Ensemble-Based Data Assimilation in POP using DART

B.T. Nadiga

OMWG, January 17, 2014

Nadiga, Casper, and Jones. "Ensemble-based global ocean data assimilation." Ocean Modelling 72 (2013): 210-230.

Outline

Brief Introduction

Problem Formulation

Results

- SST Changes due to Assimilation
- Sub-Surface Changes due to Assimilation
- Utility of Assimilations for Initialization?

Conclusions

Outline

Brief Introduction

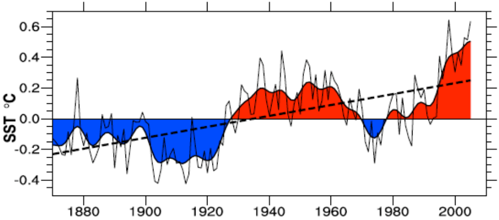
Problem Formulation

Results

- SST Changes due to Assimilation
- Sub-Surface Changes due to Assimilation
- Utility of Assimilations for Initialization?

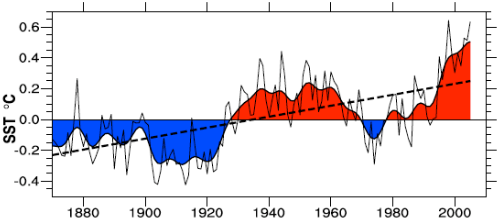
Conclusions

Climate Variability

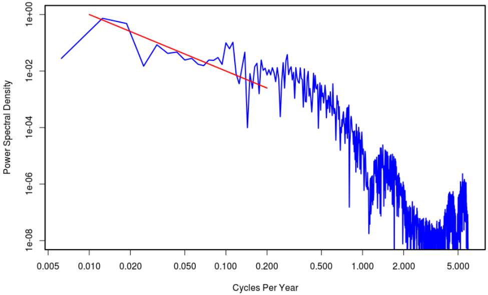


Annual SST anomalies averaged over the North Atlantic (0 to 60 N, 0 to 80 W) for 1870–2005, relative to 1901–1970 (C) (Trenberth & Shea, 2006)

Climate Variability



Annual SST anomalies averaged over the North Atlantic (0 to 60 N, 0 to 80 W) for 1870–2005, relative to 1901–1970 (C) (Trenberth & Shea, 2006)



Inter-annual Variability (1856-2013)

Slow variations *suggest* but *do not* imply extended predictability:

$$\frac{\partial T}{\partial t} = -\gamma T + \epsilon \Rightarrow$$

$$S_T(\omega) = \frac{S_\epsilon}{\omega^2 + \gamma^2} \quad (\text{Red})$$

(Hasselman 1976)

This Talk

- How well can the ocean state be initialized with DA?
- How useful is the initialization? (seasonal fore., climate pred.)

Differences from Atmospheric DA

- More strongly coupled (def. rad $\approx$ 30 km)
- Wider range of timescales
- More difficult to observe
- Surface forcing, balance

This Talk

- How well can the ocean state be initialized with DA?
- How useful is the initialization? (seasonal fore., climate pred.)

Differences from Atmospheric DA

- More strongly coupled (def. rad $\approx$ 30 km)
- Wider range of timescales
- More difficult to observe
- Surface forcing, balance

Data Assimilation

Model Evolution:

$$\mathbf{x}^{(f)}_{n+1} = \mathbf{M}(\mathbf{x}^{(a)}_{n})$$

Model Error:

$$\mathbf{x}^{(t)}_{n+1} = \mathbf{M}(\mathbf{x}^{(t)}_{n}) + \epsilon_{\mathbf{M}}$$

Forecast Error:

$$\epsilon^{(f)}_{n+1} = \mathbf{M}(\mathbf{x}^{(a)}_{n}) - \mathbf{M}(\mathbf{x}^{(t)}_{n}) - \epsilon_{\mathbf{M}}$$

$$= M(x^{(t)}_{n} + \epsilon^{(a)}_{n}) - M(x^{(t)}_{n}) - \epsilon_{M}$$

DA corrects for both by periodic re-initialization

$$p(E|O) = \frac{p(O|E)p(E)}{p(O)}$$

Data Assimilation

Model Evolution:

$$\mathbf{x^{(f)}_{n+1}} = \mathbf{M}(\mathbf{x^{(a)}_{n}})$$

Model Error:

$$\mathbf{x^{(t)}_{n+1}} = \mathbf{M}(\mathbf{x^{(t)}_{n}}) + \epsilon_{\mathbf{M}}$$

