Large-scale error covariances for the ocean component of historical coupled reanalysis

Daniel Lea and Matthew Martin

CDAW, Toulouse Oct 2016

Motivation 1/2

ERA Clim 2 coupled reanalysis run from 1900 to approx present day

Problem the ocean observation network has changed markedly in the last 100 years or so.

Observations were sparse and sampling was inhomogeneous. Observations now much less sparse and more globally homogeneous, but we still have sparse sampling at depth

Currently ERA Clim 2 relaxes to a separate SST product HadISST2

Motivation 2/2

R&D: How can 3D ocean assimilation 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

NEMOVAR what is it

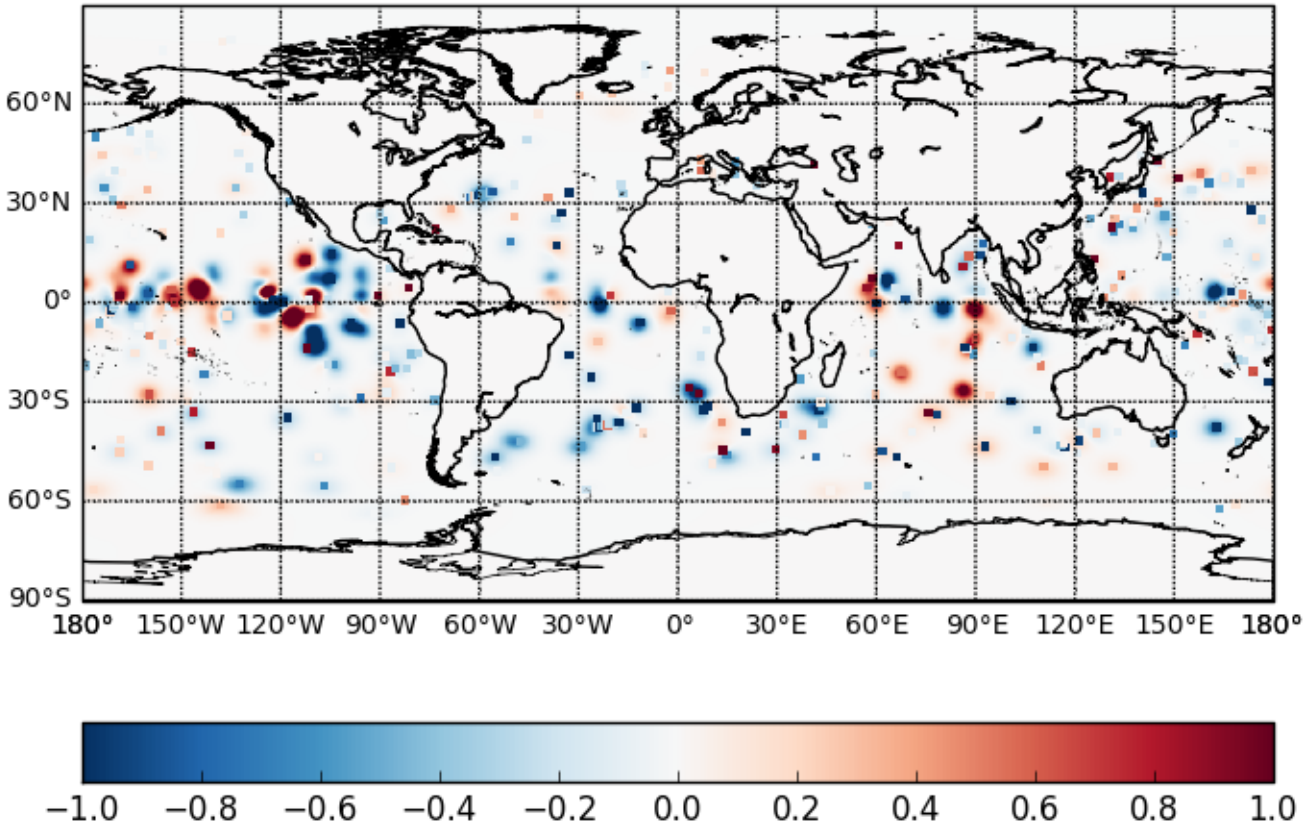
- Ocean data assimilation system for NEMO (3D/4D variational)
- It's the result of a collaboration between CERFACS, ECMWF, INRIA, Met Office ...
- At the Met Office this is used in FOAM (deep and shelf ocean) and OSTIA
- Main benefits: Compatible with the NEMO. It works with the ORCA grids natively. Efficient. There's an effective balance operator.

