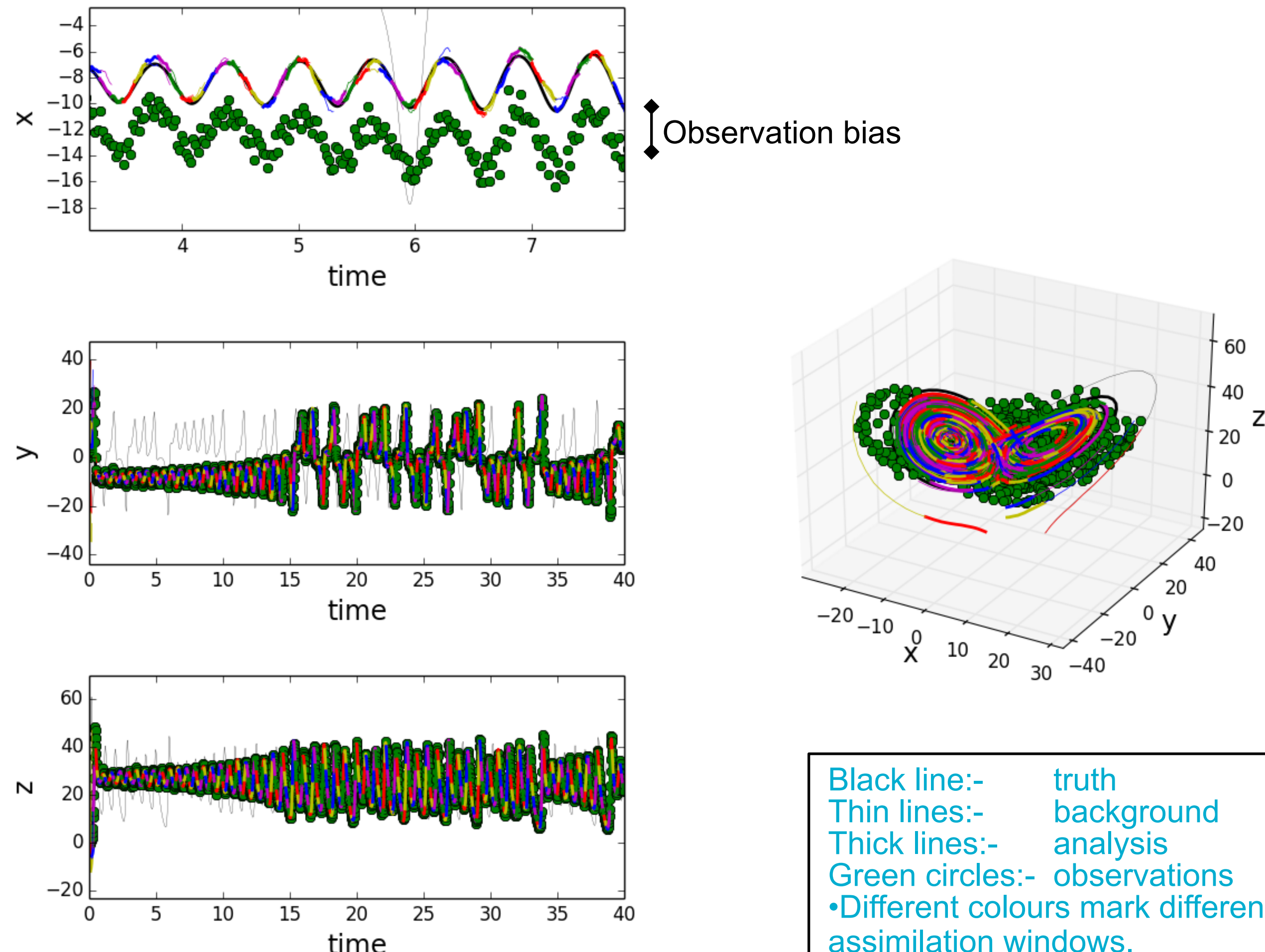
Test system

Met Office

We have tested our ideas using the Lorenz 63 system.

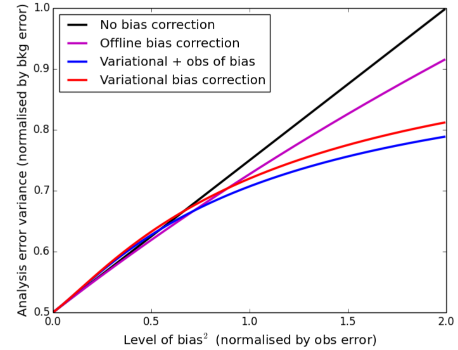
We apply noise to all three dimensions, but only the x direction contains any bias.

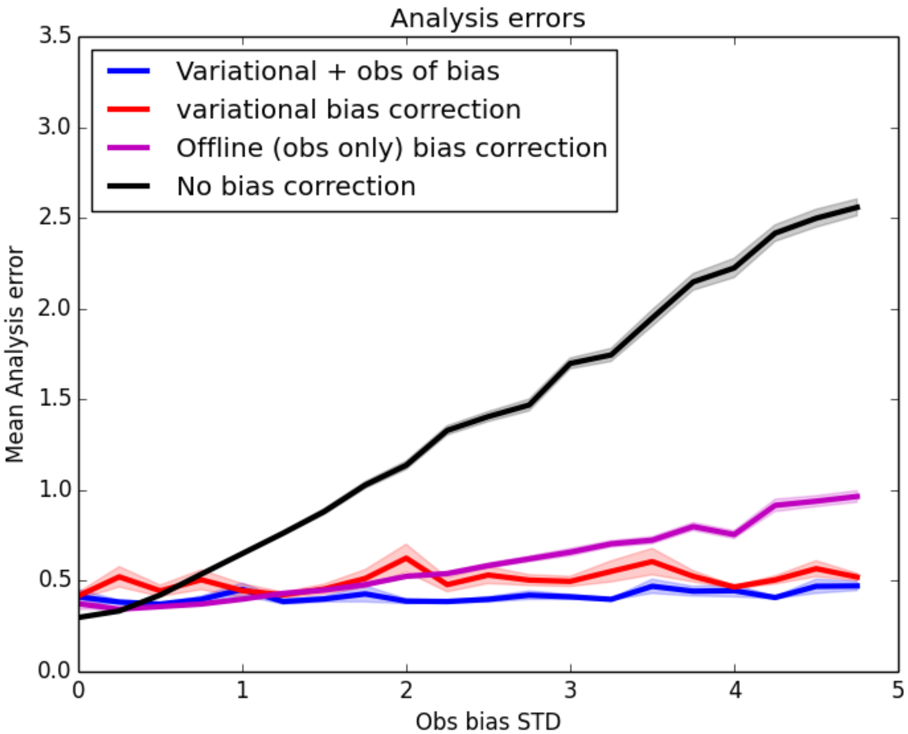
Observations are available for all 3 directions, but only x observations are biased.