Forecast Error:

$$\epsilon^{\mathbf{(f)}}_{\mathbf{n+1}} = \mathbf{M}(\mathbf{x^{(a)}_{n}}) - \mathbf{M}(\mathbf{x^{(t)}_{n}}) - \epsilon_{\mathbf{M}}$$

$$= M(x_n^{(t)} + \epsilon_n^{(a)}) - M(x_n^{(t)}) - \epsilon_M$$

DA corrects for both by periodic re-initialization

$$p(E|O) = \frac{p(O|E)p(E)}{p(O)}$$

Ensemble integration: $\mathbf{x}^{(f)}$
Sample forecast error covariance:

$$\mathbf{P}^{(f)} \approx \mathbf{X}^{(f)}\mathbf{X}^{(f)\,\mathbf{T}}$$

Analysis of ensemble mean:

$$\bar{\mathbf{x}}^{(a)} = \bar{\mathbf{x}}^{(f)} + \mathbf{K}(\mathbf{y}^{(o)} - \mathbf{H}\bar{\mathbf{x}}^{(f)})$$

$$\mathbf{K} = \mathbf{P}^{(f)}\mathbf{H}^{\mathbf{T}}(\mathbf{H}\mathbf{P}^{(f)}\mathbf{H}^{\mathbf{T}} + \mathbf{R})^{-\mathbf{1}}$$

Analysis error covariance:

$$\mathbf{P}^{(a)} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^{(f)}$$

Choose analysis perturbations consistent with analysis error covariance—Different flavors

With no assimilation, trajectories diverge (IC Error)

With assimilation, realistic trajectories (IC Error)

Outline

Control Ensemble Integration and Problem Formulation

- Performed ensemble integrations of a standard configuration of POP with interannually varying forcing CORE v2.
- Control ensemble contains practically none of the WOD09 T & S observations
- Time averaged difference of ensemble-mean SST and NOAA OI SST v2 shows large errors.

Can we obtain a state more consistent with WOD09?
If so, how useful is it?

Ocean Model and Assimilation Setup

- Configuration like in CCSM3
- Greenland displaced pole
- Nominal 1 degree with 40 levels (25 to 400m depths)
- KPP, GM (values like in CCSM3)
- Six month weak SSS restoring
- 15.8 day SST restoring under diagnostic ice
- COREv2 daily forcing
  - 260 year normal-year spinup; last 20 Jan 1 state used as IC
  - Interannual forcing for all experiments
- Daily assimilation of WOD09 T & S obs. using DART starting Jan 1 1990 for about 2 years

A New Inflation Scheme: Climatology-based Spread Inflation

- Supplement ensemble spread $\mathbf{S}$ with an apriori estimate $\mathbf{S}^{(\mathbf{bg})}$:

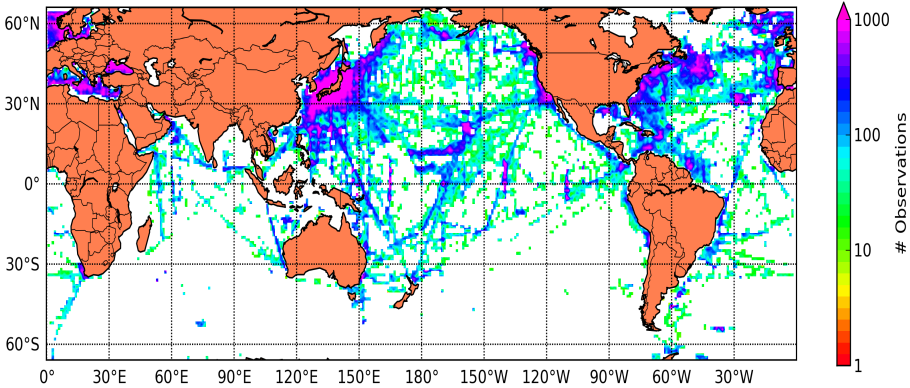
$$\mathbf{S} \mapsto (\mathbf{1} - \epsilon)\mathbf{S} + \epsilon\mathbf{S}^{(\mathbf{bg})}$$

