

# Progress in the Development of Subseasonal Ensemble Forecast techniques

*CTB Proposal: “Development of Subseasonal Ensemble Forecast Techniques”*

36th Annual Climate Diagnostics and Prediction Workshop  
Ft. Worth, TX  
October 3-6, 2011

Siegfried Schubert<sup>1</sup>, Malaquías Peña<sup>2</sup>, Yoo-Geun Ham<sup>3</sup>, Yehui Chang<sup>4</sup>, Michele Rienecker<sup>1</sup>, Max Suarez<sup>1</sup>, Yuejian Zhu<sup>2</sup>, Weiyu Yang<sup>2</sup>, Mark Iredell<sup>2</sup>

<sup>1</sup> NASA, Global Modeling and Assimilation Office

<sup>2</sup> NOAA, Environmental Modeling Center

<sup>3</sup> GESTAR, Universities Space Research Association

<sup>4</sup> GESTAR, Morgan State University

# Goals of Proposed Work

Develop and Test Initialization Techniques that focus on Improving the Prediction of Subseasonal Times Scales with a Focus on the MJO, and Implement those at EMC

- Adapt and Test Breeding Approach Used at NASA for Seasonal Prediction to Subseasonal Time scales
- Adapt and Test the EMC Ensemble Transform (ET) Technique within current EMC Systems (CFS, GEFS)
- Carry out Coordinated (NASA and EMC) Experimentation to Evaluate Approaches
- Develop an ESMF-compatible coupled model at NCEP using the GEOS-5 coupled model as a prototype, and implement the “best” initialization approach

# Experimental Framework (NASA)

- Model

- early version of GEOS-5 AGCM coupled to MOM4 Ocean

- Initial Conditions

- atmosphere and land: MERRA
- ocean: replay of MERRA with Coupled GEOS-5 model

- Examine Two Perturbation strategies

- **Breeding** and **Empirical Singular Vectors** (Yoo-Geun Ham)

- Control consists of either 1-day LAF or random

perturbations

with large-scale spatial structure

# GEOS-5 Coupled Model Replay System

## AGCM

- Finite-volume dynamical core (S.J. Lin)
- Moist physics (J. Bacmeister, S. Moorthi and M. Suarez)
- Physics integrated under the Earth System Modeling Framework (**ESMF**)
- Generalized vertical coord to 0.01 hPa
- Catchment land surface model (R. Koster)
- Prescribed aerosols (P. Colarco)
- Interactive ozone
- Prescribed SST, sea-ice

## CGCM

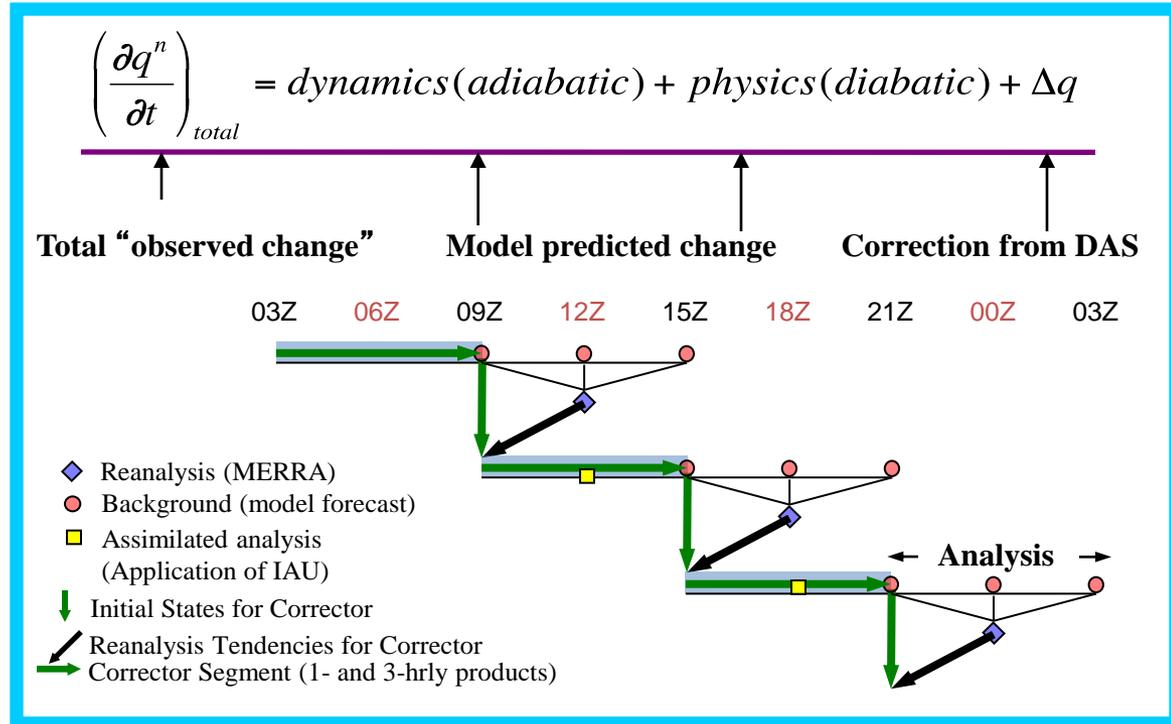
- GFDL ocean model (MOM4)
- A replay of the atmospheric data analysis in the CGCM.
- Has the potential to substantially reduce initialization shocks