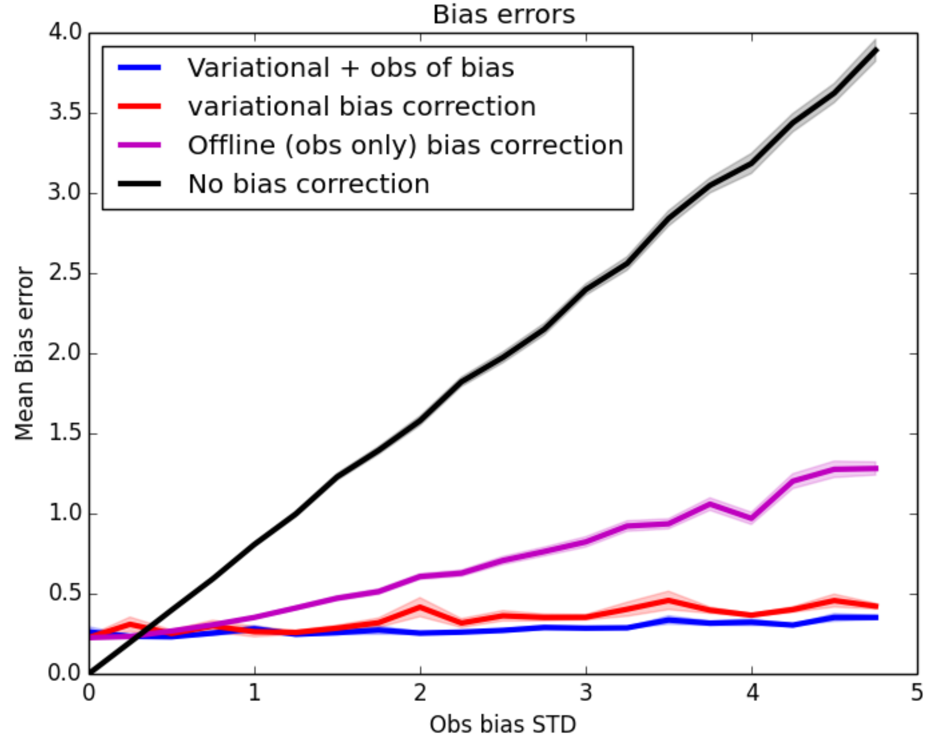


Black line:- truth
Thin lines:- background
Thick lines:- analysis
Green circles:- observations
•Different colours mark different assimilation windows.