New hybrid B NEMOVAR – what's new?

- Recently there has been a major improvement to NEMOVAR (mainly developed at CERFACS thanks to Anthony Weaver) (v4)
- We are still using an earlier version (v3) operationally
- The major changes from v3 to v4 are to improve the spatial spreading of observation information:
    - 2D implicit solver for correlation modelled by diffusion
    - Allows the use of ensemble information
    - This can be adapted to use EOF based error covariances

Standard DA

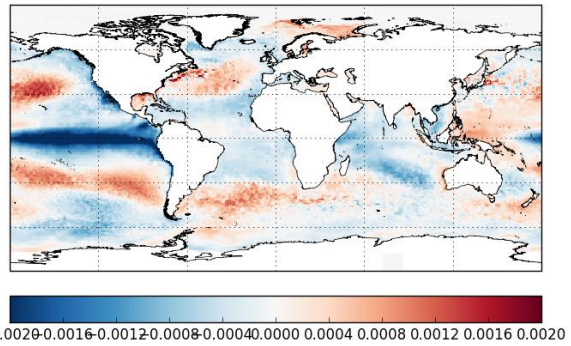
Example increments using just the diffusion modelled covariances, profile only assimilation modern day

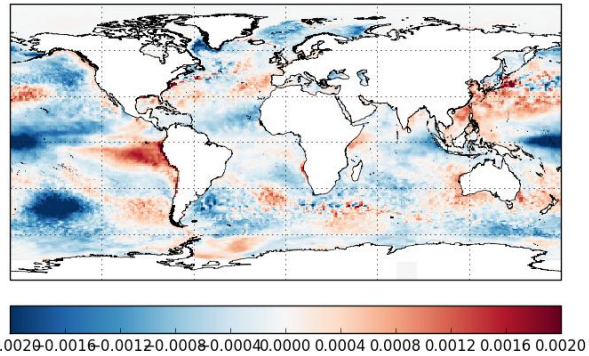


EOF DA

• I've adapted the ensemble covariance code for EOF DA. EOFs are just another "space" to spread the observation information in.

•Top two SST EOFs from variability of the Glosea 5 reanalysis >>>





Question:
The best source of EOFs:

1. EOFs of the actual/estimated initial condition error (variances accurate but little spatial information)
2. EOFs from model/previous reanalysis variability (good spatial pattern but variances may be too large)

Plan is to rescale EOFs estimated from the model variability to more closely represent initial condition error

EOF DA

Working in model space

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

Working in EOF space

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

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

Hybrid

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

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

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

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

Background from climatology

NEMO ORCA025 grid

Compare increments using EOF and standard data assimilation.

EOFs are 3D multivariate T&S from the most recent Glosea5 reanalysis

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

(NEMO standalone observation operator used heavily in this work)