## Replay

- Apply Incremental Analysis Increments (IAU) to reduce shock of data insertion (Bloom et al.)
- IAU gradually forces the model integration throughout the 6 hour analysis period

## Reanalysis

- MERRA
- NCEP, JRA-25, ERA

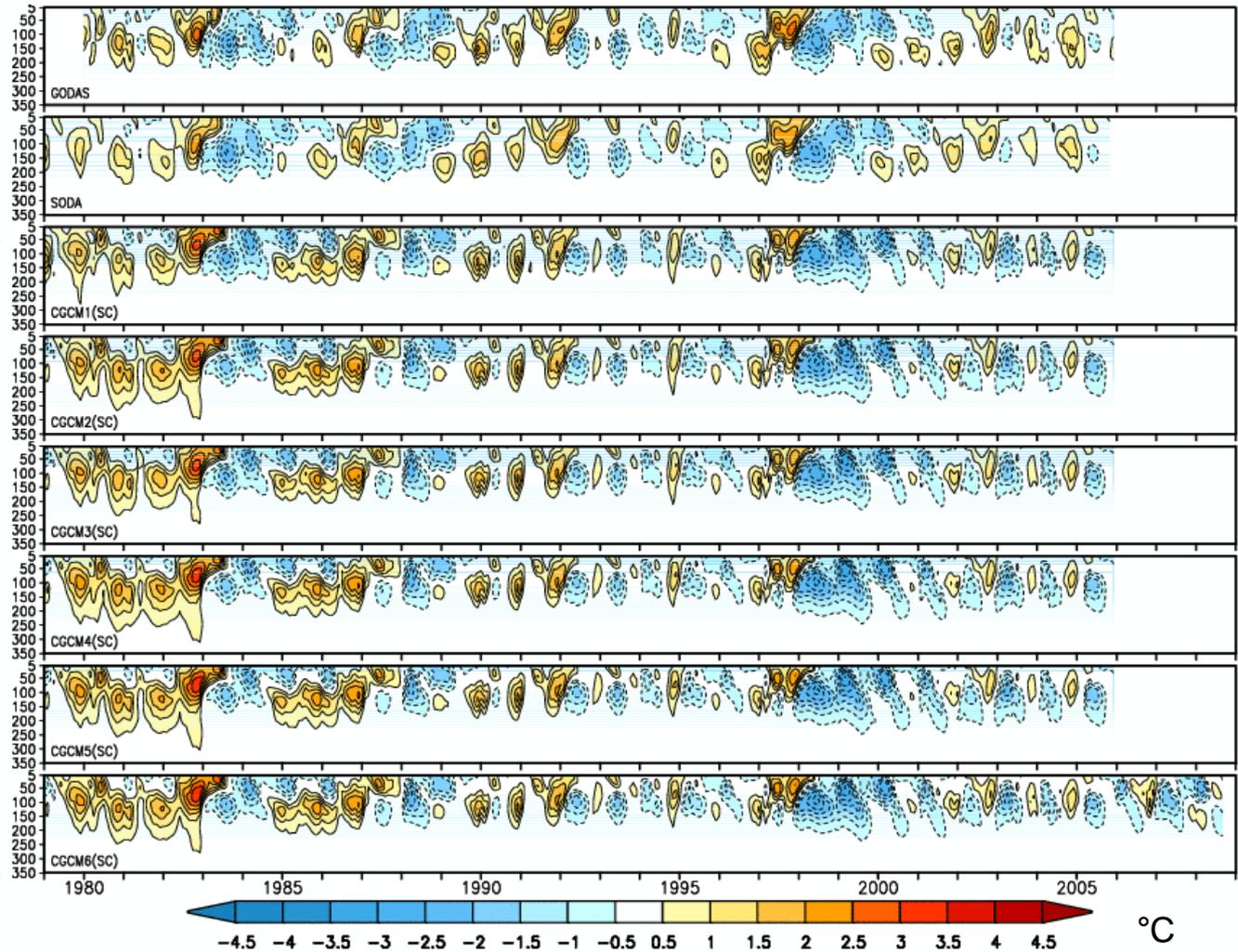


# Subsurface Ocean Temperature Anomalies (5S-5N, 130E-80W)

GODAS