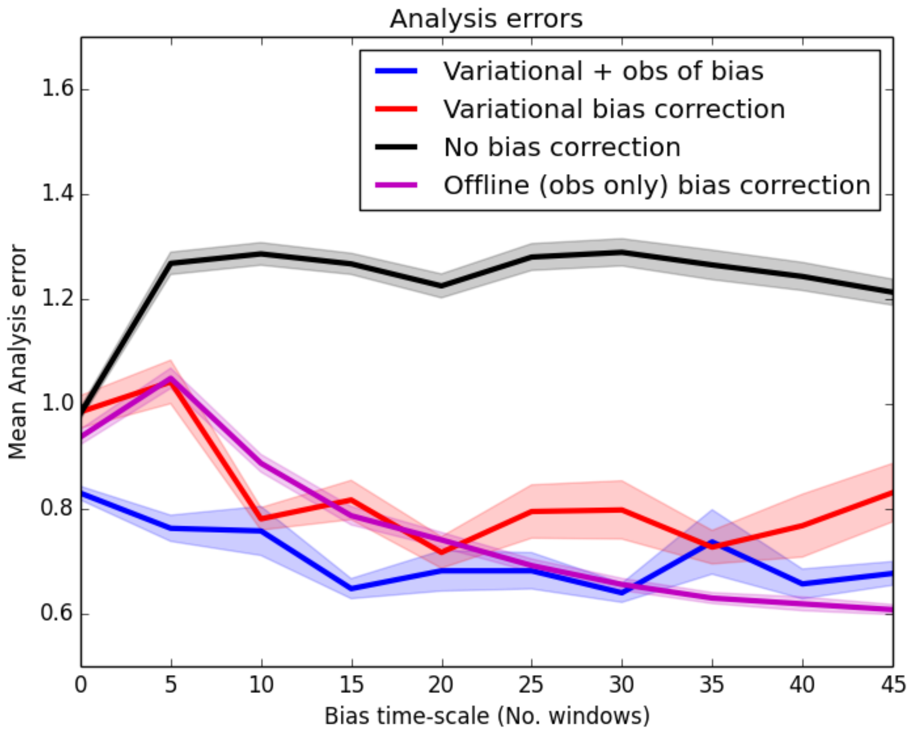
Response to increasing bias

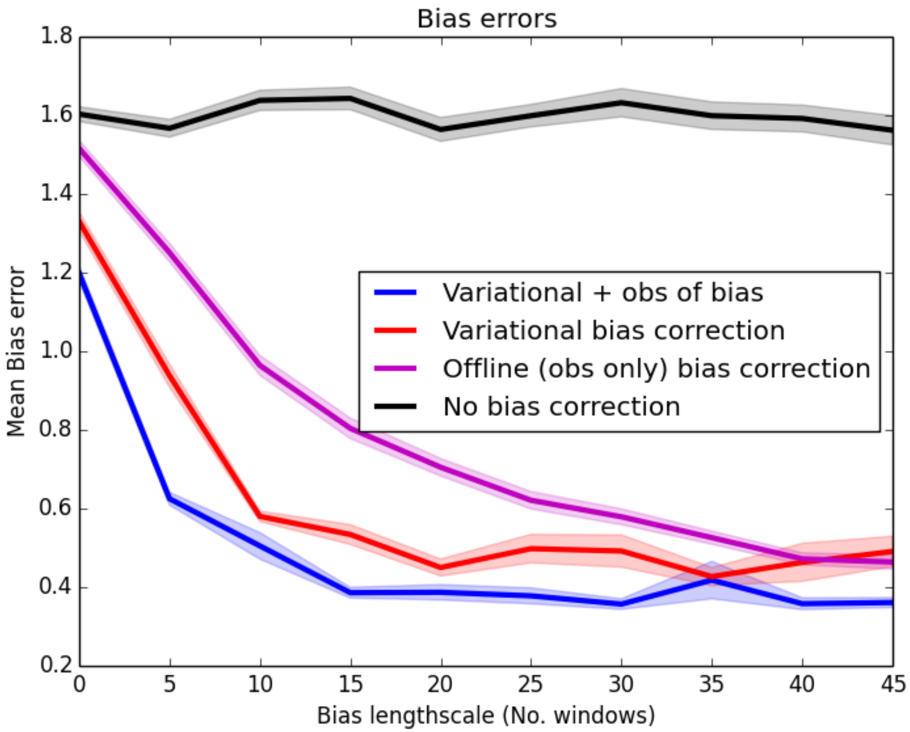




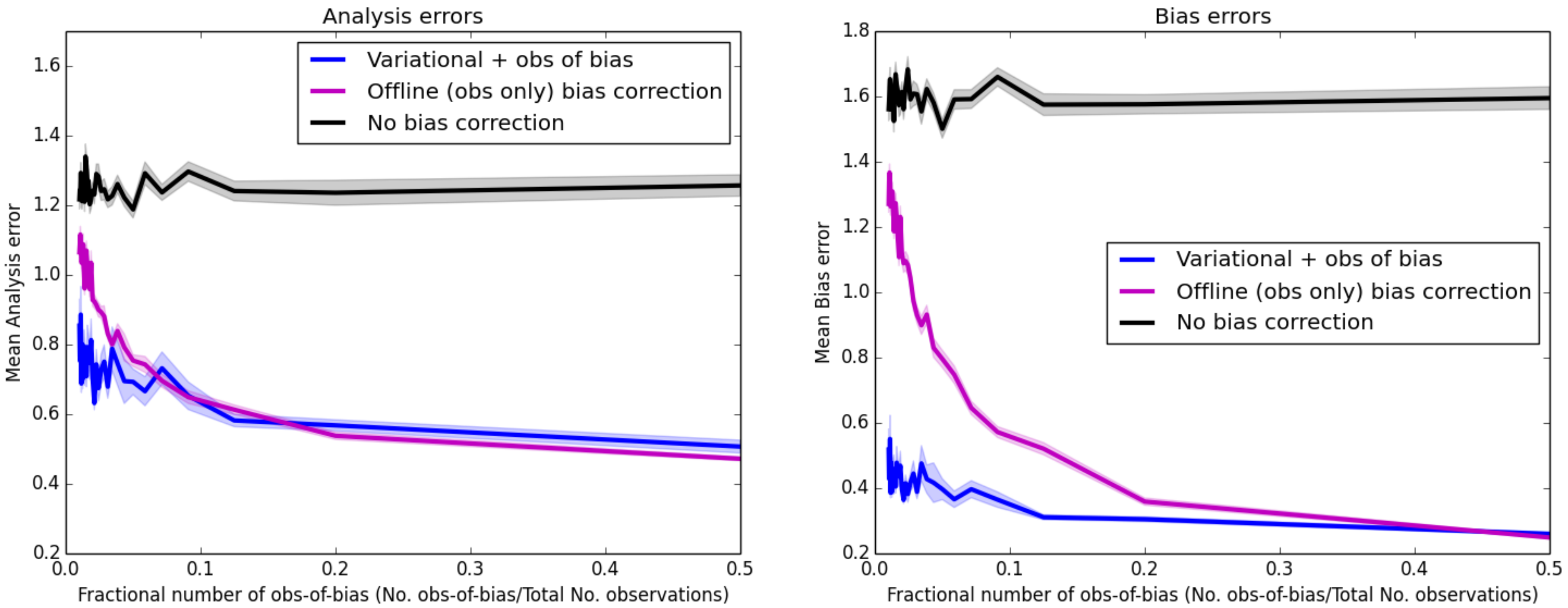


Response to increasing bias time-scale





Response to increasing numbers of observations-of-bias



Current status of the work

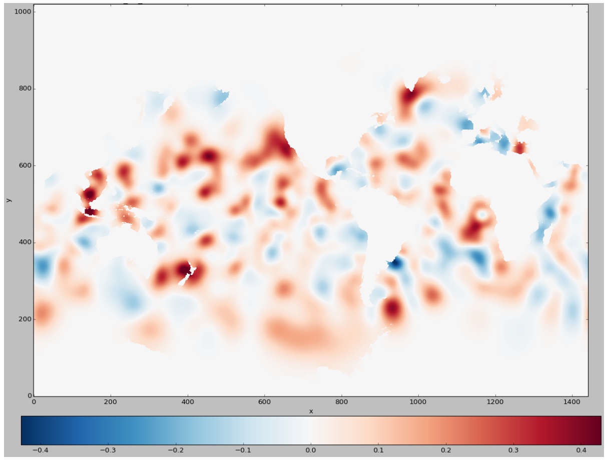
• Theoretical work showed the combined scheme to outperform the existing scheme, and to outperform the standard variational bias correction scheme.

