# Representation Error in Ocean Data Assimilation

Robert N. Miller
James G. Richman
David Behringer
Yvette H. Spitz

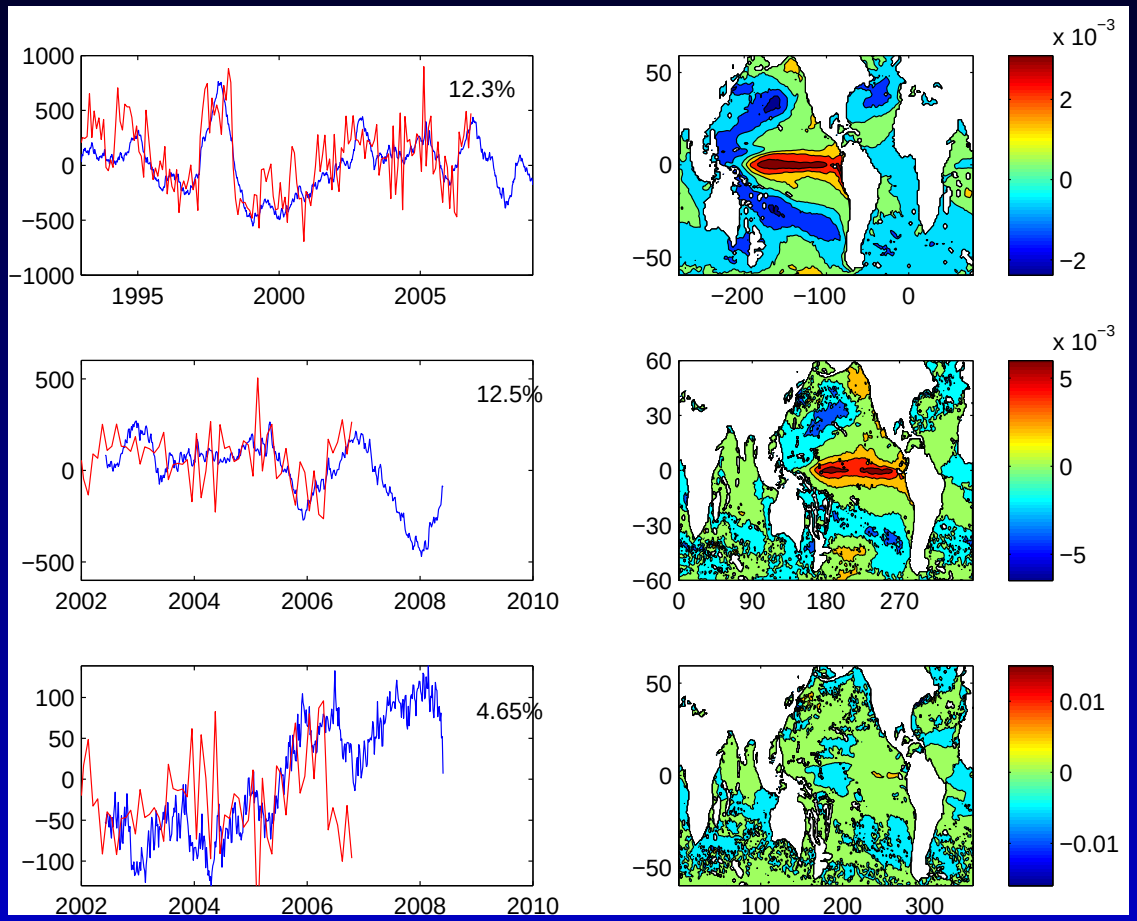
# Outline

- Data assimilation works by using model-data misfits to correct the model state of the system
- The causes of some of the observed variability are not reflected in the model, and that portion of the observed variability cannot be usefully assimilated
- We propose a method for constructing statistical error estimates that account for representation error explicitly
- We describe the results of our first implementation of our methods within the framework of the operational climate forecast system.

