Met Office

Assimilation using large scale EOF error covariances

Daniel Lea and Matthew Martin

ERA CLIM2, University of Bern, Bern Dec 2017

Motivation for using large scale covariances

Ocean data is often sparse and inhomogeneous particularly historically

The key thing which gives data assimilation its power is the background error covariance which allows us to spread information from the observation locations. Generally done with simple Gaussian like structures

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?

Applications:

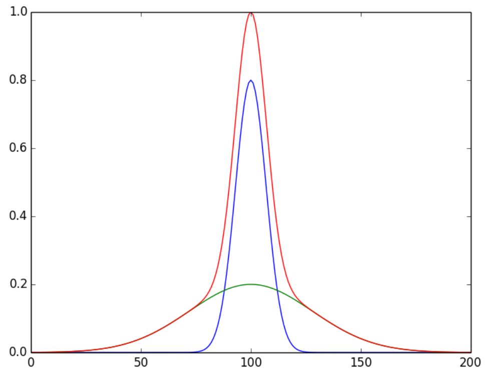
Historical reanalyses

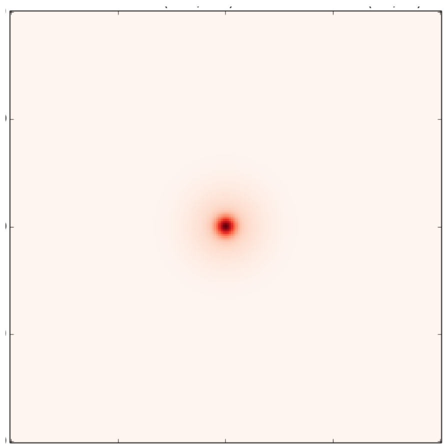
Decadal prediction which requires calibration to a historical reanalyses.

Modern day subsurface data assimilation

Ocean data assimilation system: NEMOVAR

- Ocean data assimilation system for NEMO (3D/4D variational)
- It's the result of a collaboration between CERFACS, ECMWF, INRIA and the Met Office
- At the Met Office this is used (3D var) in FOAM (deep and shelf ocean), Coupled DA (ocean component) and OSTIA (demo)
- Main benefits: Compatible with NEMO. It works with the ORCA grids natively. Efficient. There's an effective balance operator.
- Diffusion modelled background error covariances with multiple length scales (in 2D/3D)





EOFs (Empirical Orthogonal Functions)

Met Office

Change of variables.

Take a time series of field data $\mathbf{X}$ (columns are each time slice of a field $\mathbf{x}$). Separate into normalised orthogonal modes (Principal Components) with an associated variance and time series. Each mode maximises the variance explained.

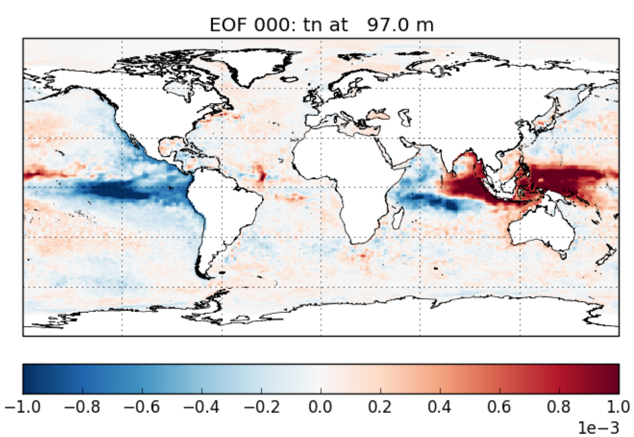
i.e $\mathbf{X} = \mathbf{E\Lambda P}^{\mathrm{T}}$ Single time i $\mathbf{x}_{\mathbf{i}} = \mathbf{E\Lambda a}_{\mathbf{i}}$

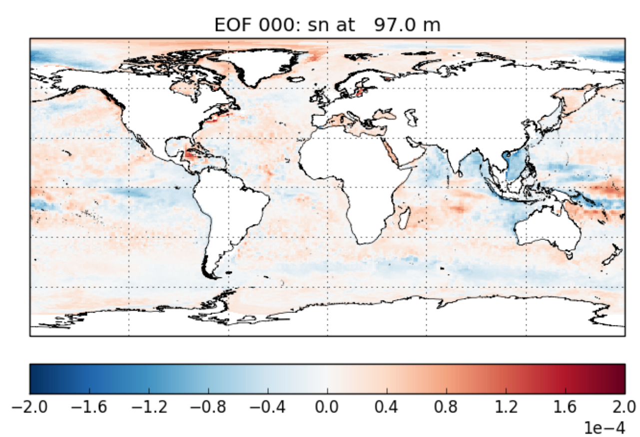
$\mathbf{E}$= matrix of EOFs (columns are each EOF pattern)

$\mathbf{\Lambda}$= diagonal matrix of eigenvalues associated with each EOF pattern

$\mathbf{P}$ = Matrix of time series for each pattern

$\mathbf{a}$= vector of coefficients/weights for each EOF ($x = \mathbf{E\Lambda a}$)





EOF 1 - 24.2% of the total variance

Testing EOF DA: One year objective analysis

- One year of monthly analyses
- Each month the assimilation background is from the GloSea5 climatology (for that month).
- EOFs calculated from the equivalent month of the GloSea5 reanalysis. Total of 20 EOFs for each month (excluding 2010 the test year)
- Assimilating subsampled modern day data (2010 data subsampled to 1953)
- Independent assessment of performance using statistics calculated from the set of modern day data not assimilated.

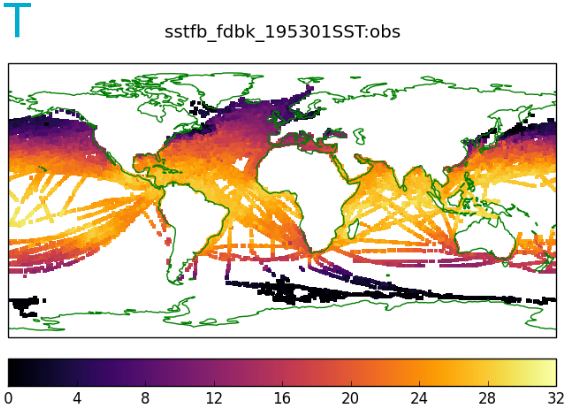
Bkg	No assim.
EOF	100% EOF
STD	100% standard
HYB1	10% EOF
HYB2	50% EOF
HYB3	90% EOF