Testing the EOF DA system: Observing system experiments

Met Office

Subsample modern day observations to look like historical data

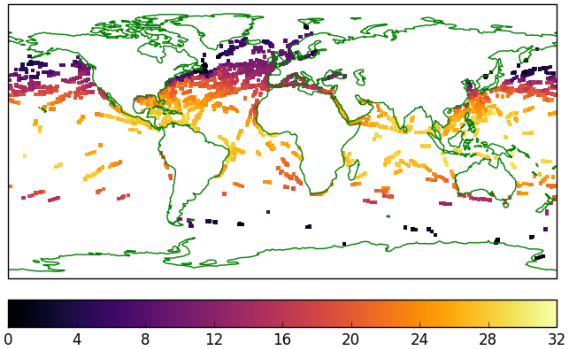
1 Jan 1960

1 Jan 2010

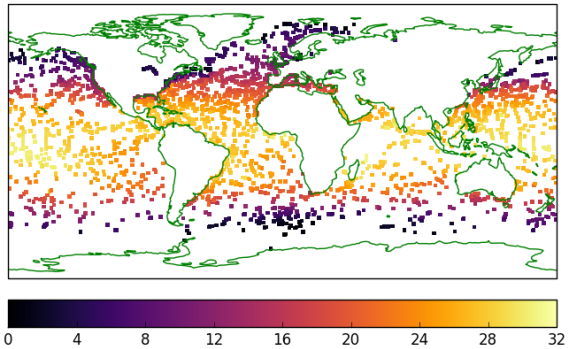
2010 data subsampled

In-situ SST

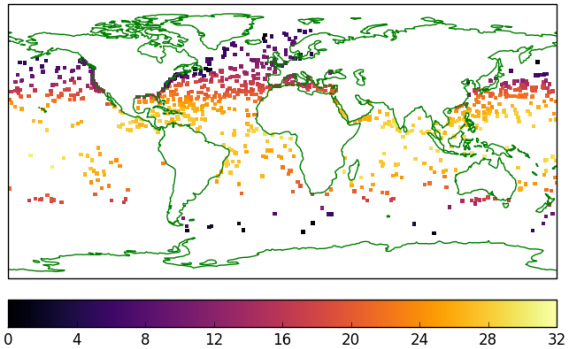
sstfb_fdbk_19600101SST:obs



sstfb_fdbk_20100101SST:obs