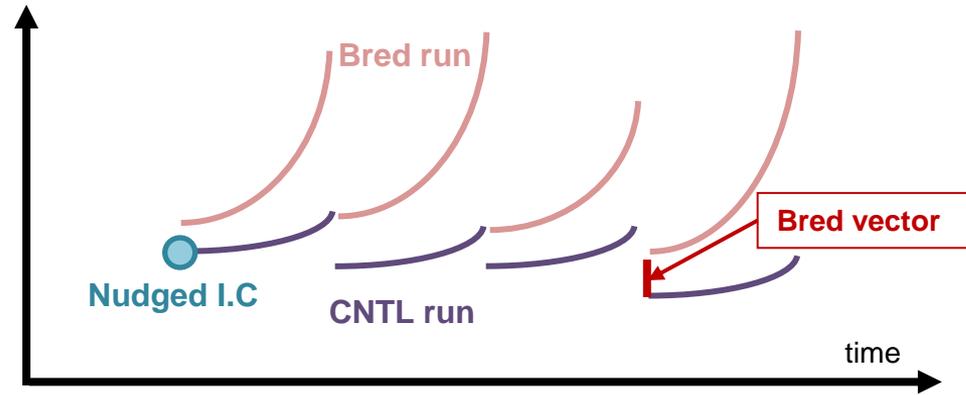
SODA

Results of repeated  
Replay with GEOS-5



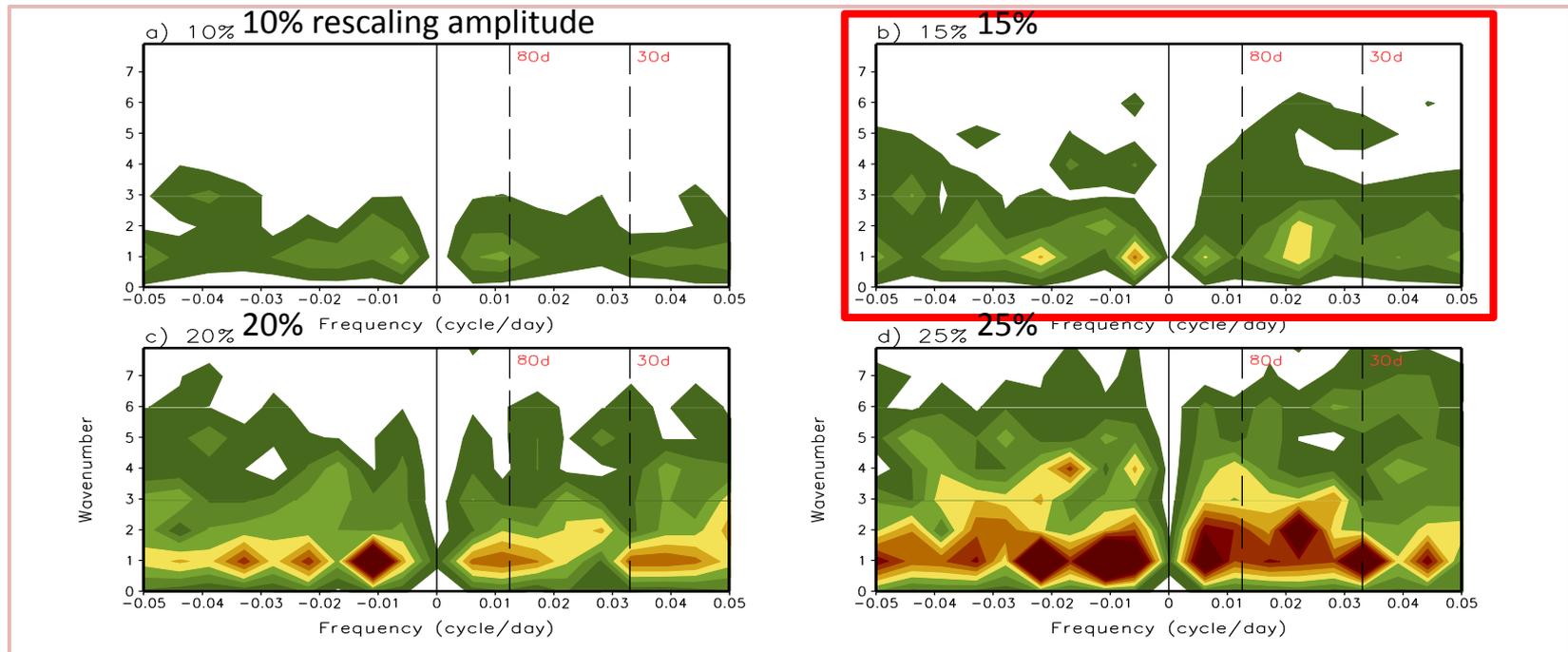
# Breeding Experiments

Following:  
Chikamoto et. al. 2007  
Yang et al. 2009



\* Rescaling magnitude :  
Based on natural variability of VP200 over 40-180E, 20S-20N