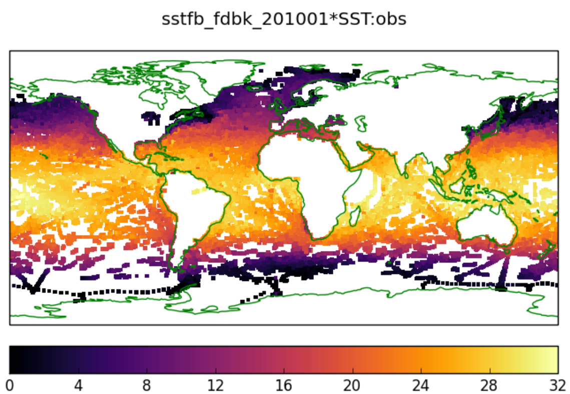
Observing system experiments

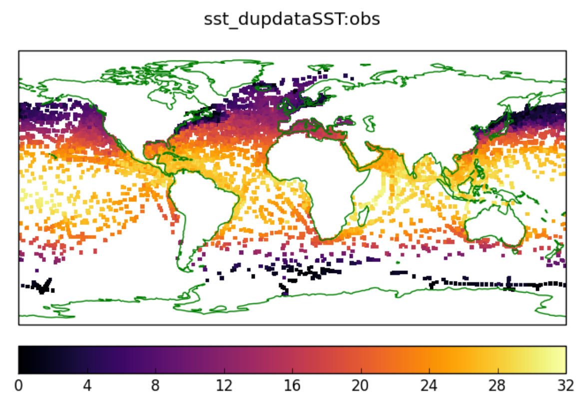
Subsample modern day observations to look like historical data (data from HadIOD)

Jan 1953 | Jan 2010 | 2010 data subsampled

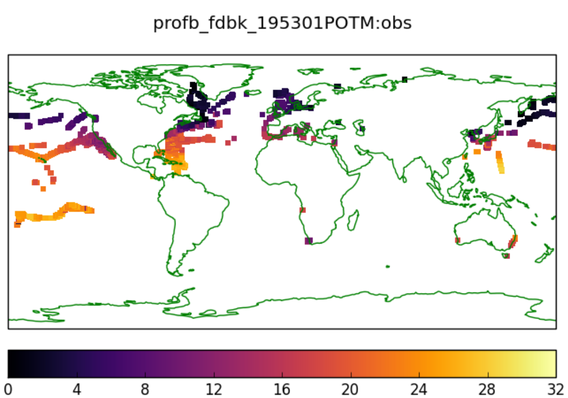
In-situ SST

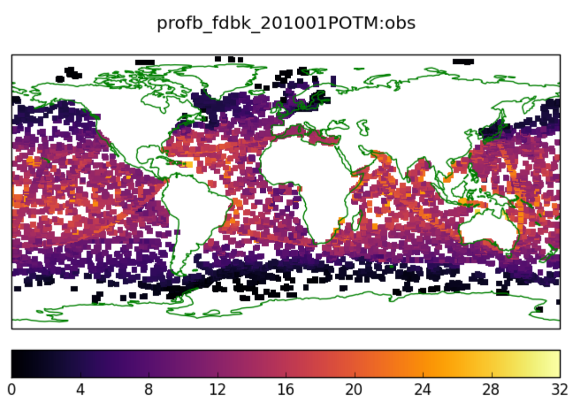


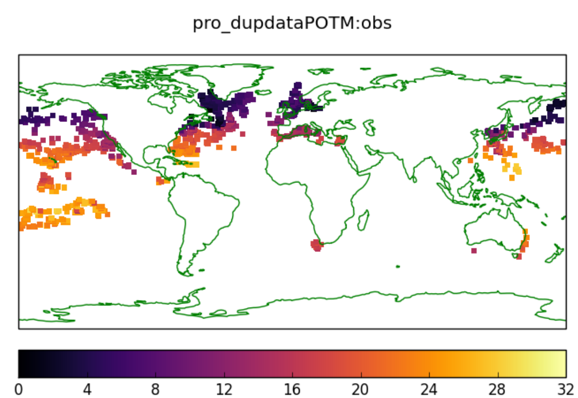




Profile T







One year objective analysis

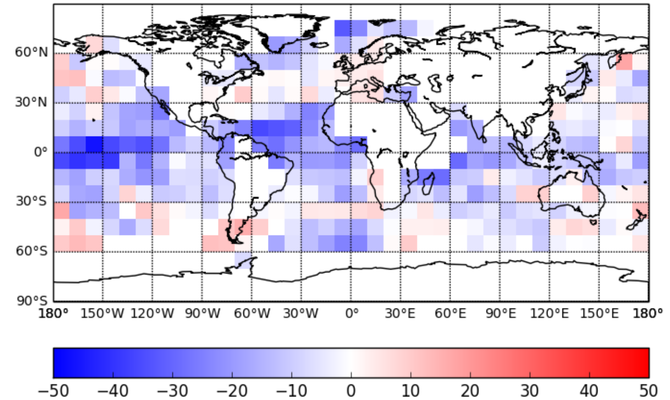
Expts assimilating subsampled SST & profile data

Results for the whole year
compared to unassimilated data

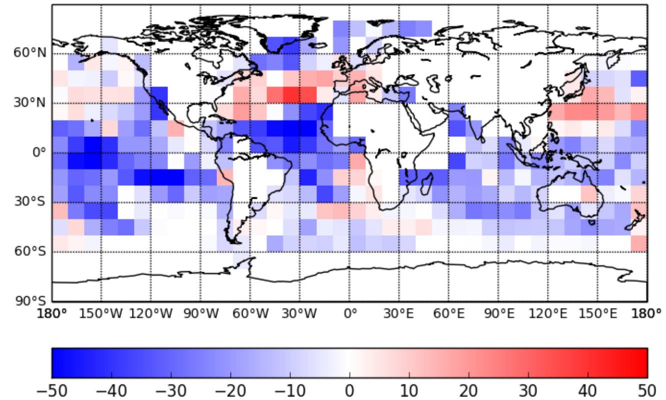
	SST RMS	T prof RMS	S prof RMS
bkg	1.036	1.158	0.233
EOF	0.955	1.132	0.216
Std	0.938	1.114	0.226
Hyb1 10% EOF	0.921	1.084	**0.209**
Hyb2 50% EOF	0.915	**1.084**	0.209
Hyb3 90% EOF	**0.910**	1.089	0.210

% reduction in SST error
compared to background

EOF



STD



Test in FOAM - a cycling system

EOFs from a coupled climate run (Met Office GC3.1) using 100 years of simulation

Gives 99 EOFs for each month - using top 80 EOFs (explains 90% of the variance).

System adapted the FOAM reanalysis system which cycles daily.