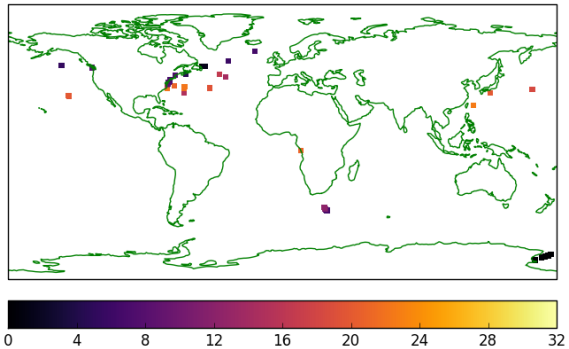


sst_dupdataSST:obs

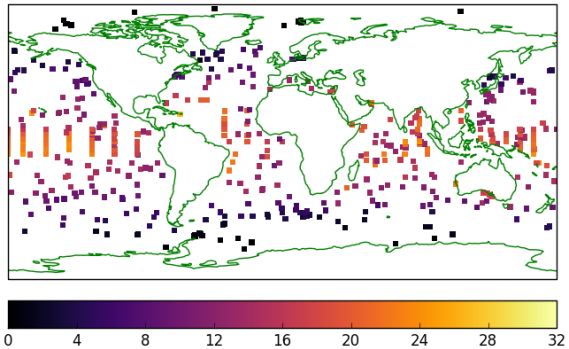


Profile T

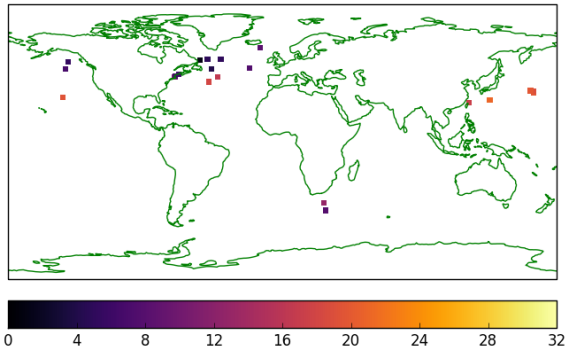
profb_fdbk_19600101POTM:obs



profb_fdbk_20100101POTM:obs



pro_dupdataPOTM:obs



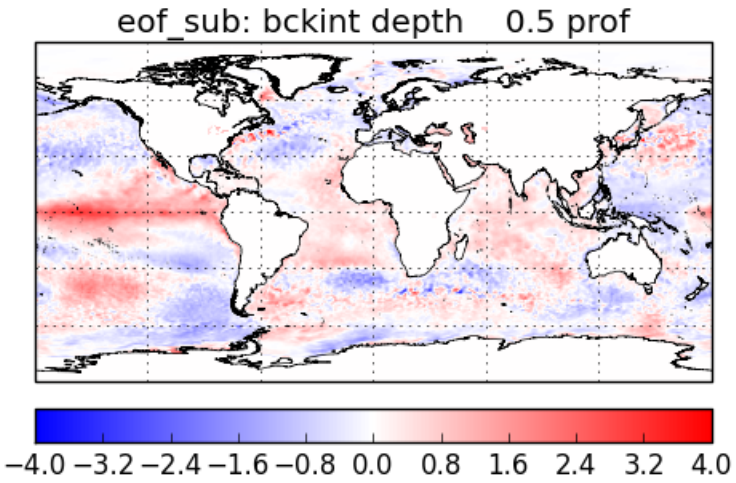
Test results (profile only surface)

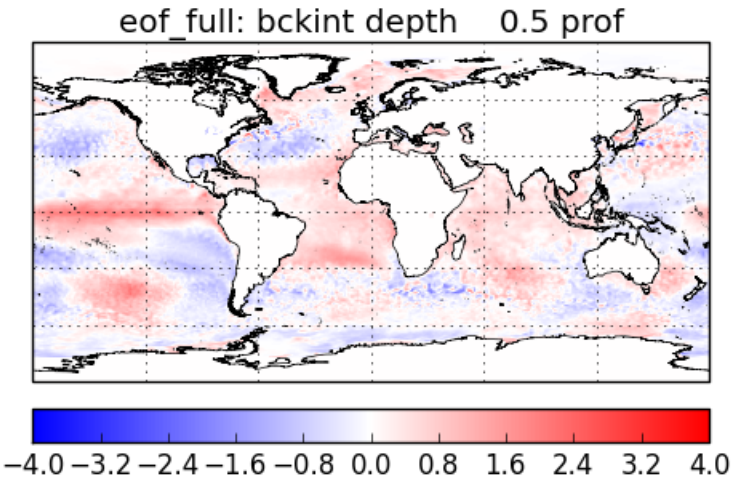
ocean surface temperature increments /ºC