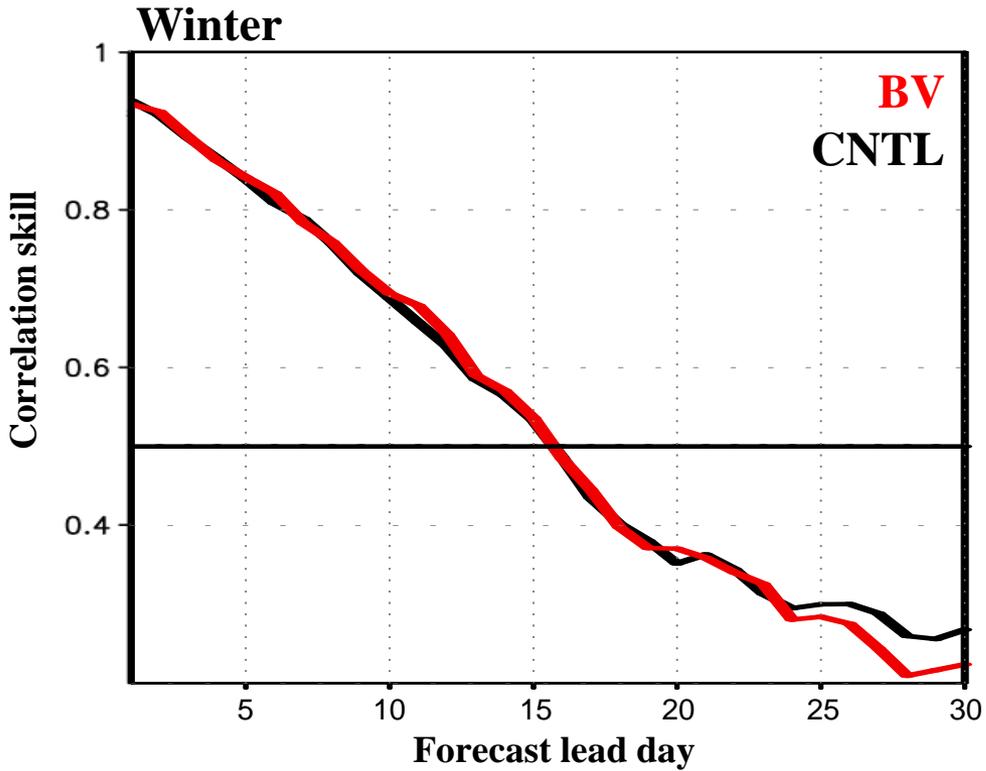
## Space-Time Power Spectrum (2 day rescaling)



# Forecast skill of bi-variate RMM index - BV

$$\text{Cor of RMM index } (\tau) = \frac{[\sum_{i=1}^N a_{1i}(t) \cdot b_{1i}(t) + a_{2i}(t) \cdot b_{2i}(t)]}{\sqrt{\sum_{i=1}^N [a_{1i}^2(t) + a_{2i}^2(t)]} \cdot \sqrt{\sum_{i=1}^N [b_{1i}^2(t) + b_{2i}^2(t)]}}$$

$a_{1i}(t), a_{2i}(t)$  : observed RMM1,2 at day t  
 $b_{1i}(t), b_{2i}(t)$  : simulated RMM1,2 at day t  
 $\tau$  : Forecast lead day  
 N : Number of forecasts



	Red	Black
Ensemble perturbation	BV (2dy 15%)	LAF (1-day)
Observation	MERRA	
Initialized method	MERRA replay with coupled GCM	
Ensemble member	2	
Model	GEOS5 CGCM	
Prediction period	1992.11.1-1996.04.30 (Total : 180 cases)	

# Empirical Singular Vector (ESV)

## 1. Define initial (X) & final (Y) variables with **forecast data**

- (1) Select optimal time : 10 days → Time-lag between X & Y : 10 days
- (2) Initial variable (X) : U850,U200,VP200 at initial time  
Final variable (Y) : U850,U200,VP200 at 10-day after
- (3) Calculate anomaly
  - Subtract daily climatology & previous 120 day mean
  - Divide by standard deviation of each variables
- (4) X : PC time series of EOF 5 modes of initial variable  
Y : PC time series of combined EOF 2modes of final variable



## 2. Formulate the Empirical Operator (L)

Linear inverse modeling ( $L_{\text{linear}}$ )

$$\begin{array}{ccc} \mathbf{X} & \xrightarrow{\mathbf{L}} & \mathbf{Y} \\ \text{(Initial)} & & \text{(Final)} \end{array}$$

$$\begin{aligned} \mathbf{Y} &= \mathbf{L} \cdot \mathbf{X} \\ \mathbf{L} &= \mathbf{YX}^T (\mathbf{XX}^T)^{-1} \end{aligned}$$