Assimilates 2015 data subsampled to 1953 distribution

Expts	
STD	Standard FOAM system
EOF02	Hybrid 1% EOF
EOF03	Hybrid 5% EOF

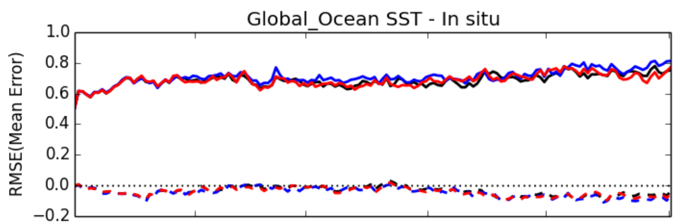
Cycling system Preliminary results

Met Office

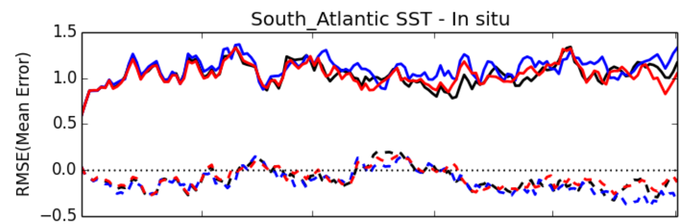
2015 run

Forecast vs observations (including those not assimilated)

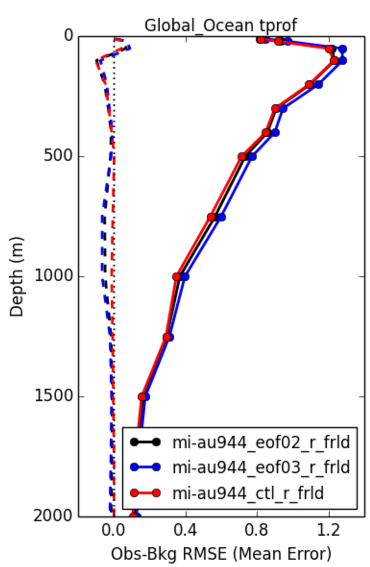
STD	Standard FOAM system
EOF02	Hybrid 1% EOF
EOF03	Hybrid 5% EOF

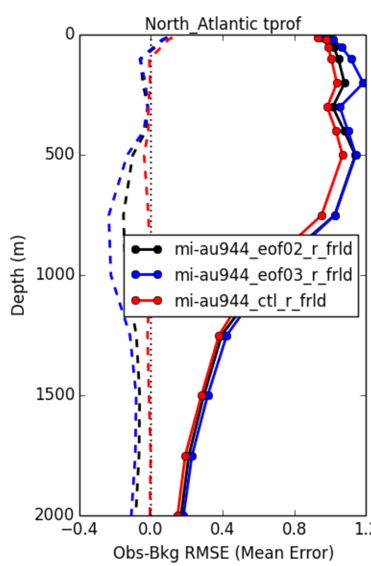


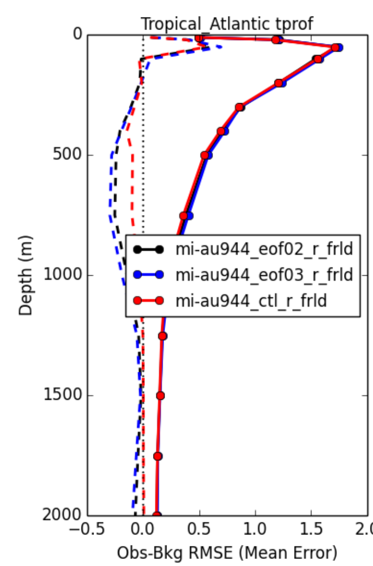
1 Jan 30 Jun



1 Jan 30 Jun







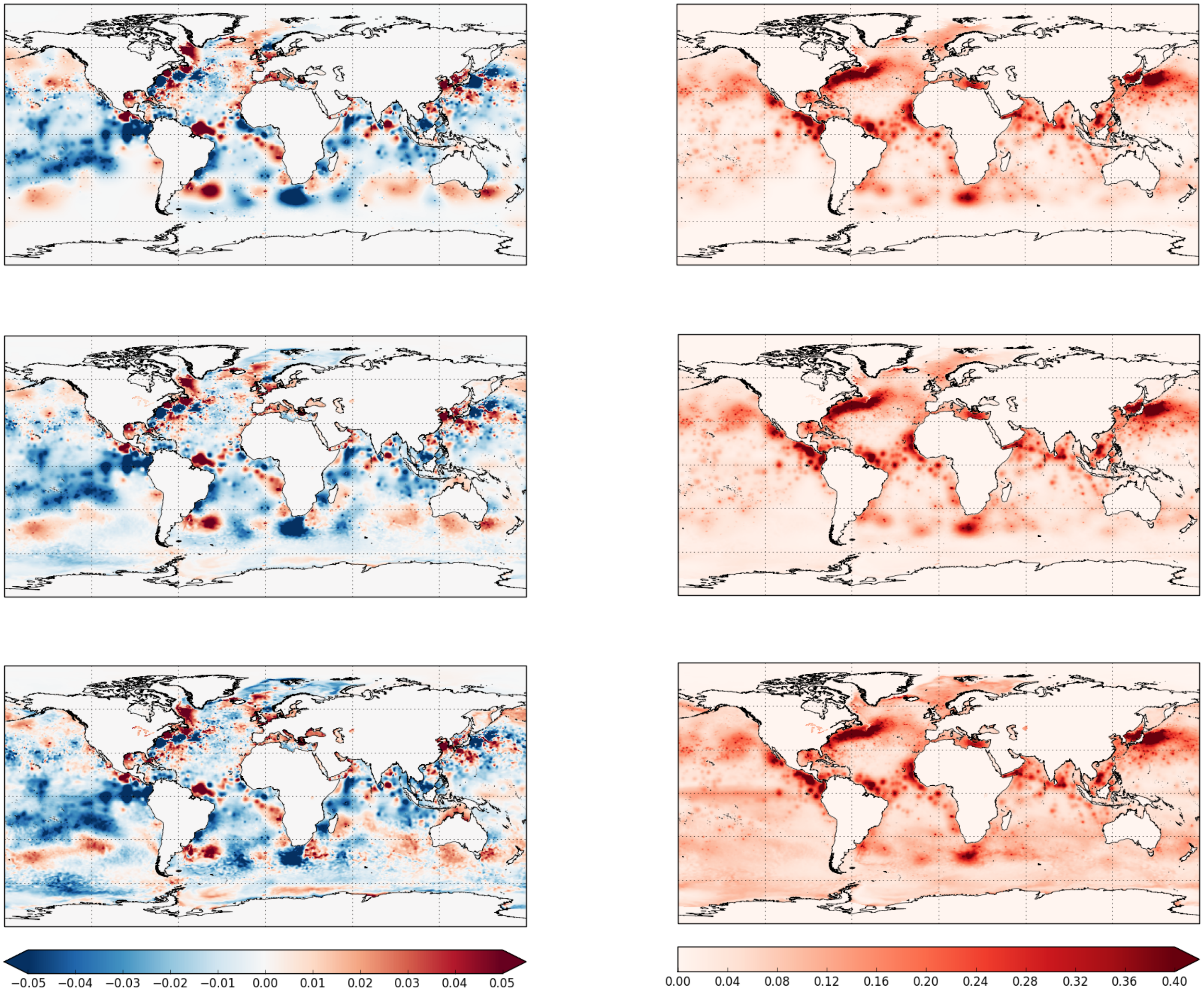
Jan 2015 increments

mean | std dev

Std

Hyb EOF 1% (EOF02)