# Estimating Representation Error

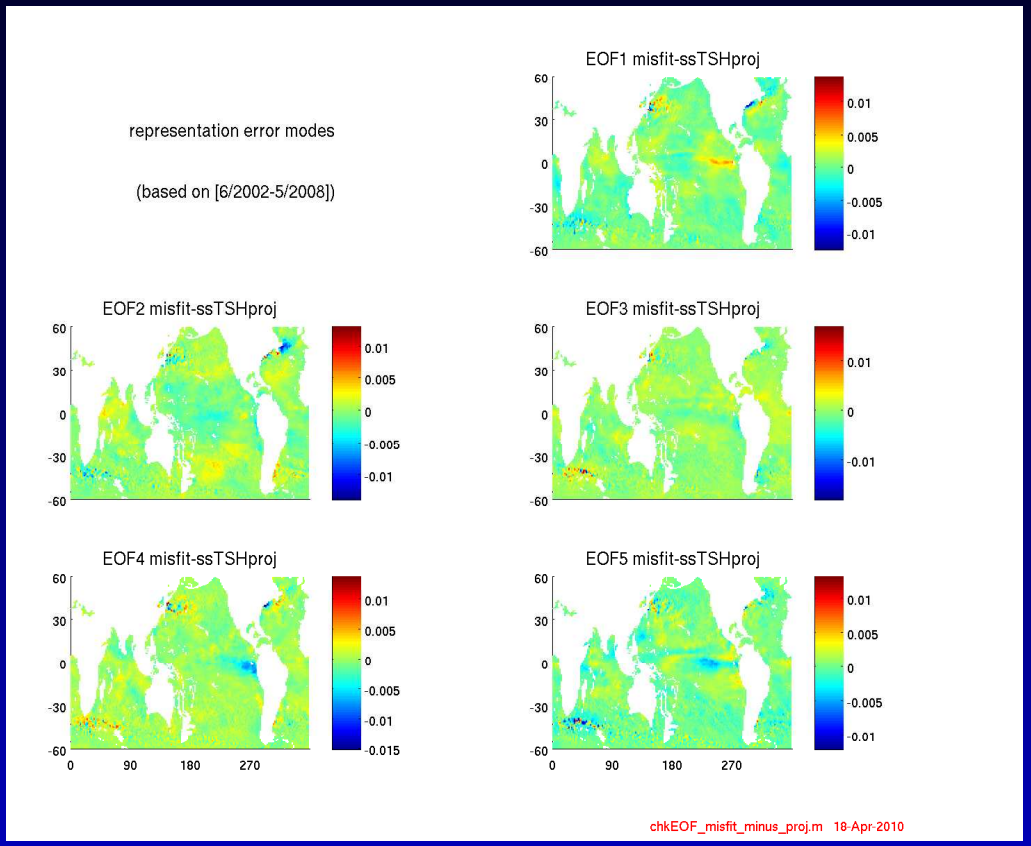
1. Compute multivariate EOFs of a 16 year run of the ocean component of CFS
2. Estimate number of significant degrees of freedom (DOF) by the Preisendorfer test. The span of the significant DOF is the "model space"
3. Project a series of model-data misfits into the model space. Assume: innovations - their projections on the model space = instrument error + representation error.
4. The significant EOFs of the the innovations - their projections into model space are assumed to span the space of representation errors.

# Surface Variability




Lead PCs and EOFs of SST variability. Top to bottom: SST variability derived from surface multivariate EOF; PC and EOF of satellite analysis; PC and EOF of misfit

# Representation Error EOFs




Preisendorfer test shows 106 significant EOFs, about 75% of the total misfit variance.

# The New Assimilation Scheme

- Assimilate data by 3DVAR:

$$
\begin{aligned}
J &= \frac{1}{2}(\mathbf{x}-\mathbf{x}^{(b)})^T E^{-1}(\mathbf{x}-\mathbf{x}^{(b)}) + \\
&\quad \frac{1}{2}(z-H\mathbf{x})^T F^{-1}(z-H\mathbf{x}) \\
\delta J &= \delta\mathbf{x}^T\left(E^{-1}(\mathbf{x}-\mathbf{x}^{(b)}) - H^T F^{-1}(z-H\mathbf{x})\right) \\
\nabla J &= E^{-1}(\mathbf{x}-\mathbf{x}^{(b)}) - H^T F^{-1}(z-H\mathbf{x})
\end{aligned}
$$

- Obs error covariance $F$ appears only in the term $H^T F^{-1}(z-H\mathbf{x})$
- Write $F = D + uu^T$ where $D$ is diagonal and columns of $u$ span representation error space.