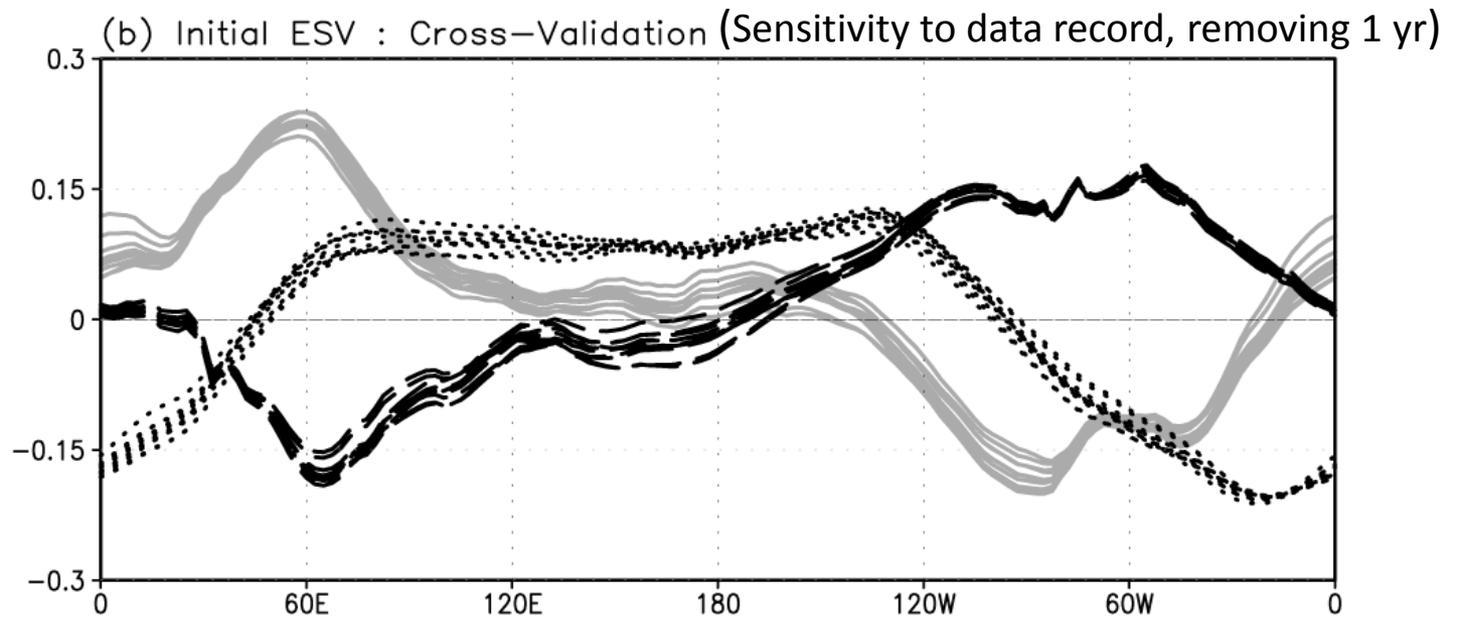
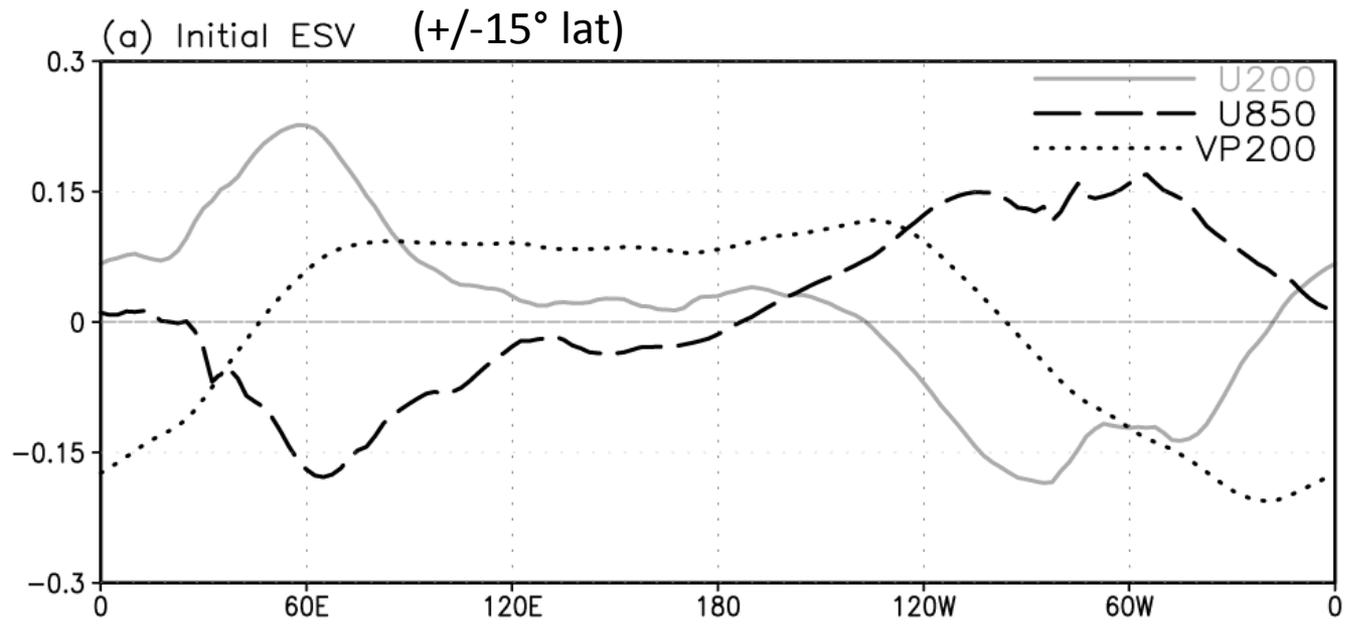
## 3. Find fast growing perturbation using SVD