Hyb EOF 5% (EOF03)



Summary/plans for EOF DA

- Developed a method for EOF DA in NEMOVAR
- A hybrid of EOF DA and standard DA gives good results overall perhaps better than either EOF only or standard assimilation
- Preliminary tests of EOF DA in a reanalysis system (FOAM like) is mostly promising
- The correct setting of the hybrid weights looks important

- Plan to further explore the hybrid EOF DA including understanding aspects where it is not working so well
- EOF DA may be particularly unsuited to correcting biases in the model
- Plan to test EOF DA in combination with a T-S bias correction scheme

- Compare EOF DA and hybrid DA to the Met Office decadal prediction system (DePreSys) (initial work looks promising)

Met Office

Thank you

Assimilation cost function

Met Office

Working in model space (Std)

$J(\delta \mathbf{x}) = \frac{1}{2}\, \delta \mathbf{x}^T\, \mathbf{B}^{-1}\, \delta \mathbf{x}$

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

Working in EOF space

$\delta \mathbf{x} = \mathbf{E}\boldsymbol{\Lambda}\delta \mathbf{a}$

$\mathrm{J}(\boldsymbol{\delta}\mathbf{a}) = \frac{1}{2}\, \delta \mathbf{a}^T\, (\boldsymbol{\Lambda}^2)^{-1}\, \delta \mathbf{a}$

$+\ \frac{1}{2}\, (\mathbf{y}-\mathbf{H}(\mathbf{x}_b+\mathbf{E}\boldsymbol{\Lambda}\delta \mathbf{a}))^T\, \mathbf{R}^{-1}\, (\mathbf{y}-\mathbf{H}(\mathbf{x}_b+\mathbf{E}\boldsymbol{\Lambda}\delta \mathbf{a}))$

Hybrid

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

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

a= vector of coefficients/weights for each EOF (model field anomaly = $\mathbf{E}\boldsymbol{\Lambda}\mathbf{a}$)

E= EOF patterns matrix

Λ= diagonal matrix of eigenvalues associated with each EOF pattern

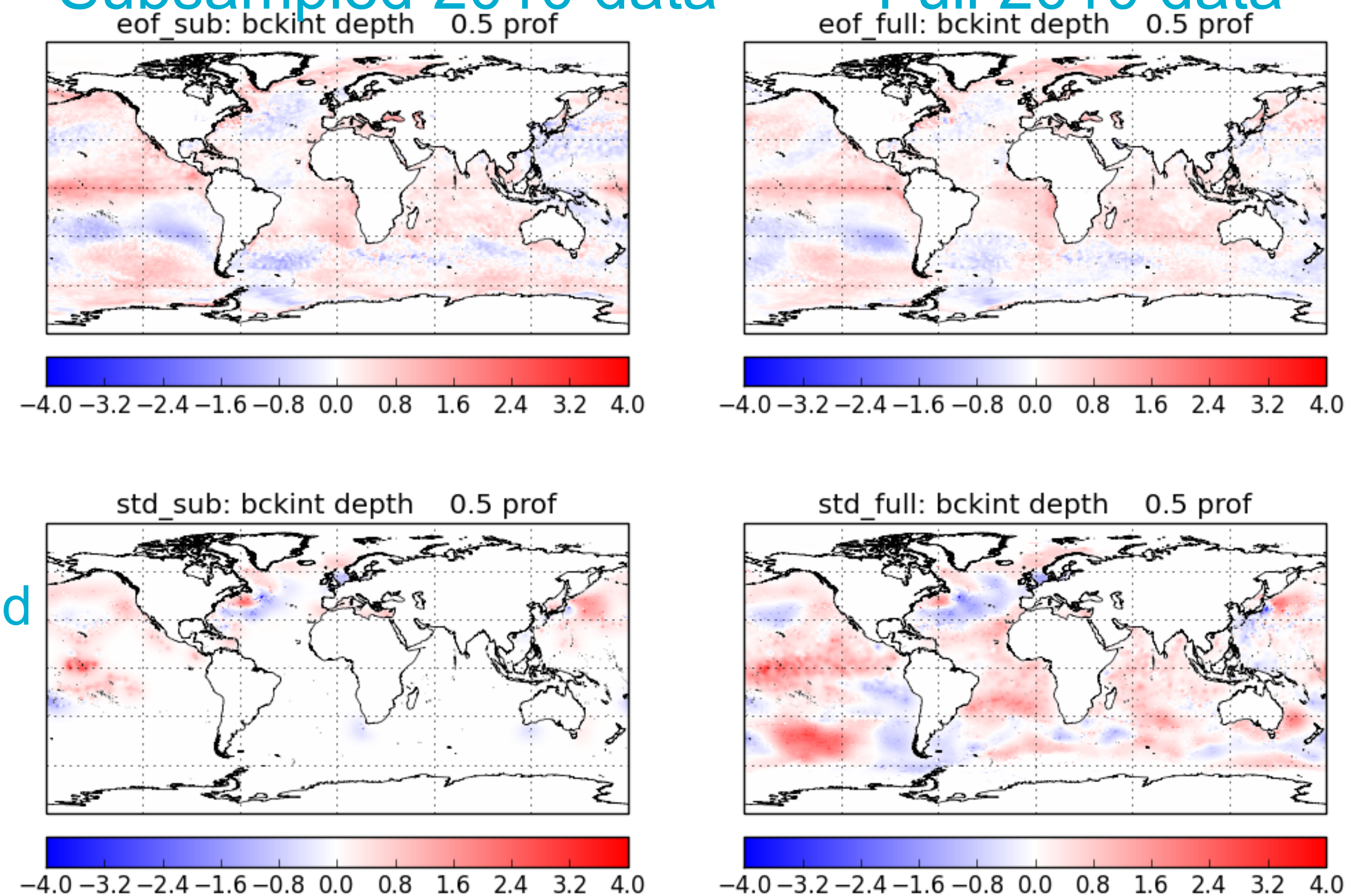
Ocean surface temperature increments /ºC (profile data only assimilated)