Subsampled 2010 data

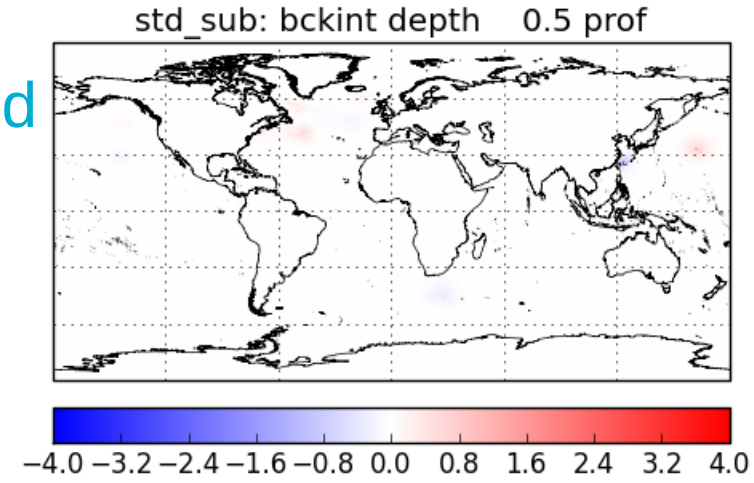
Full 2010 data

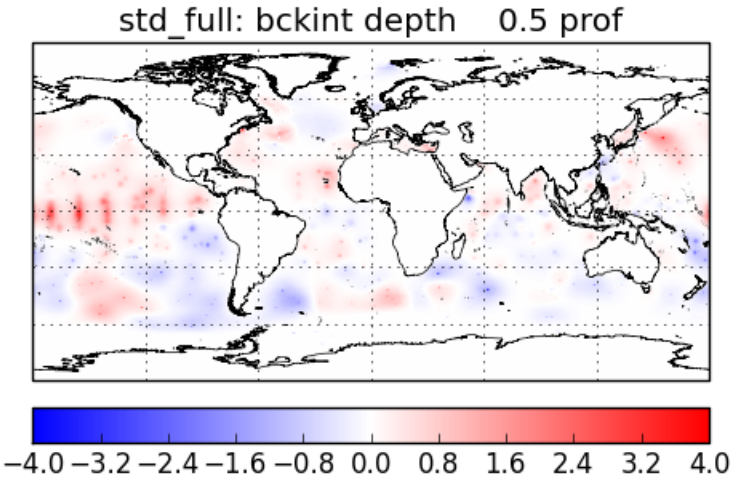
EOF DA





Standard DA



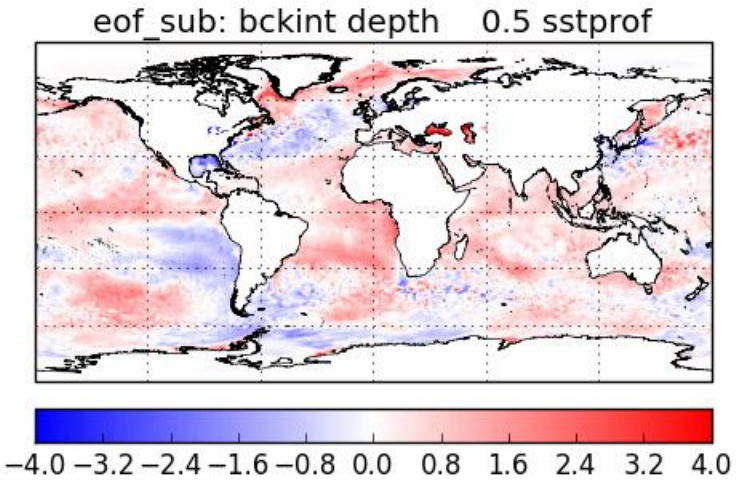


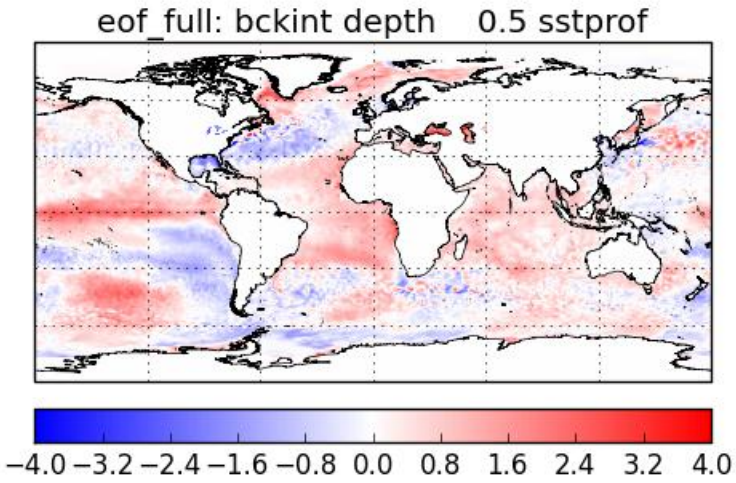
Test results (SST + profile only surface) ocean surface temperature increments /ºC