$$L_{\text{linear}} = \mathbf{USV}^T$$

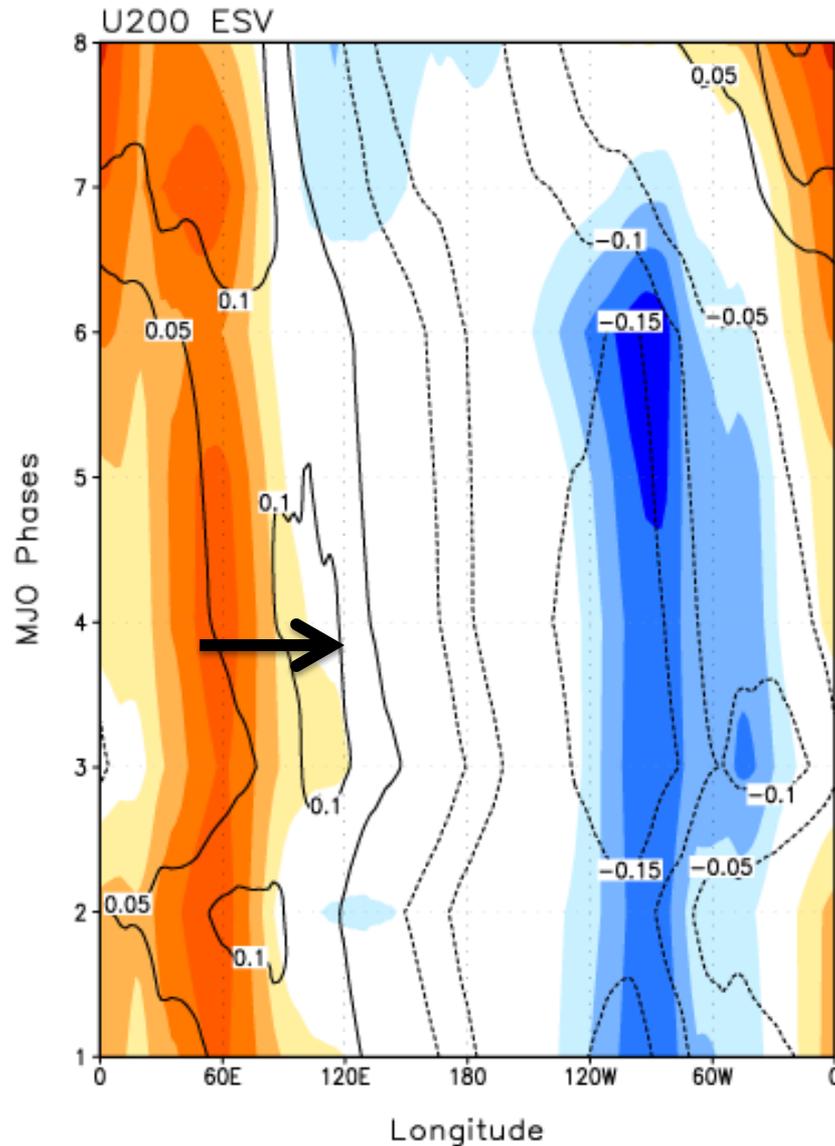
Fast growing perturbations :