• Coding work has been completed in NEMOVAR to perform the variational bias correction scheme with obs-of-bias. This is now in the ECMWF Git repository.

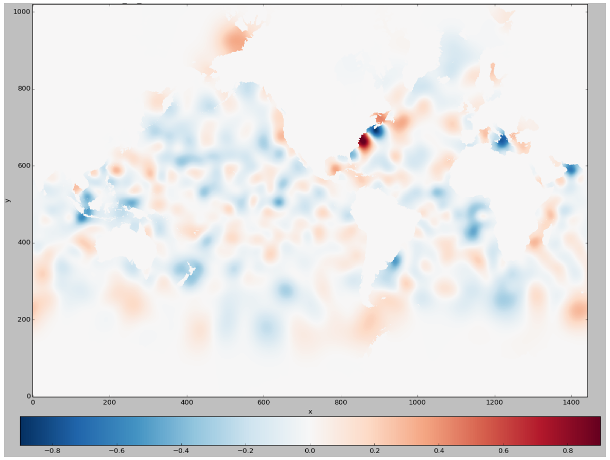
• NEMOVAR can also perform an off-line bias correction scheme (equivalent to the Met Office's current bias correction scheme).

Bias fields from NEMOVAR

NOAA-AVHRR



METOP-AVHRR



Plans for the SST bias correction work

- Parameters for the scheme (error covariances, number of obs-of-bias, etc) need to be estimated.

- Test the bias correction scheme in a cycling suite, and compare to doing no bias correction, and to using an ‘offline’ scheme.

- Write up the results and submit the deliverable by March.

Assimilation using large scale covariances

R&D: Improve 3D ocean assimilation to make best use of the sparse historical data while still doing a good job with today's data

The key thing which gives data assimilation its power is the background error covariance which allows us to spread information from the observation locations.

Can we improve the error covariance structures to allow us to correctly spread sparse observation information over greater distances in order to fill in the gaps?

Improving the error covariances could have other applications too:

Decadal prediction which requires calibration to a historical reanalyses.

Modern day subsurface data assimilation

EOF DA assimilation cost function

Working in model space ("standard")

$$J(\delta\mathbf{x}) = \tfrac{1}{2}\,\delta\mathbf{x}^{\mathrm{T}}\,\mathbf{B}^{-1}\,\delta\mathbf{x} + \tfrac{1}{2}\,(\mathbf{y}-\mathbf{H}(\mathbf{x}_b+\delta\mathbf{x}))^{\mathrm{T}}\,\mathbf{R}^{-1}\,(\mathbf{y}-\mathbf{H}(\mathbf{x}_b+\delta\mathbf{x}))$$

Working in EOF space ("EOF")

$$\delta\mathbf{x} = \mathbf{E}\,\mathbf{a}$$

$$\mathrm{J}(\mathbf{a}) = \tfrac{1}{2}\,\mathbf{a}^{\mathrm{T}}\,\mathbf{\Lambda}^{-1}\,\mathbf{a} + \tfrac{1}{2}\,(\mathbf{y}-\mathbf{H}(\mathbf{x}_b+\mathbf{E}\mathbf{a}))^{\mathrm{T}}\,\mathbf{R}^{-1}\,(\mathbf{y}-\mathbf{H}(\mathbf{x}_b+\mathbf{E}\mathbf{a}))$$

Hybrid

$$\delta\,\mathbf{x} = w_1\,\mathbf{E}\,\mathbf{a} + w_2\,\delta\,\mathbf{x}_{\mathrm{residual}}$$

$$J = w_1\,\tfrac{1}{2}\,\mathbf{a}^{\mathrm{T}}\,\mathbf{\Lambda}^{-1}\,\mathbf{a} + w_2\,\tfrac{1}{2}\,\delta\,\mathbf{x}_{\mathrm{res.}}^{\mathrm{T}}\,\mathbf{B}^{-1}\,\delta\,\mathbf{x}_{\mathrm{res.}} + [\text{obs cost}]$$

a= vector of coefficients/ weights for each EOF (Temperature anomaly = **Ea**)

E= EOFs

Λ= diagonal matrix of eigenvalues (calculated with the EOFs) (squared)

Testing the EOF DA system

EOF error covariance code implemented in NEMOVAR

Performed observing system experiments using real profile and in-situ SST data:

Background from climatology

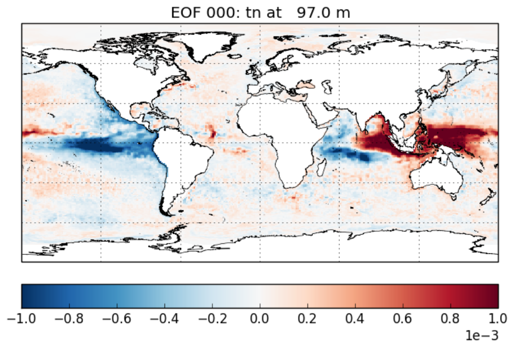
EOFs are 3D multivariate T&S from the most recent Glosea5 reanalysis of the satellite era. Total of 20 EOFs

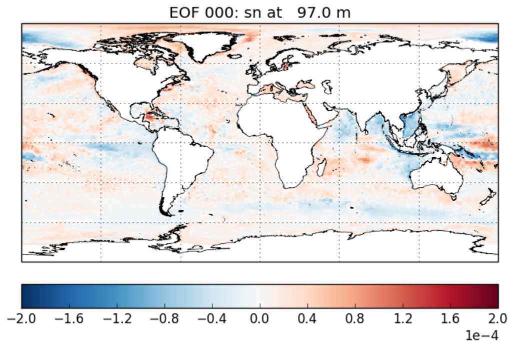
Will show:

Comparisons of increments using EOF and standard data assimilation.