Subsampled 2010 data

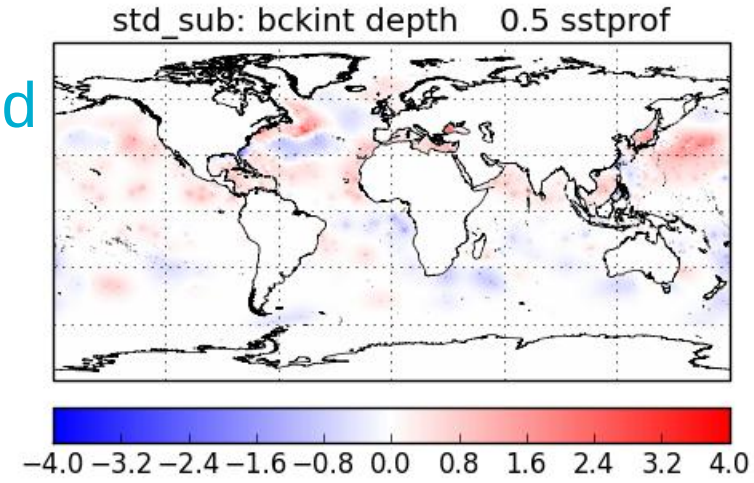
Full 2010 data

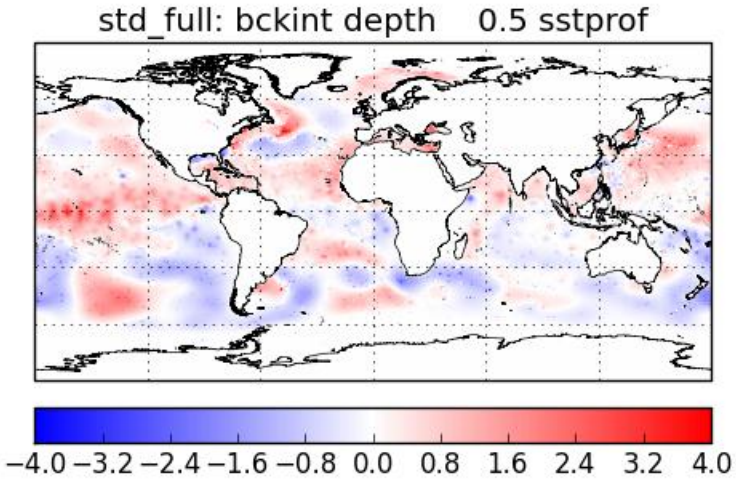
EOF DA





Standard DA





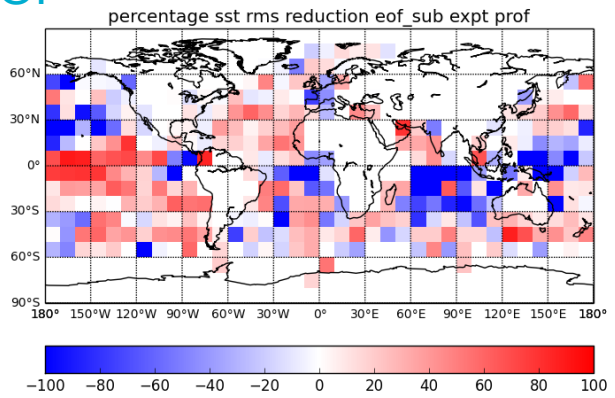
Comparison to observations not assimilated