  ϵ is a small parameter and $\mathbf{S}^{(\mathbf{bg})}$ is related to climatology
- Rewritten as

$$\mathbf{S} \mapsto \mathbf{S} + \epsilon(\mathbf{S}^{(\mathbf{bg})} - \mathbf{S}),$$

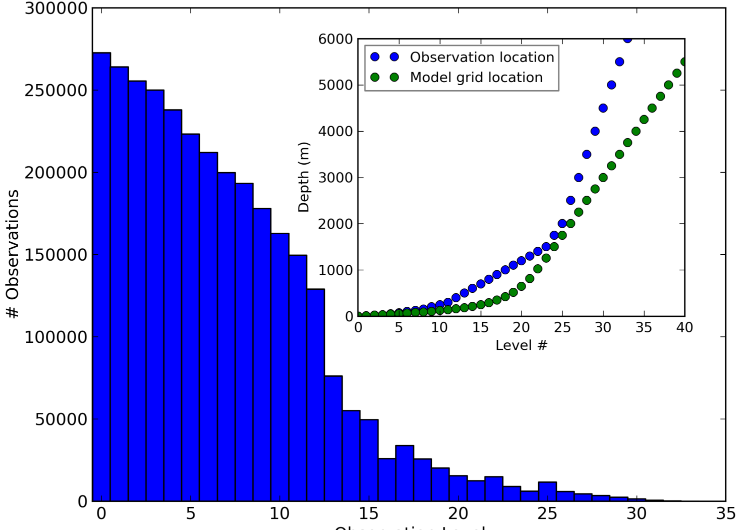
  it can be shown that this corresponds to parameterizing model error by the second term on the RHS
- Can lead to deflation (unlike most previous schemes)
  - This feature can be useful when observations are sparse
- New scheme can be additive (stochastic) or multiplicative (deterministic)
- Validated in Lorenz-96

Horizontal Density of WOD09 T & S Observations Used



Vertical Density of World Ocean Database 2009 (WOD09) T & S Observations Used

Only 12% below 400m



Outline

Brief Introduction

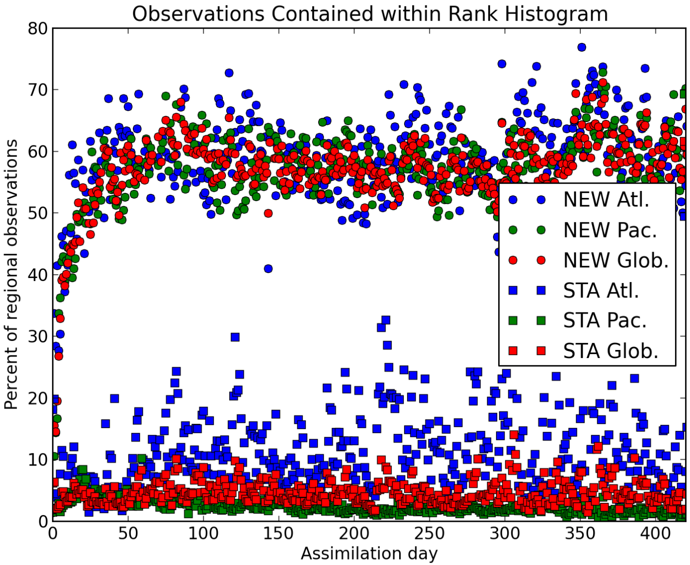
Problem Formulation

Results

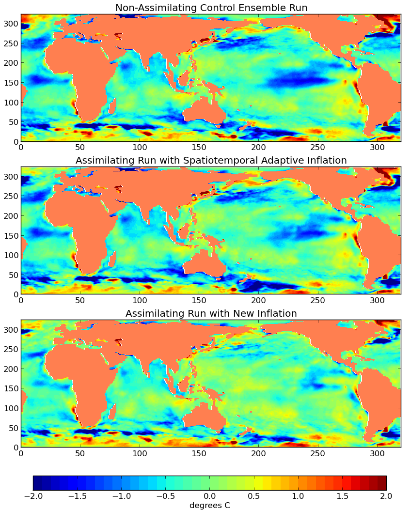
- SST Changes due to Assimilation
- Sub-Surface Changes due to Assimilation
- Utility of Assimilations for Initialization?