Look at impact on unassimilated data to test whether EOF DA is better at filling in the gaps

EOF 1 - 24.2% of the total variance





Testing the EOF DA system: Observing system experiments

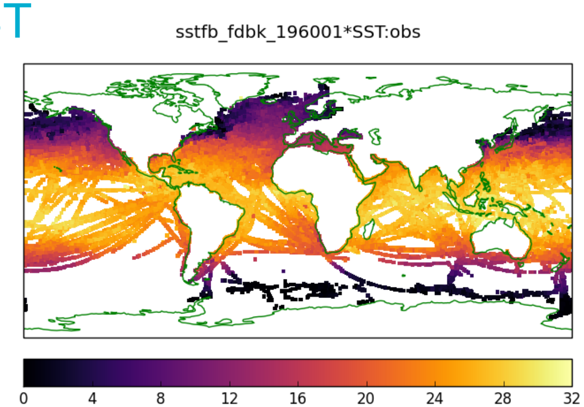
Subsample modern day observations to look like historical data

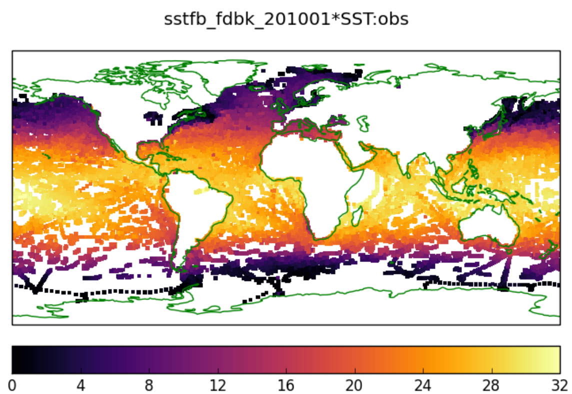
Jan 1960

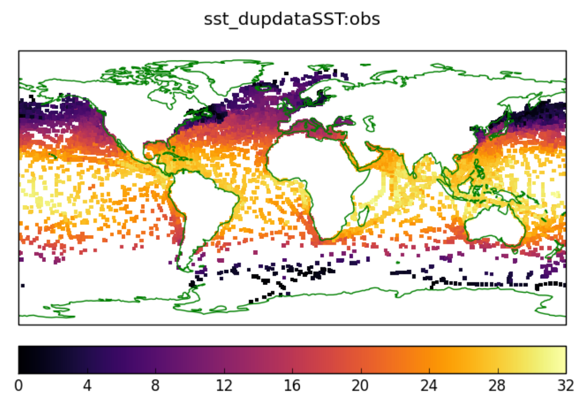
Jan 2010

2010 data subsampled

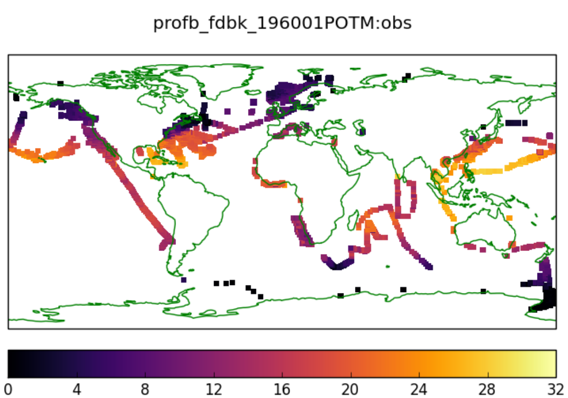
In-situ SST

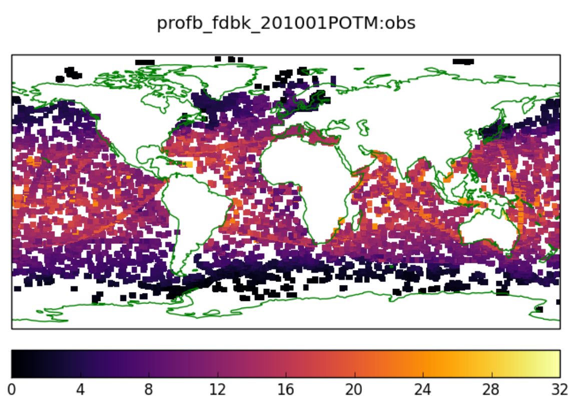