% reduction in error compared to background

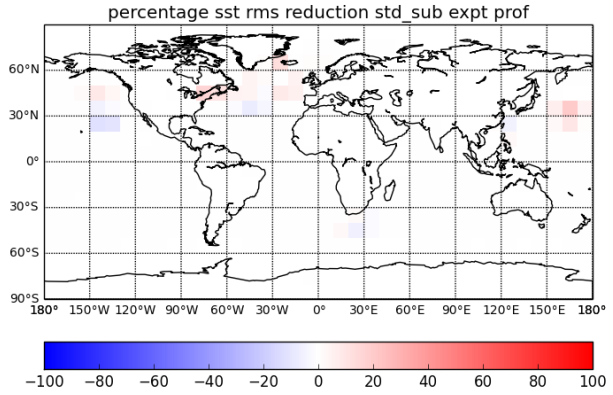
Subsampled profile data only assimilation

		RMS	Mean	N
bkg	SST	1.7256	0.0611	42894
EOF	**SST**	**1.6146**	**-0.1158**	
Std	SST	1.7148	0.0509	
Bkg	T prof	0.9943	0.1637	56178
EOF	**T prof**	**0.8675**	**0.1523**	
Std	T prof	0.9846	0.1635	

EOF



STD



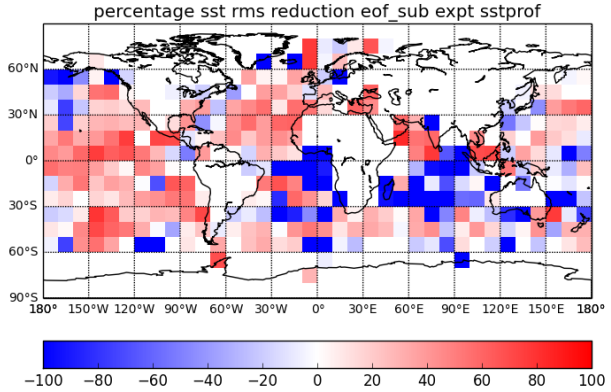
Comparison to observations not assimilated

% reduction in error compared to background

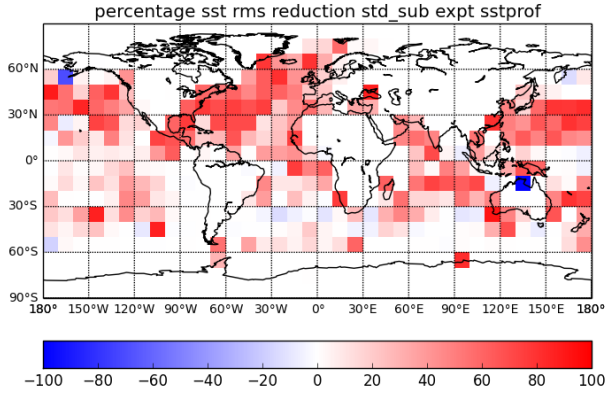
Subsampled SST & profile data assimilation

		RMS	Mean	N
bkg	SST	1.7256	0.0611	42894
EOF	SST	1.3368	-0.1663	
Std	**SST**	**0.9454**	**0.1108**	
Bkg	T prof	0.9943	0.1637	56178
EOF	T prof	1.1070	-0.2237	
Std	T prof	0.9764	0.1600	