Subsampled 2010 data

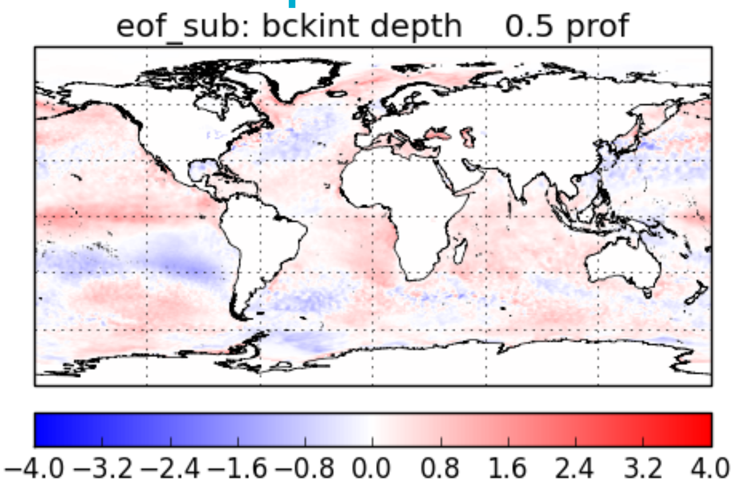
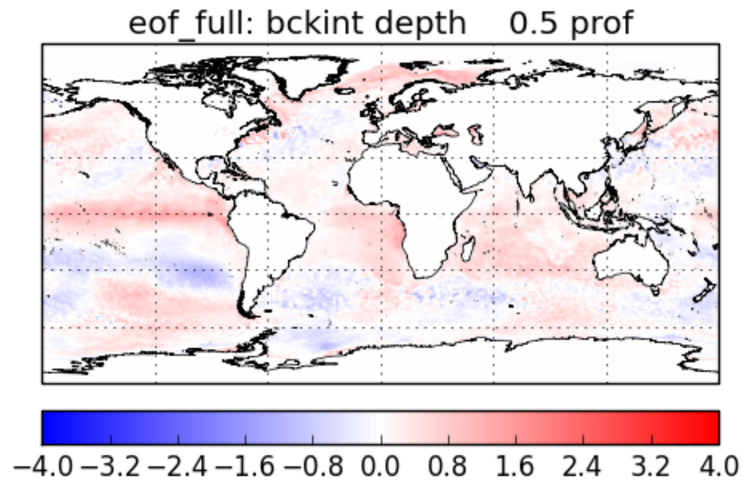
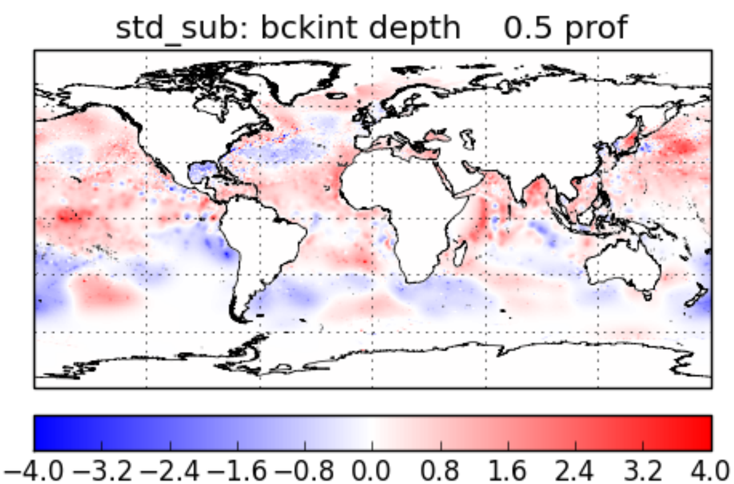
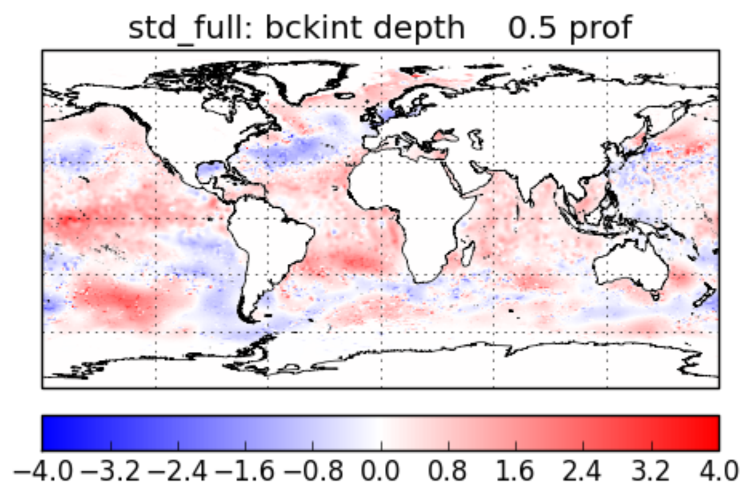
Full 2010 data

EOF DA

Standard DA

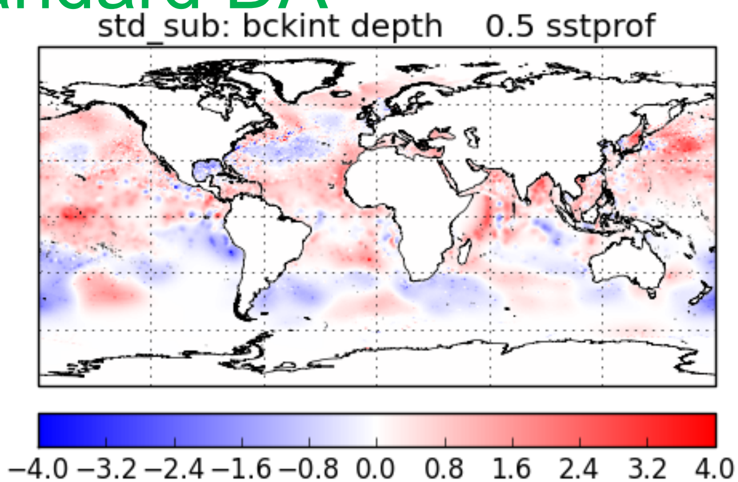


Ocean surface temperature increments /ºC (SST & profile data assimilated)

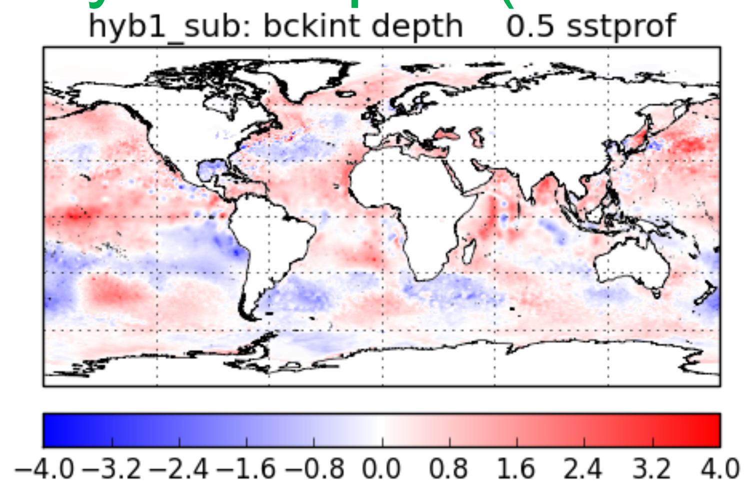
	Subsampled 2010 data	Full 2010 data
EOF DA		
Stand ard DA		