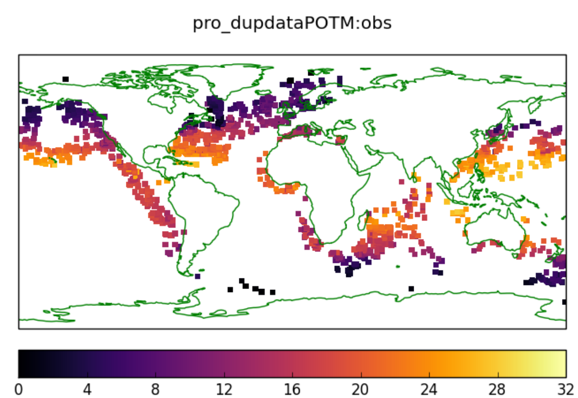




Profile T





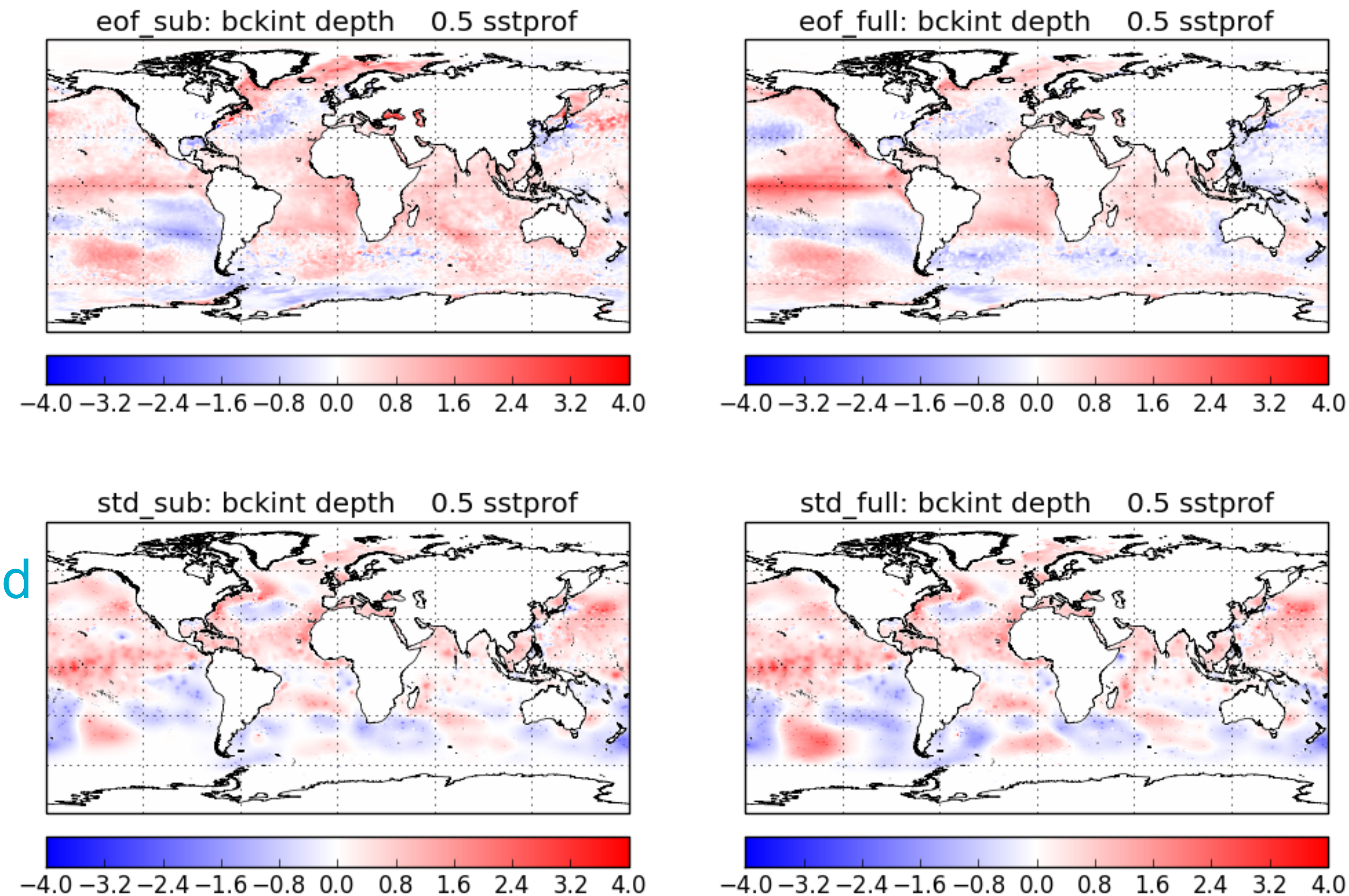


Test results (SST + profile only surface) ocean surface temperature increments /ºC

Subsampled 2010 data | Full 2010 data

EOF DA

Stand ard DA



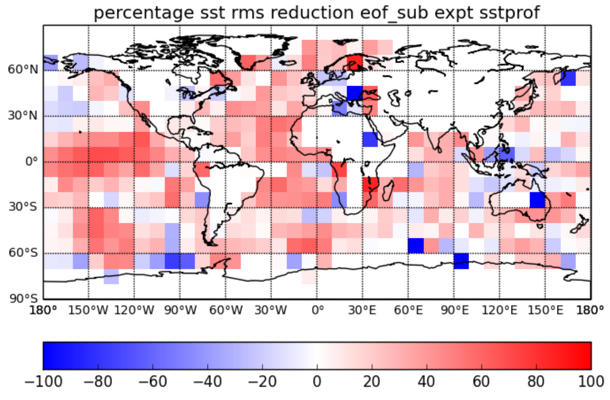
Observations statistics using unassimilated data

% reduction in SST error compared to background

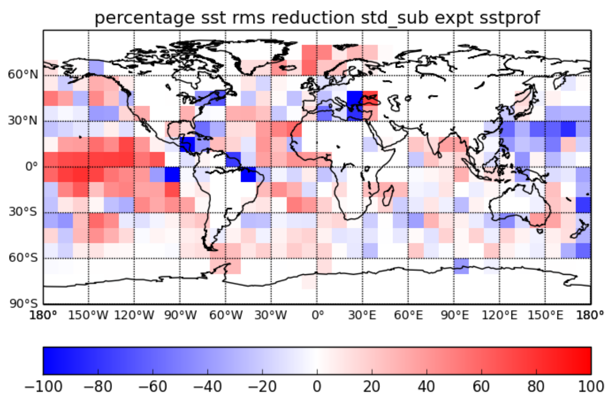
Expts assimilating subsampled SST & profile data