Conclusions

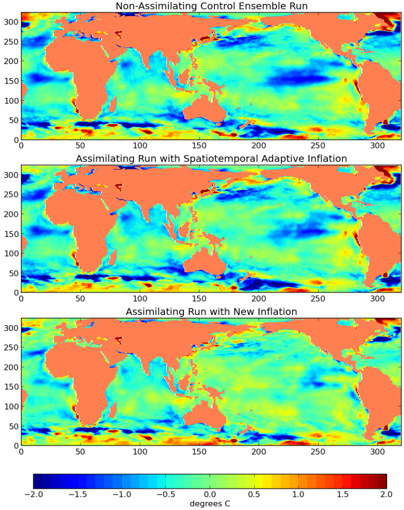
Total Observations Contained in Ensemble

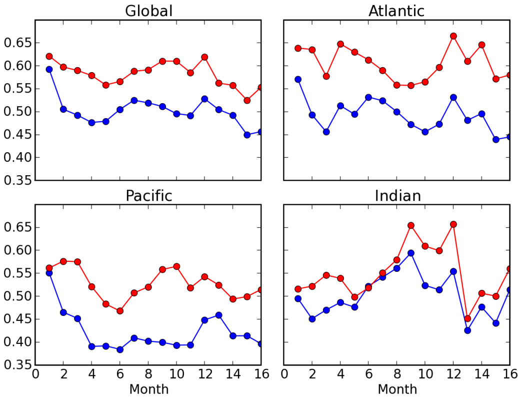


SST Errors wrt NOAA OI SST Reduced



SST Errors wrt NOAA OI SST Reduced



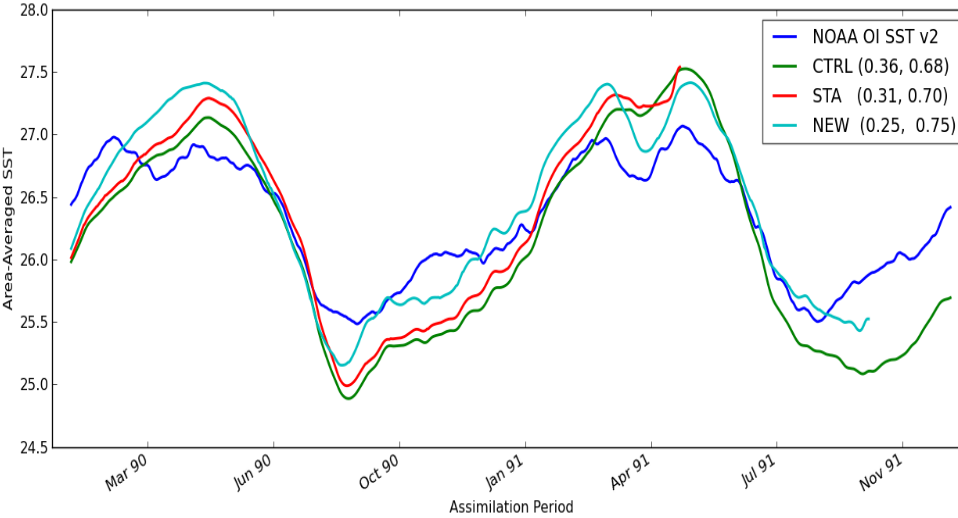


Red: Error of Control Ensemble-mean

Blue: With Assimilation

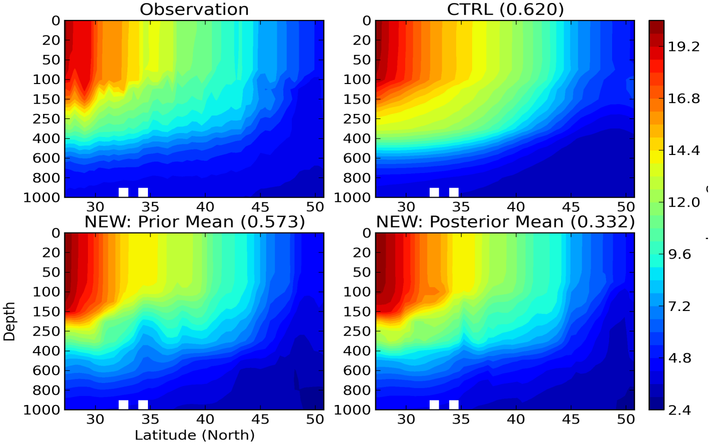
30-day running mean of (area-wt) SST in NINO3 region