Ocean surface temperature increments /ºC (Subsampled SST & profiles assimilated)

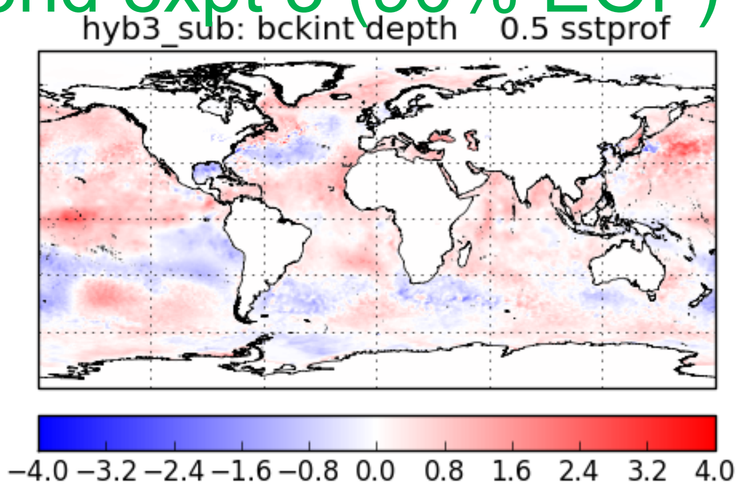
Standard DA



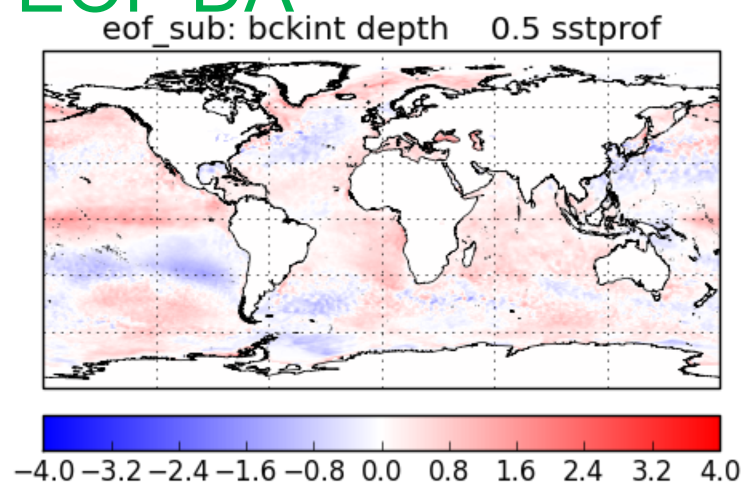
Hybrid expt 1 (10% EOF)



Hybrid expt 3 (90% EOF)

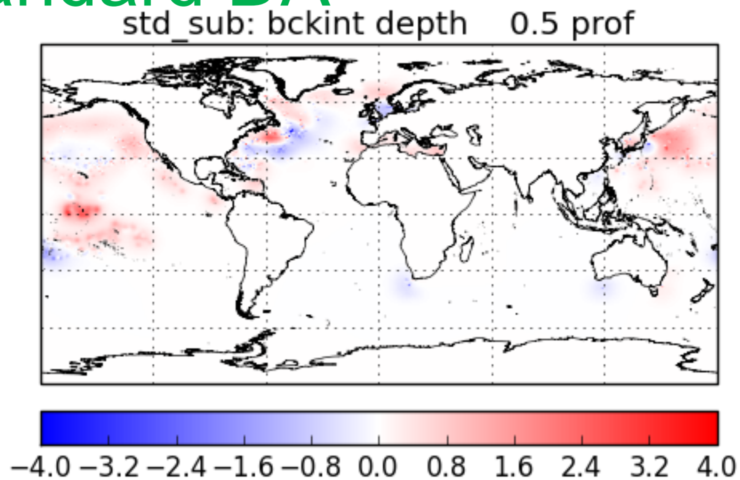


EOF DA

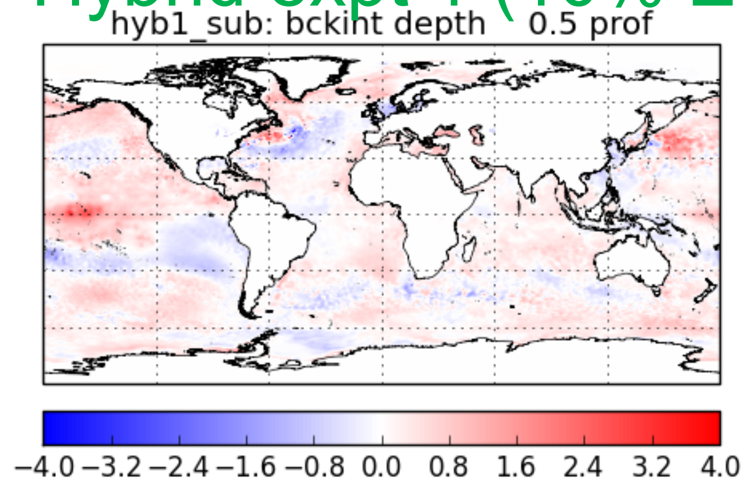


Ocean surface temperature increments /ºC (Subsampled profiles only assimilated)