		RMS	Mean	N
bkg	SST	1.0918	0.2415	1228915
EOF	**SST**	**0.8630**	**-0.0588**	
Std	SST	0.9793	-0.0462	
Bkg	T prof	0.9640	0.1601	1876372
EOF	**T prof**	**0.8441**	**-0.1417**	
Std	T prof	0.9515	0.1244	

EOF



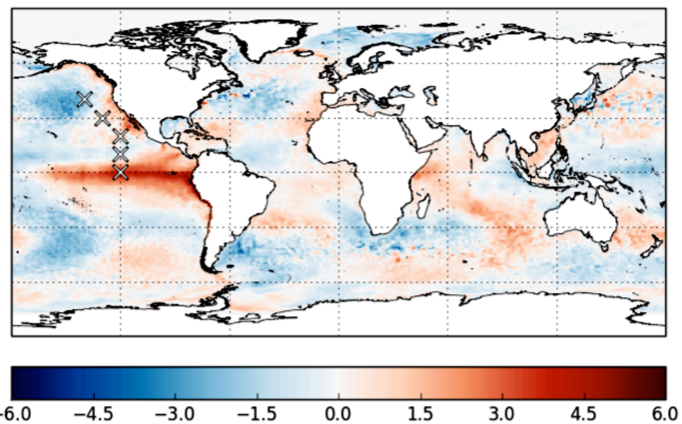
STD



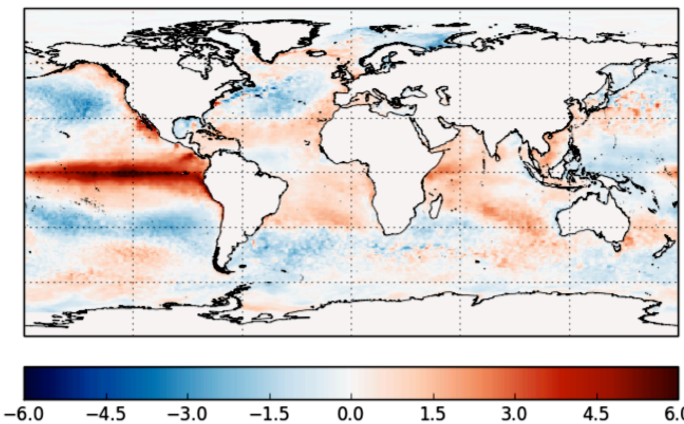
Test: Increments with different hybrid EOF and diffusion modelled covariance weights

Met Office

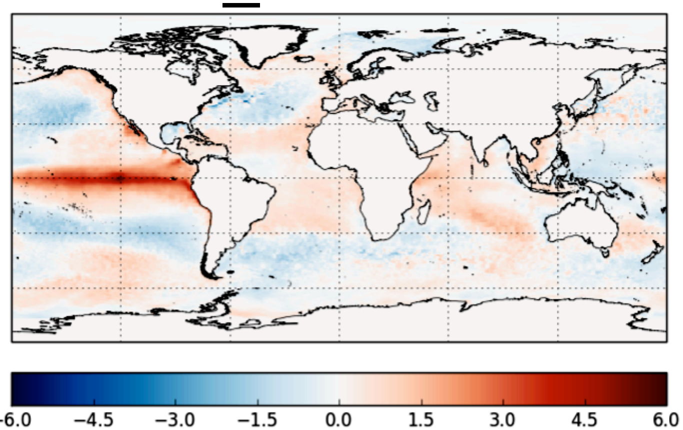
Truth and sim. obs locations



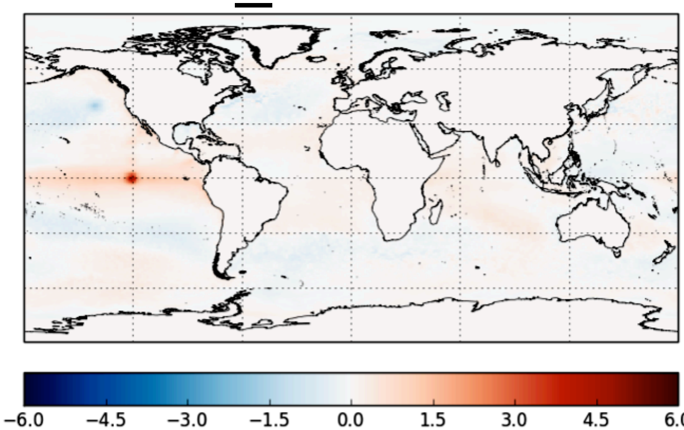
W_1=1 (W_1+W_2=1)



W_1=0.01



W_1=0.001



Summary/plans for EOF DA

- EOFs strongly constrain the results. We get good results where the data are sparse but standard data assimilation may be more effective/robust when the data are dense
- Plan to run more observing system experiments to assess the robustness of the EOF DA
- Investigate the impact of the source of the EOFs.
- Demonstrate feasibility in a reanalysis system.
- Make the code available to ECWMF through the NEMOVAR Git repository.
- Write up results and documentation by March.
- Plan to further test the hybrid of EOF and standard DA