Error and anomaly-correlation in legend



Meridional-Depth Distribution of Temperature

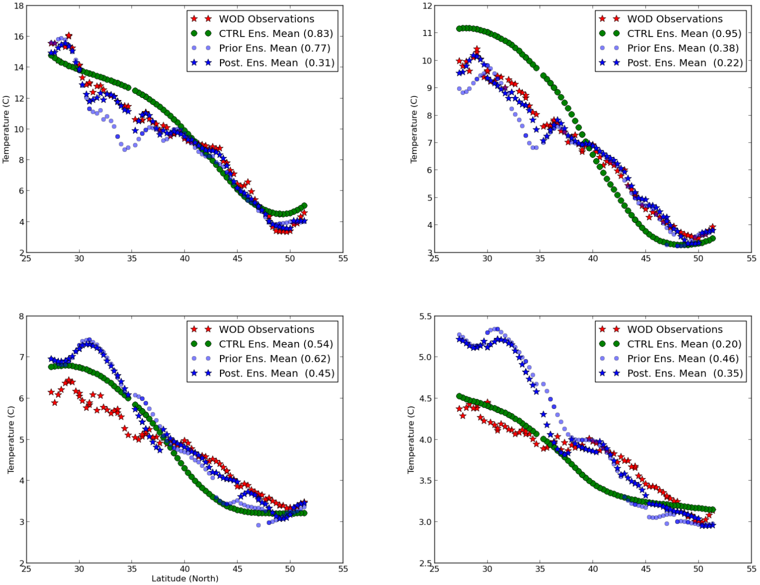
158W, north of Hawaii[1], between Feb 12^{th} and 23^{rd}, 1991



[1]Choice of section guided by an inspection of the location of observations with an aim to identifying a region with observations that had reasonable meridional-depth coverage over a short period of time

Subsurface Meridional Temperature Profiles at 158W

200m (top-left), 400m (top-right), 600m (bottom-left), and 800m (bottom-right)



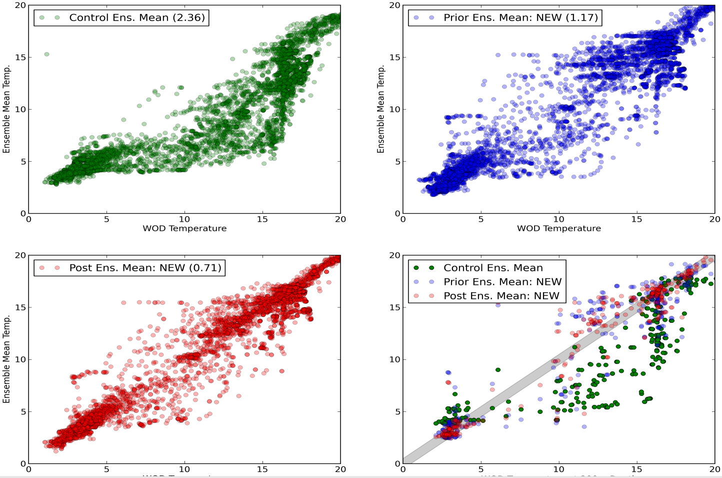
Subsurface Meridional Temperature Profiles at 158W

- Fair correspondance between observations and ensemble mean at depths of 200m and 400m with new scheme
- Poorer comaprisions at depths of 600m and 800m. Likely reason is fewer observations at depth
- Signs of ensemble collapse at lower latitudes with new scheme
- Errors with new scheme lower than with STA

Subsurface North Atlantic

- Surprisingly, over the period considered, WOD does not have a horizontal-depth section in the Atlantic (similar to 158W in the Pacific) that has reasonable observational coverage over a short period of time.
- Therefore considered temperature observations within 5 degrees of longitude and 20 degrees of latitude from 40W and 40N over the last month of the experiment.