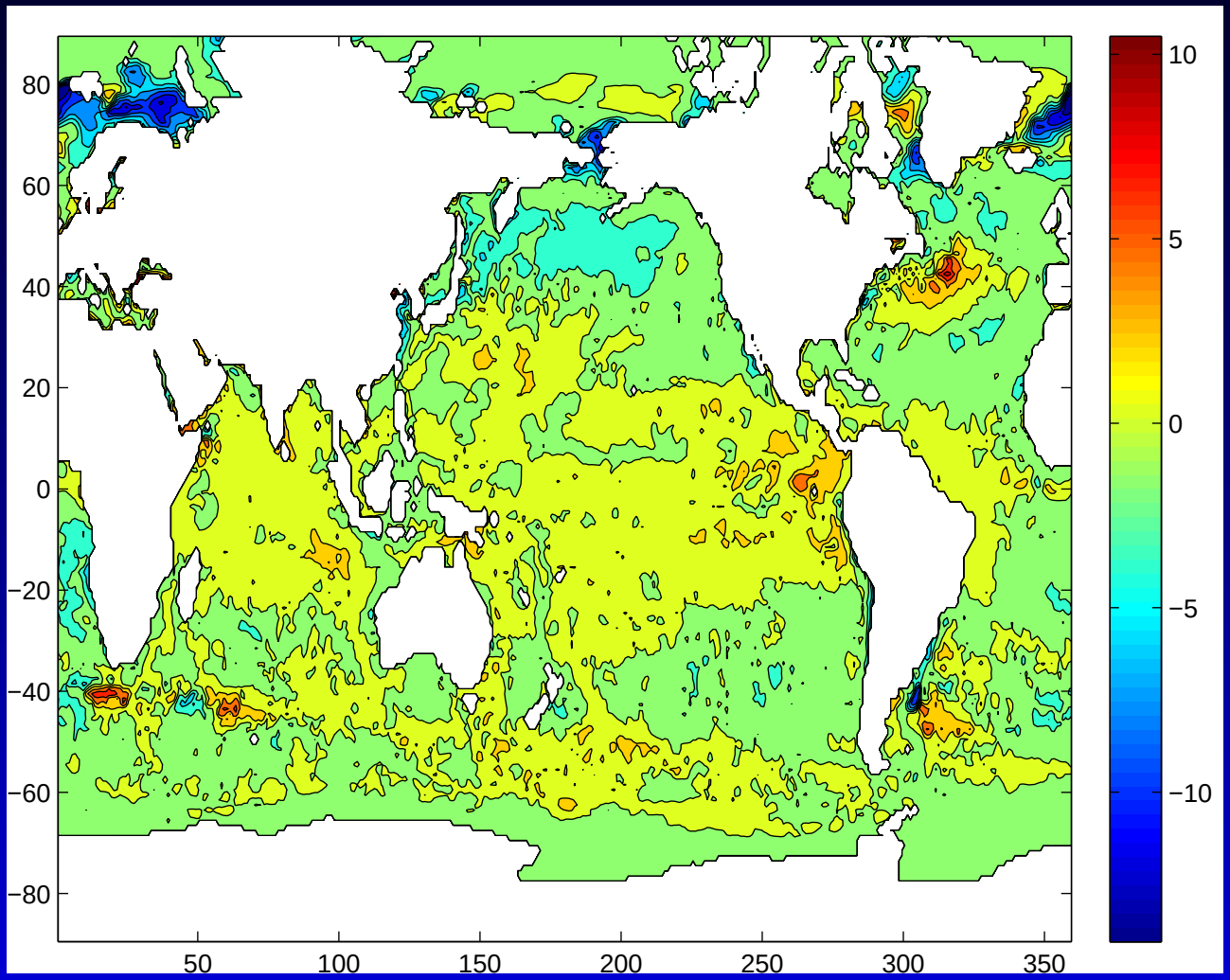
# Effect of Augmentation of $F$

- Calculation of $F^{-1}$ will be inexpensive.
- Example: a 1D representation error subspace
- Let $F = \sigma^2 I + uu^T$, $\sigma^2$ the obs error variance and $u$ a vector with $u^T u = r^2 > \sigma^2$. The Sherman-Morrison formula:

$$(\sigma^2 I + uu^T)^{-1} = \frac{1}{\sigma^2}\left(I - \frac{uu^T}{\sigma^2 + u^T u}\right)$$