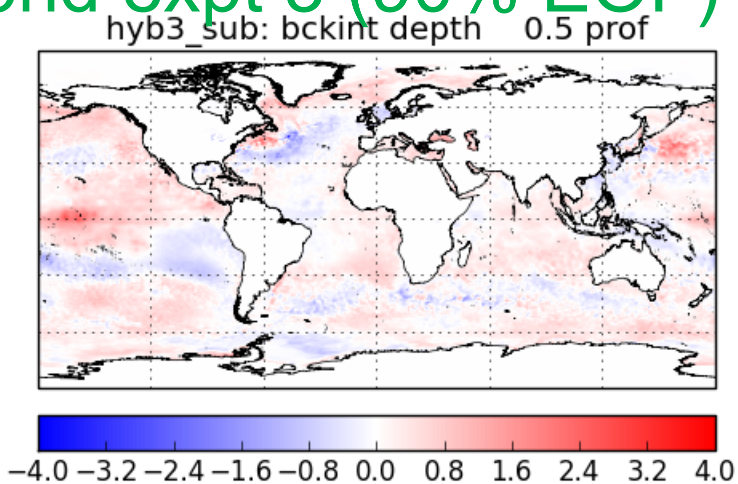
Standard DA



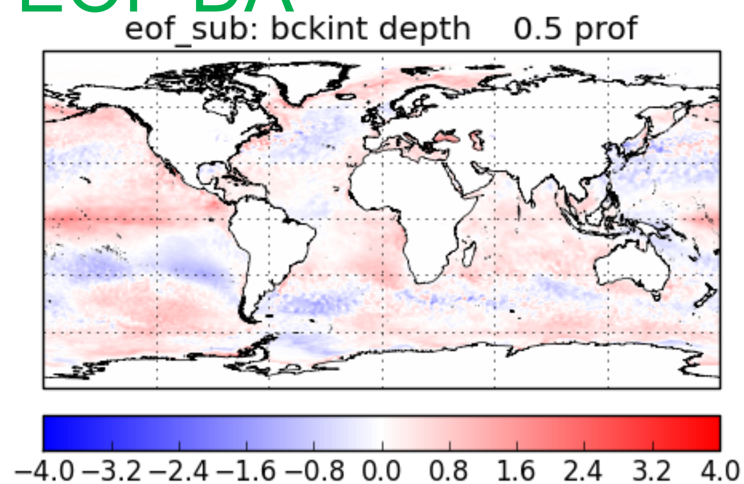
Hybrid expt 1 (10% EOF)



Hybrid expt 3 (90% EOF)



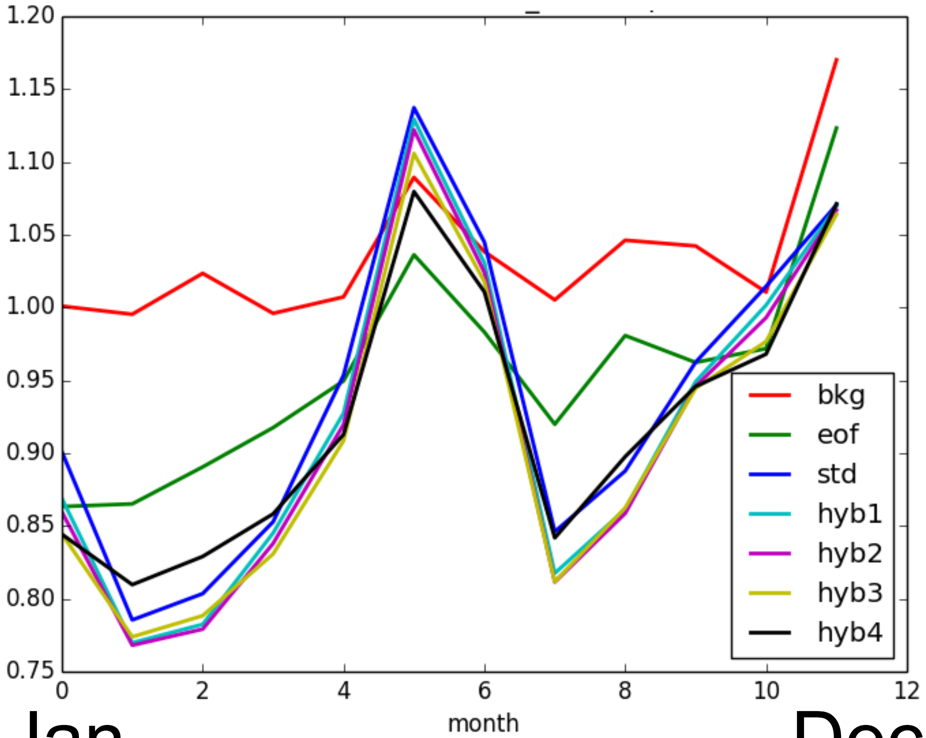
EOF DA



One year objective analysis assimilating subsampled SST and profiles

RMS SST error / ºC (vs. unassimilated data)

Bkg	No assim
EOF	100% EOF
STD	100% standard
HYB1	10% EOF
HYB2	50% EOF
HYB3	90% EOF
HYB4	99% EOF