EOF

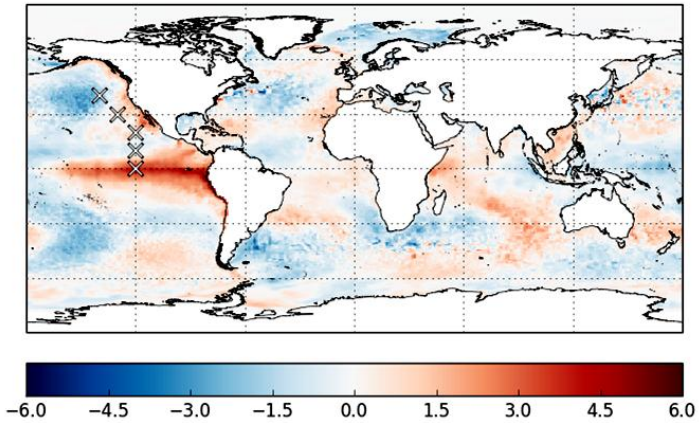


STD

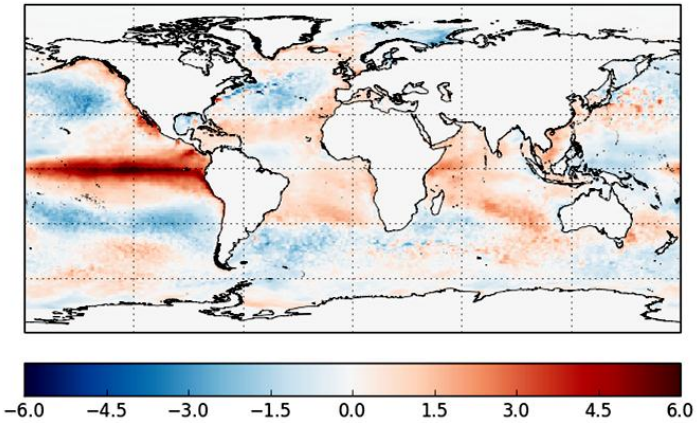


Increments with different hybrid EOF and diffusion modelled covariance weights

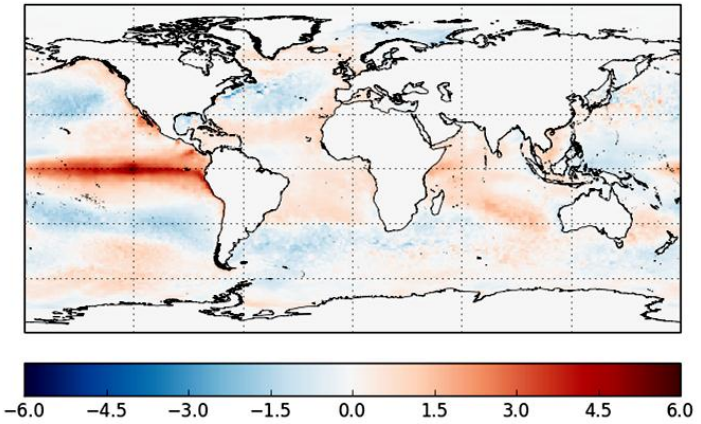
Truth and sim. obs locations



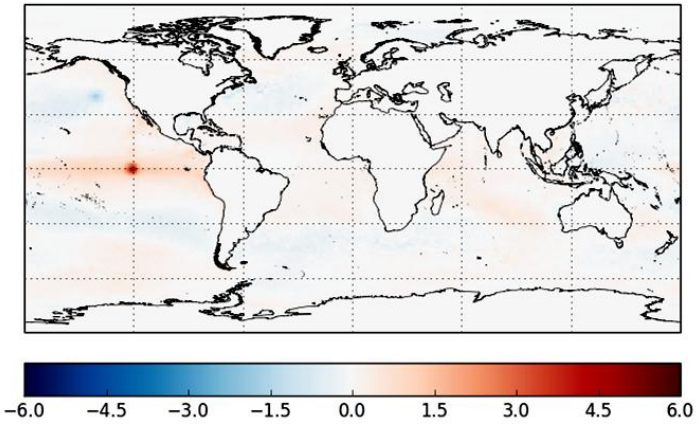
W_1=1 (W_1+W_2=1)



W_1=0.01



W_1=0.001



Plans/summary

- Run more observing system experiments to assess the robustness of the EOF DA

- Investigate at the impact of the source of the EOFs.

- EOFs strongly constrain the results. We get good results where the data are sparse but standard data assimilation is more effective/robust when the data are dense

- Plan to further test the hybrid of EOF and standard DA to get the benefits of both approaches

- Test in a reanalysis system (FOAM like)

- Ultimately test in a decadal prediction system (DePreSys)

Met Office

Thank you

Questions?