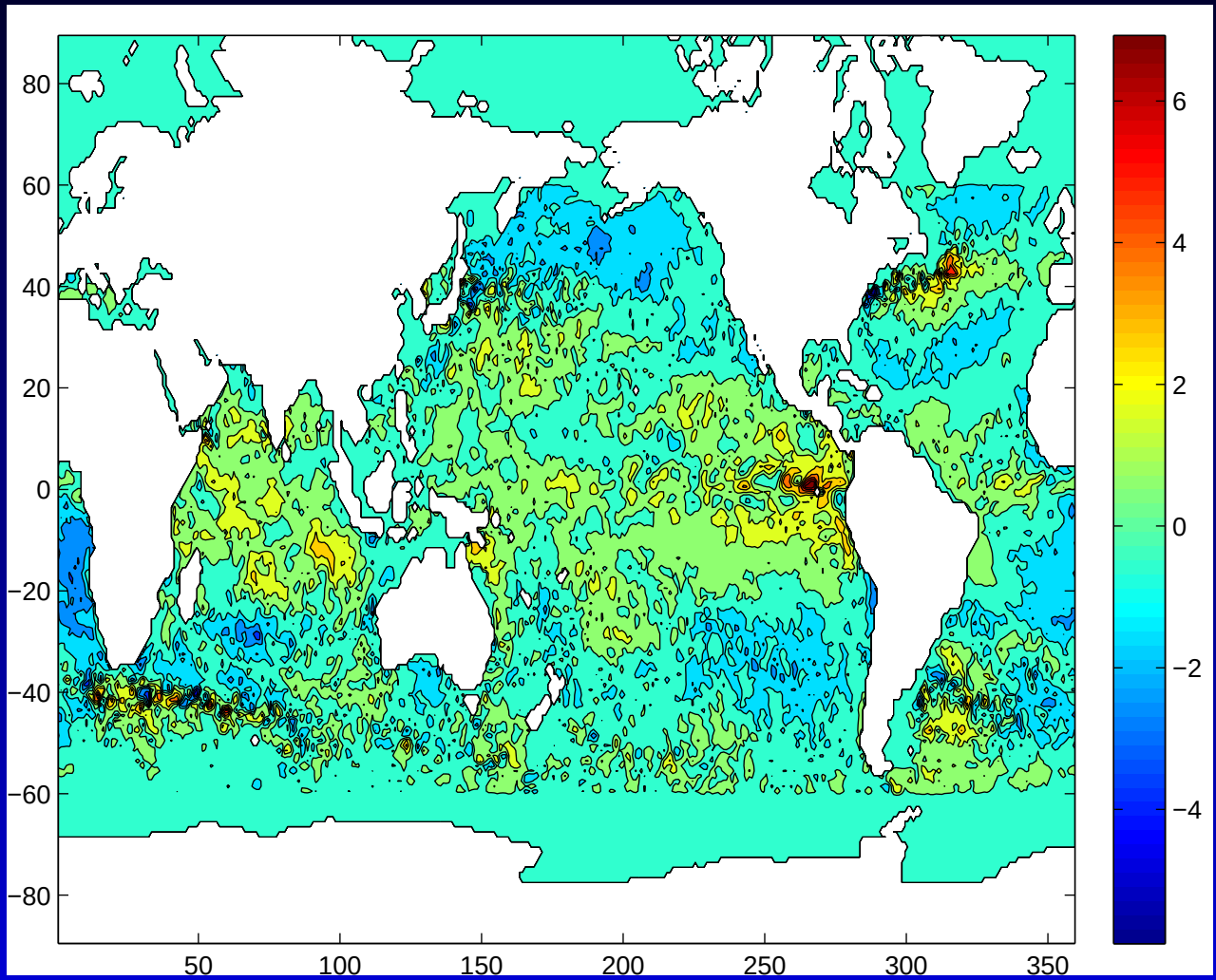
- Multiplying the innovation by $F^{-1}$ will thus have the effect of damping the component parallel to $u$ by a factor of $\sigma^2/(r^2 + \sigma^2) << 1$
- We work with 106 basis vectors and apply the generalized S-M formula