→ Right singular vectors whose singular value is maxima

MJO  
Phase 4



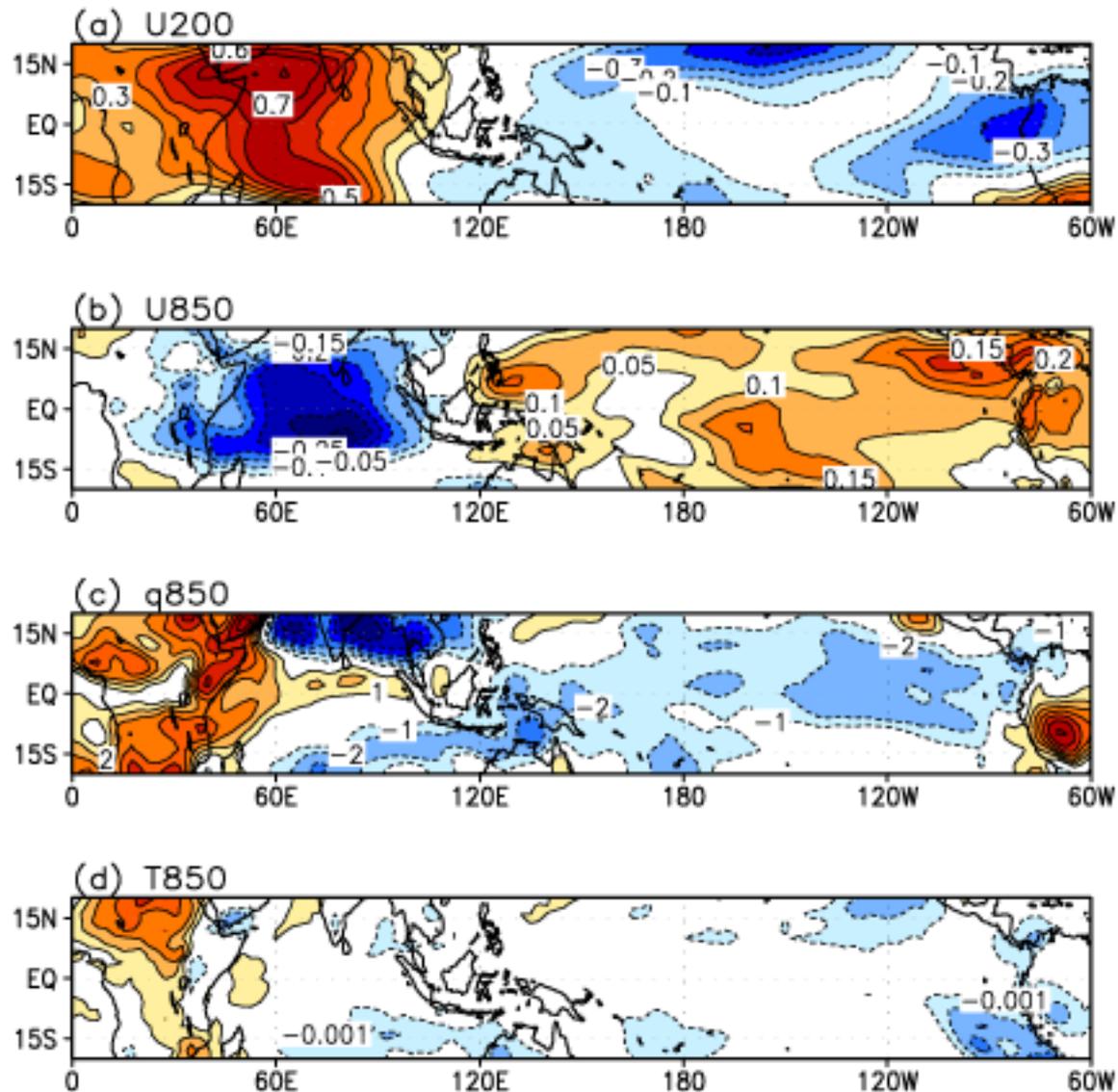
# ESV of U200 over equator



Shading : Initial  
Contour : Final (10 days after)

ESV mode : Eastward propagating signal over Indian Ocean

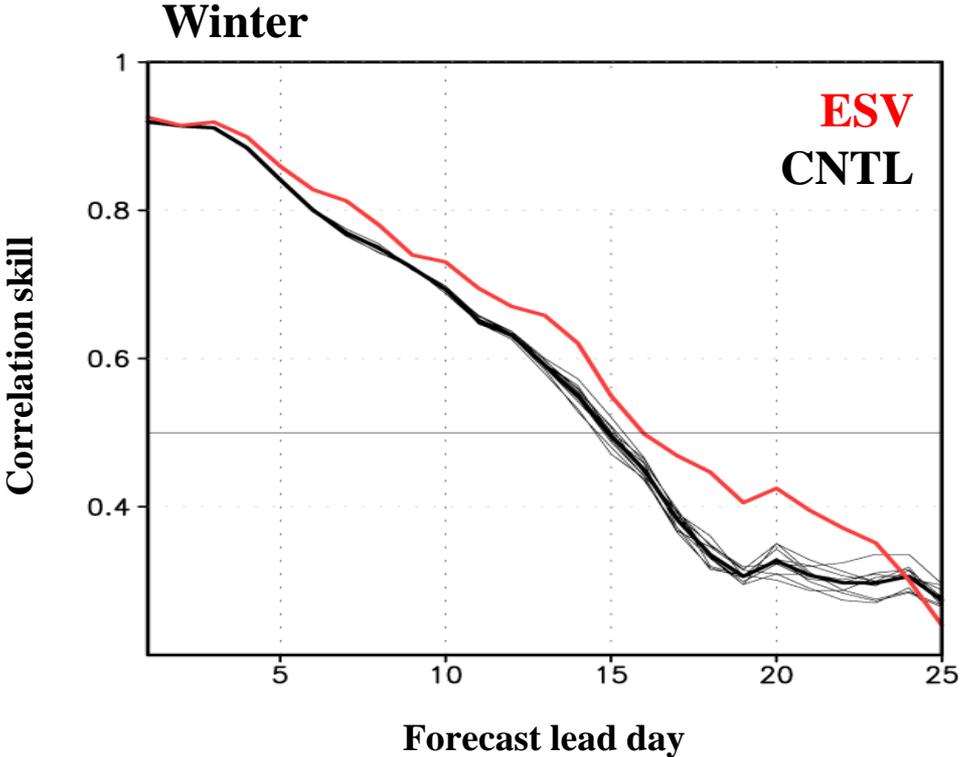
# Empirical Singular Vector at phase 4



# Forecast skill of bi-variate RMM index

$$\text{Cor of RMM index } (\tau) = \frac{[\sum_{i=1}^N a_{1i}(t) \cdot b_{1i}(t) + a_{2i}(t) \cdot b_{2i}(t)]}{\sqrt{\sum_{i=1}^N [a_{1i}^2(t) + a_{2i}^2(t)]} \cdot \sqrt{\sum_{i=1}^N [b_{1i}^2(t) + b_{2i}^2(t)']}}$$