Subsurf. N. Atlantic: Scatter of model T vs. obs. T



- Diagonal alignment indicates lower error
- Clustering of points at low and high ends indicates front.
- Cold bias in CTL. Reduced bias with DA

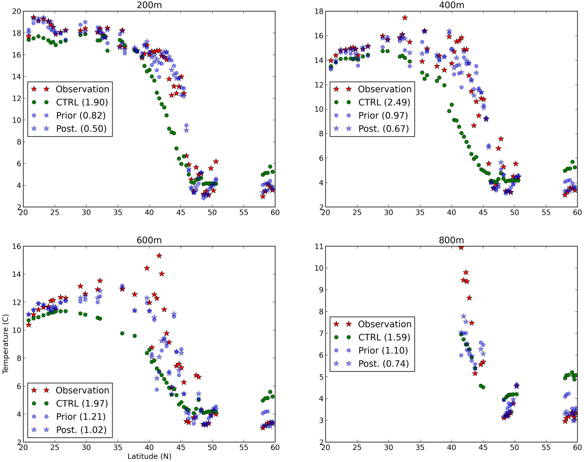
Forecast and Analysis Errors in Temperature Subsurface N. Atlantic

Depth	# Obs.	STA		NEW	
		Prior	Posterior	Prior	Posterior
200	339	1.88	1.03	1.29	0.71
400	281	1.42	0.91	1.37	1.02
600	217	1.13	0.61	1.06	0.80
800	68	0.90	0.50	0.96	0.48

Experiment NEW achieves lower prior (and posterior) error than experiment STA at depth 200m where most observations are present. Other differences are likely insignificant

Subsurface N. Atlantic Meridional Temperature Profiles

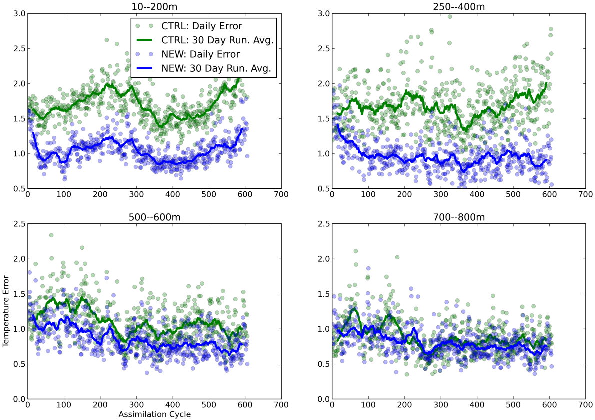
200m (top-left), 400m (top-right), 600m (bottom-left), and 800m (bottom-right)



Reasonable representation of NW Atlantic current

Global Subsurface Temperature Errors

Top Left: 10m to 200m; Top Right: 250m to 400m
Bottom Left: 500m to 600m; Bottom Right: 600m to 800m



Significant improvements above 400m
(where a greater number of observations are present)

Utility of Assimilations for Initialization

- Diagnostics demonstrate improvement in the performance of the DA system with the new inflation scheme
- Reductions in error while significant are modest ($\approx 20\%$)
- Furthermore, the DA system has limitations
  - Insufficient observations
  - Small ensemble size, ...

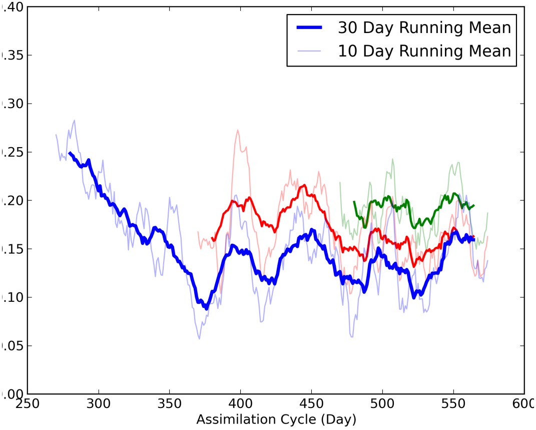
Given these limitations, and only modest reductions in error, are the assimilations of any use?

Do *initialized* trajectories stay close to observations for useful periods of time?