# Results from New 3DVAR



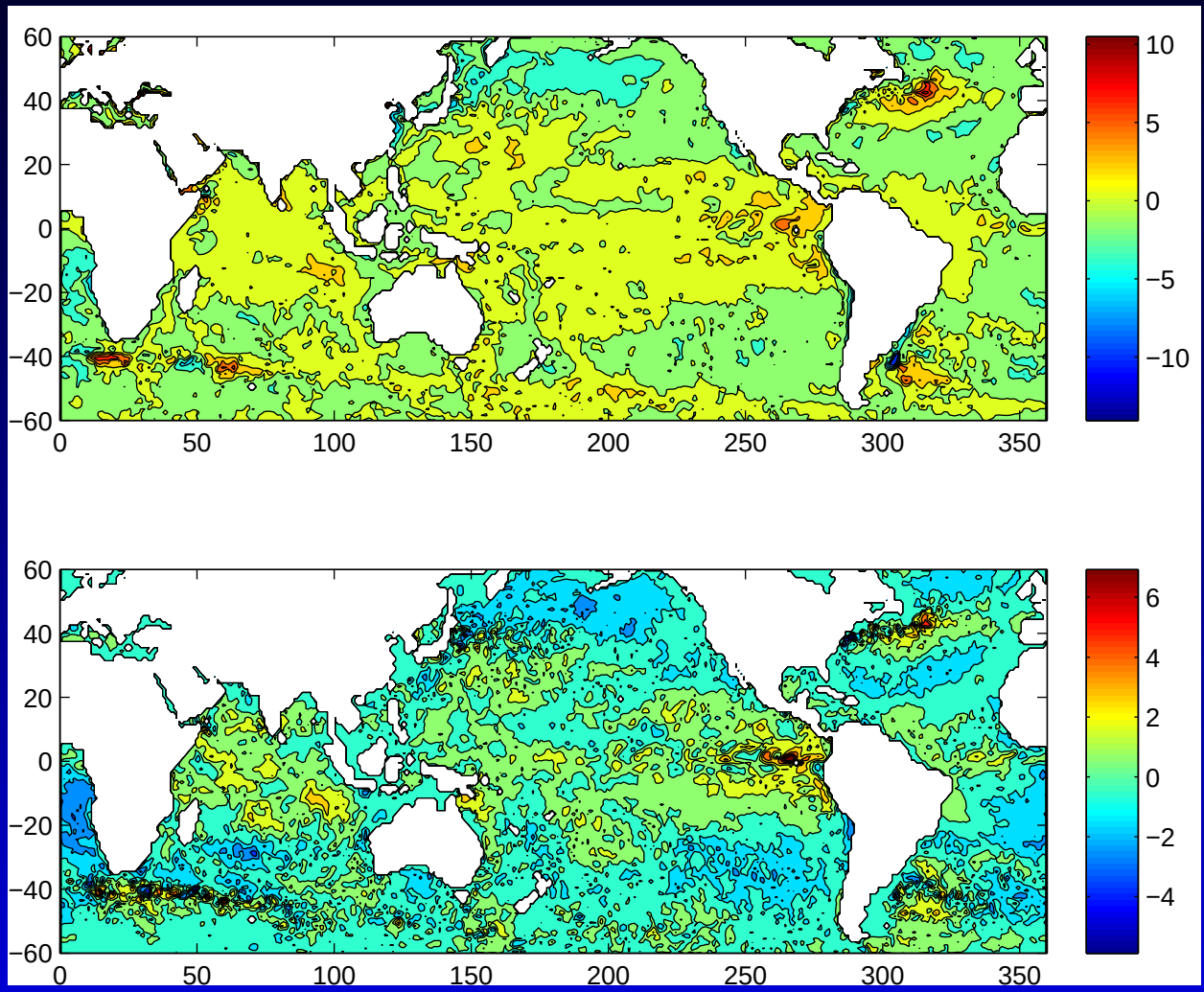

SST misfit on June 30, 2003, following a 6 month run

# Projection into Representation Error