Jan

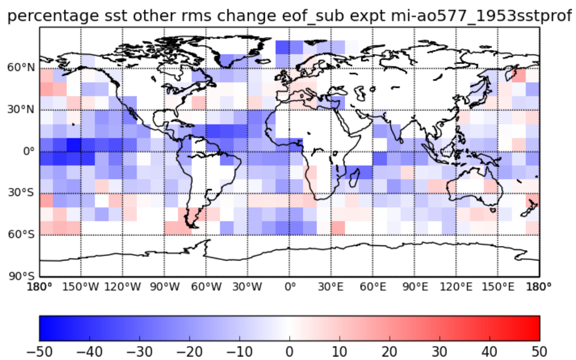
Dec

One year run

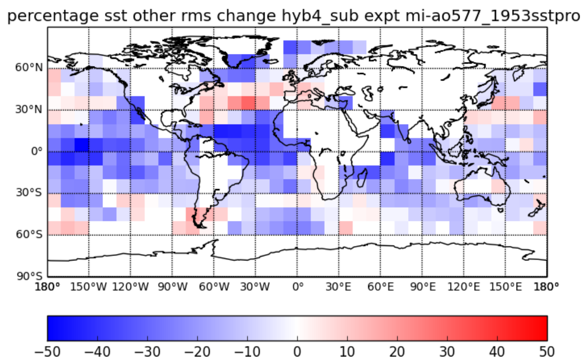
Expts assimilating SST & profile data

% reduction in SST error compared to background

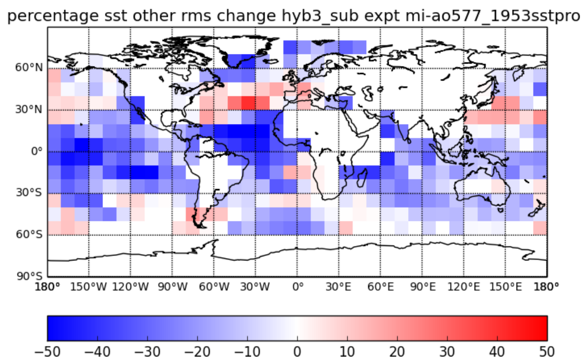
EOF



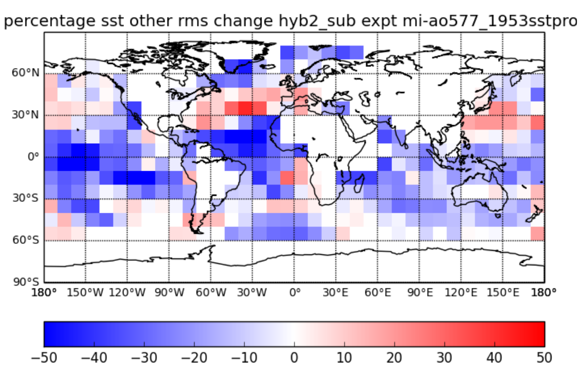
Hyb 99% EOF



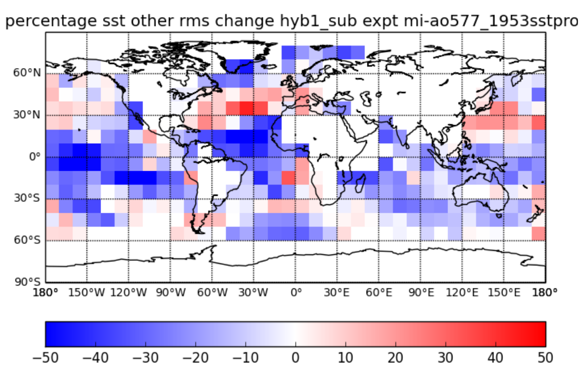
Hyb 90% EOF



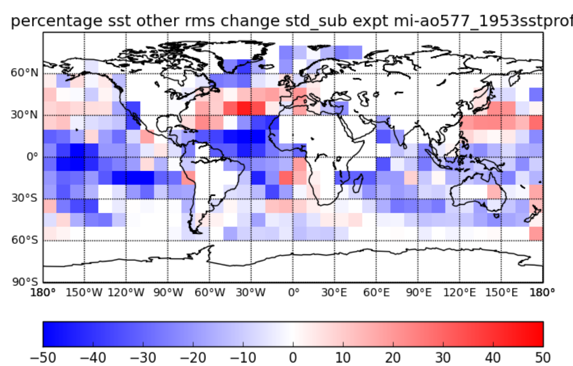
Hyb 50% EOF



Hyb 10% EOF



Std

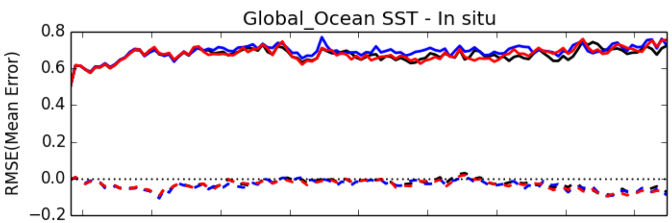


Cycling system
4 month results

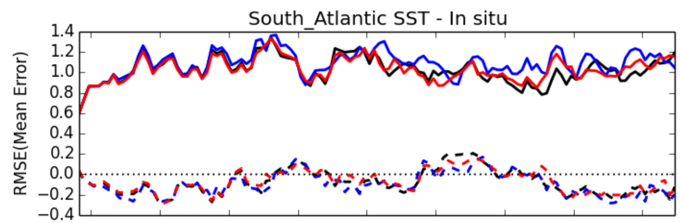
2015 run

Forecast vs observations (including those not assimilated)

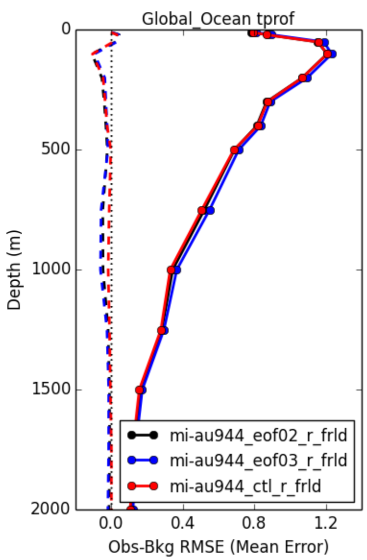
STD	Standard FOAM system
EOF02	Hybrid 1% EOF
EOF03	Hybrid 5% EOF

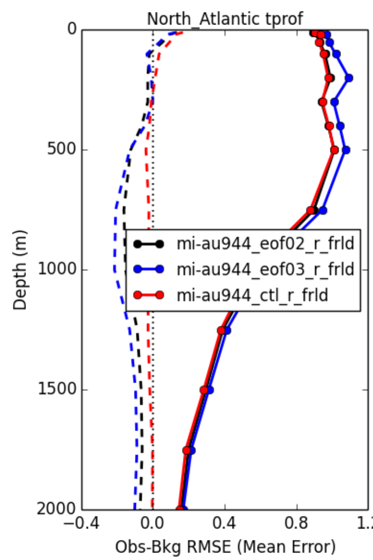


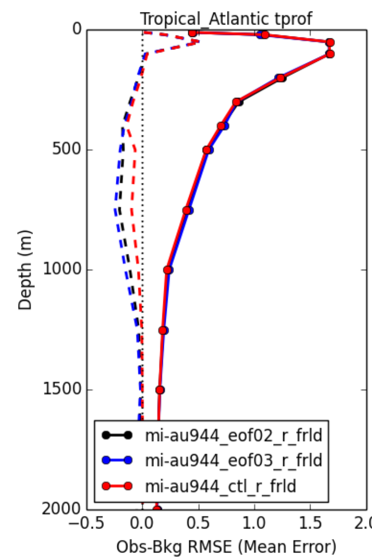
1 Jan 30 April



1 Jan 30 April

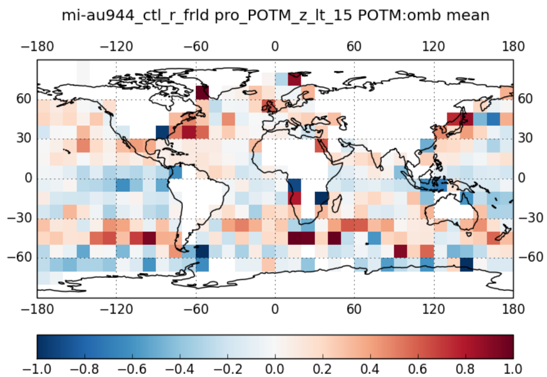




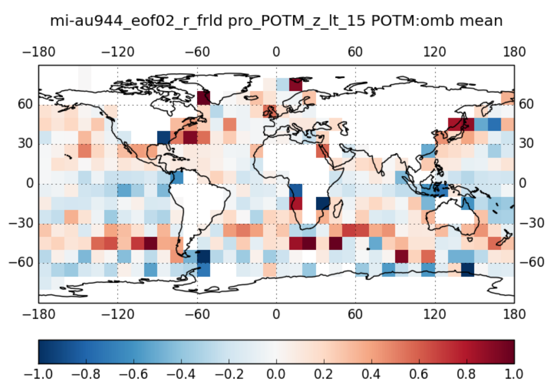


Jan 2015 mean innovations

Ctl



Hyb
EOF
(EOF02)



Change in mean T innovations

EOF - Ctl