Met Office

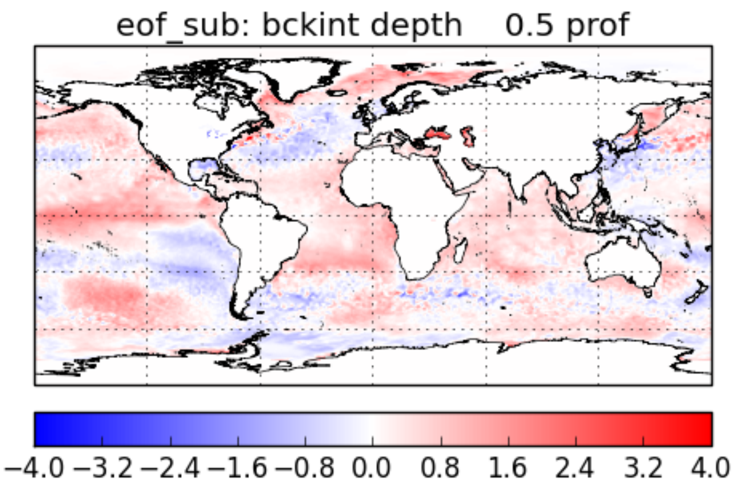
Thank you

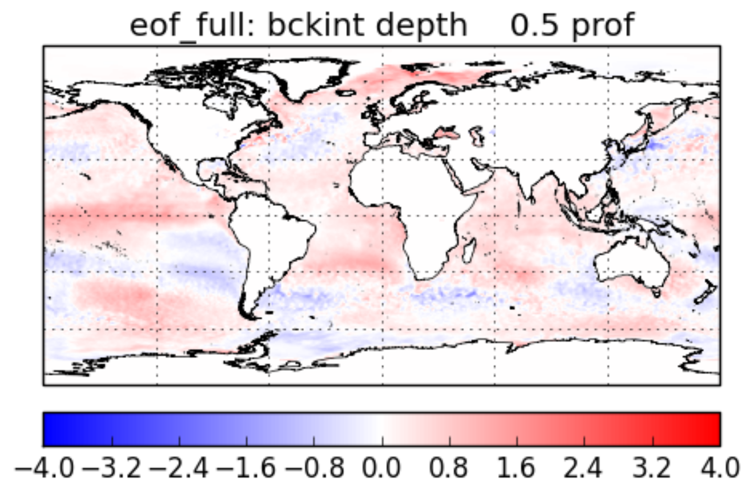
Questions?

Test results (profile data only assimilated) ocean surface temperature increments /ºC

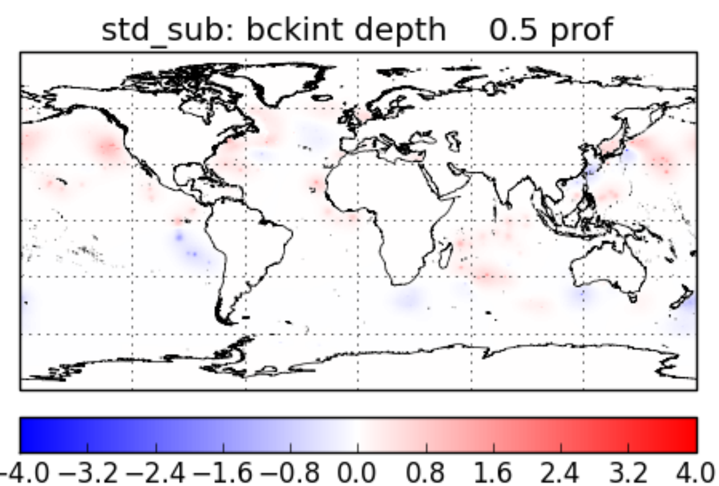
Subsampled 2010 data | Full 2010 data

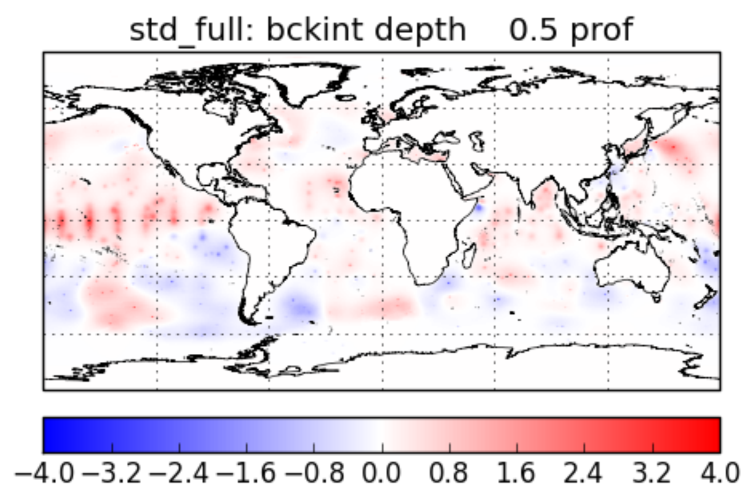
EOF DA





Standard DA





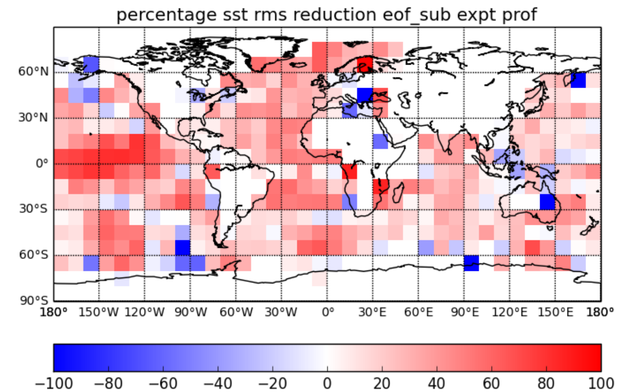
Observations statistics using unassimilated data

% reduction in SST error compared to background

Expts assimilating subsampled profile data only

		RMS	Mean	N
bkg	SST	1.0962	0.1986	1228915
EOF	**SST**	**0.8048**	**-0.0384**	
Std	SST	1.0962	0.1986	
Bkg	T prof	0.9640	0.1601	1876372
EOF	**T prof**	**0.7536**	**-0.0390**	
Std	T prof	0.9600	0.1403	

EOF



STD