Space



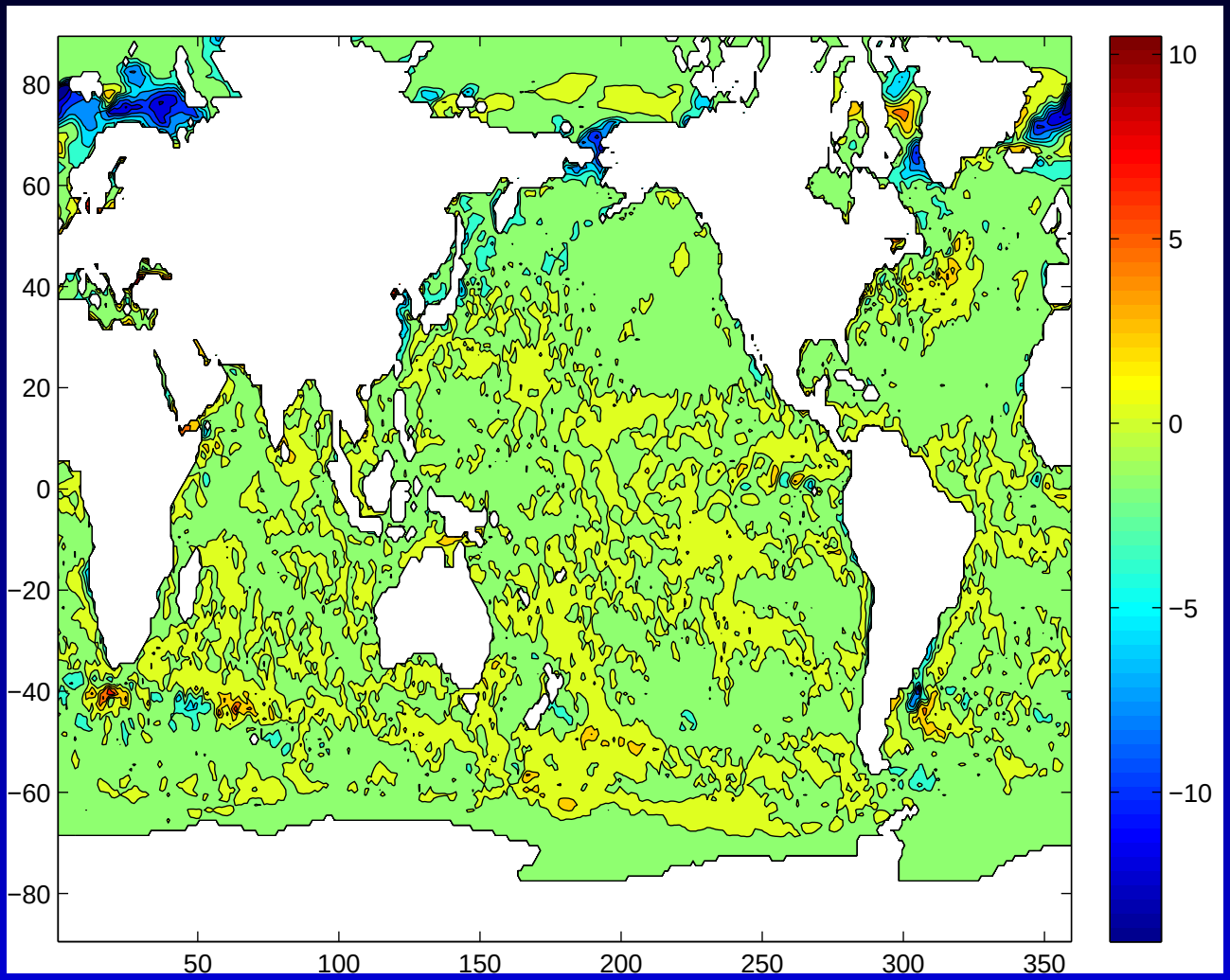

Best fit of a linear combination of representation error EOFs to misfit. $R^2 = 57\%$ Note difference in scales.