Or are remaining inconsistencies so large as to overwhelm the minor reductions in error?

n

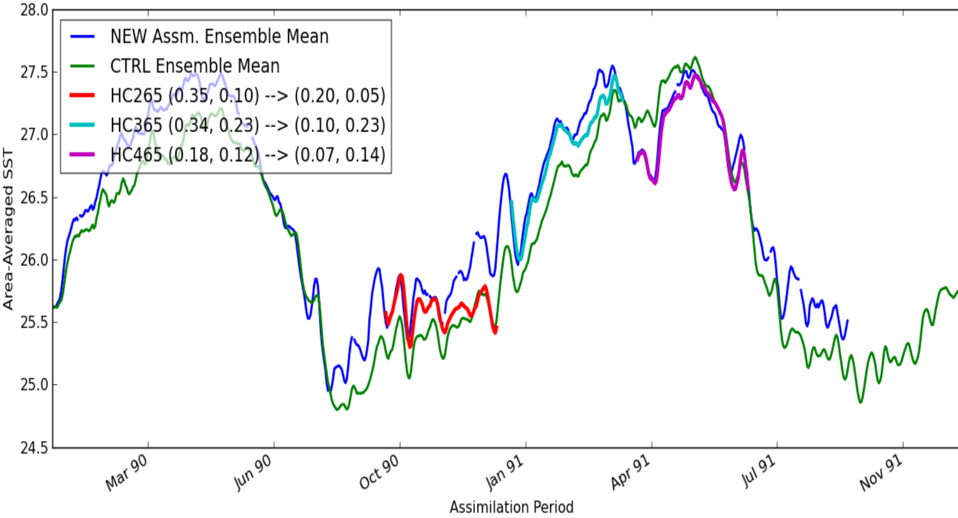
Hindcast Skill $(1 - \frac{E_{assim}}{E_{control}})$



- Skill at a level of about 0.20 is seen for a period of about 3 months following initialization
- A small level of skill persists for longer
- Commensurate with level of error reduction in assimilation runs

Hindcasts of (area-wt) SST in NINO3 region

Error and anomaly-correlation in legend



Outline

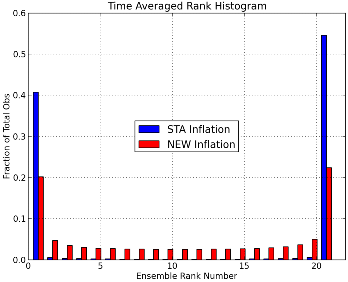
Conclusions and Questions

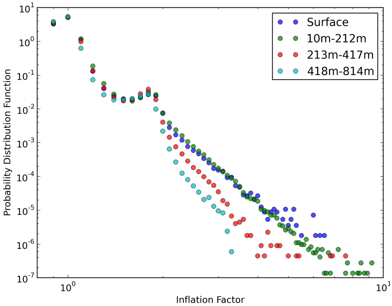
- DA over 1990-91 period reduced errors over control-ensemble wrt WOD09 by 15-20%
- Utility verified as significant hincast skill over about 3 months

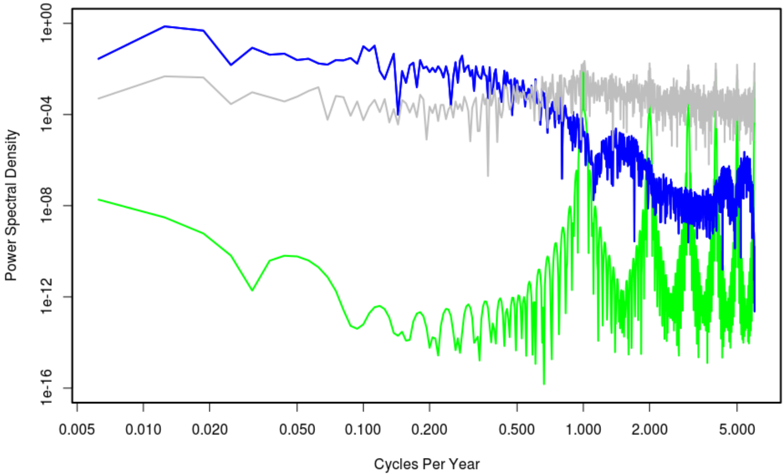
Is the 3 month timescale related to *model-error*?

- Error reduction and skill improvement with
  - ARGO
  - CESM

Rank Histogram and Inflation Factor Distribution







Power Spectral Density
1e+00
1e-04
1e-08
1e-12
1e-16
0.005
0.010
0.020
0.050
0.100
0.200
0.500
1.000
2.000
5.000
Cycles Per Year