# Projection into Representation Error Space

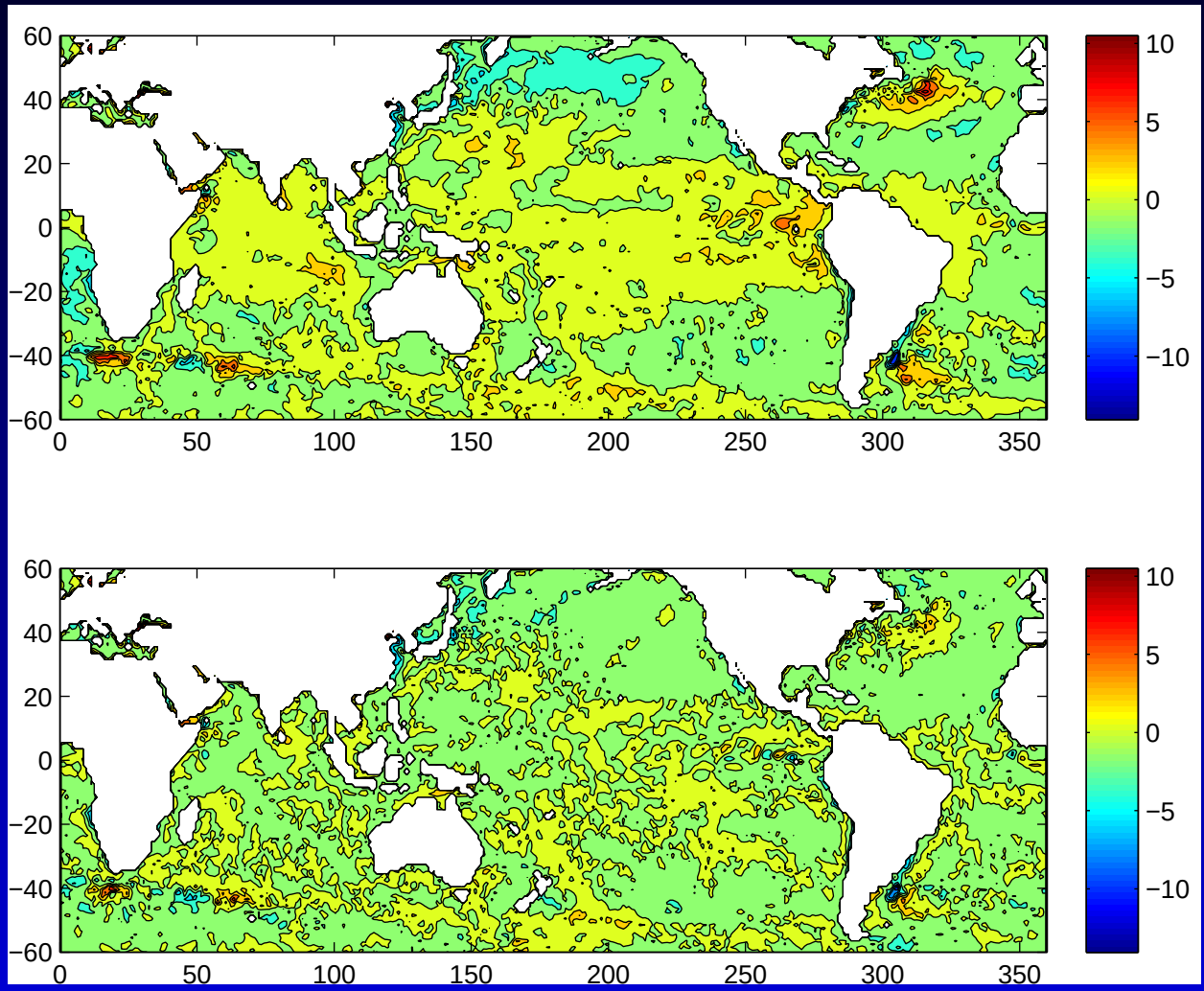



Best fit of a linear combination of representation error EOFs to misfit. $R^2 = 57\%$

# Misfit-Representation Error



# Misfit-Representation Error




Top: misfit; Bottom: misfit-projection into EOF space

# Summary

- We have devised a methods for estimating the statistics of representation error for OGCMs
- Following a first experiment in the north Pacific, we have incorporated our representation error estimates into a 3DVAR data assimilation scheme applied to $0.5^o$ global MOM4
- First results show the scheme behaving as expected, i.e., the representation error accounts for a significant fraction of the variance of the misfit
- The patterns assigned to representation error are evidently not participating in the assimilation

# Next Steps

- Simulate representation error as an autoregressive process
- Model SST as a sum of model output and simulated representation error
- Ensemble experiments
- The EnKF
- Extend to use with altimeter data

# nagging worries

- drifts and biases
- underestimate of model capabilities due to limitations of available data sets

# Future directions

- Extension to estimation of model error
- Investigation of response of model atmosphere
- Climate applications