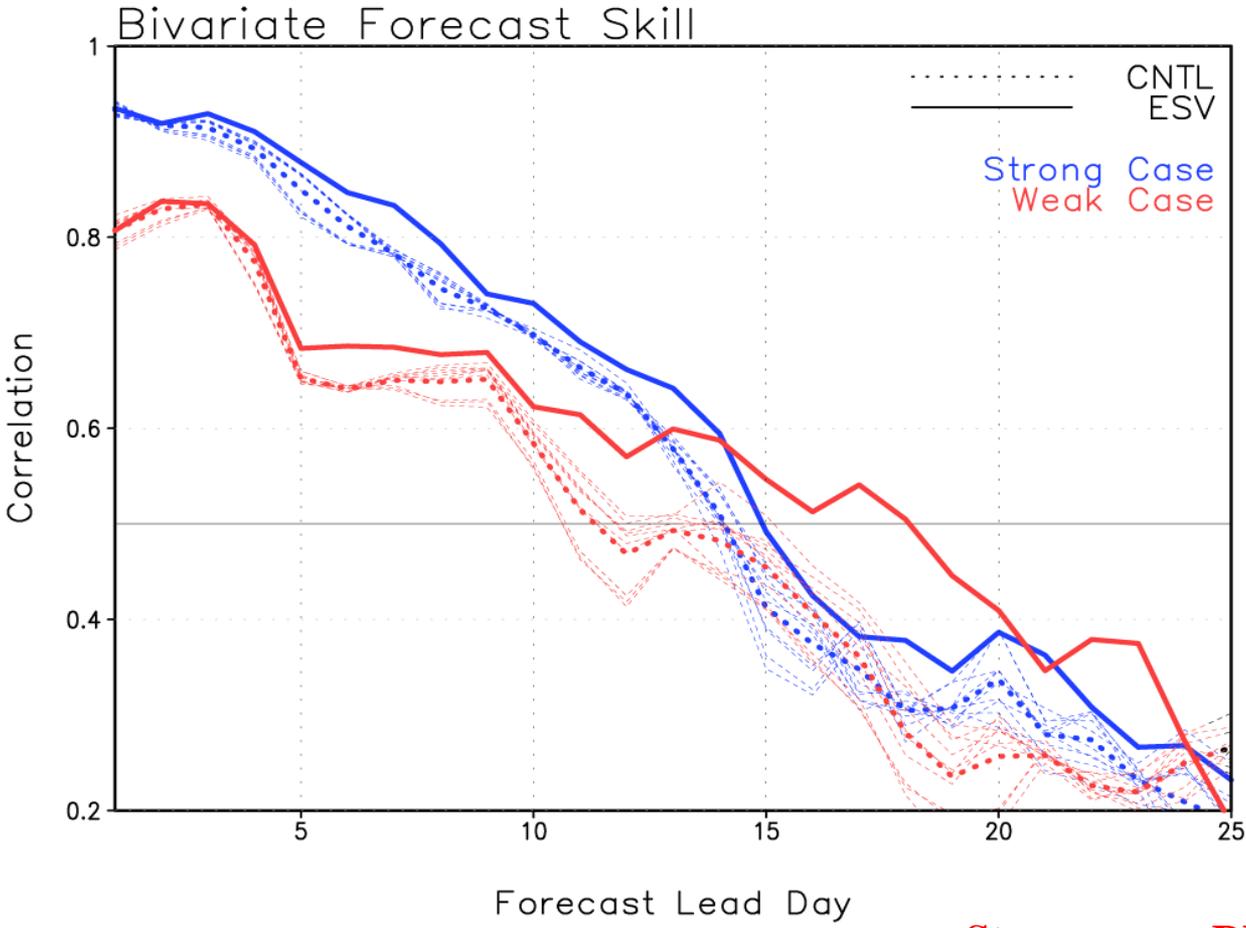
$a_{1i}(t), a_{2i}(t)$  : observed RMM1,2 at day t  
 $b_{1i}(t), b_{2i}(t)$  : simulated RMM1,2 at day t  
 $\tau$  : Forecast lead day  
 N : Number of forecasts



	<b>Red</b>	<b>Black</b>
Ensemble perturbation	<b>ESV</b>	<b>Random Pert.</b>
Observation	MERRA	
Initialized method	MERRA replay with coupled GCM	
Ensemble member	<b>2</b>	
Model	GEOS5 CGCM	
Prediction period	1990.1.1-2000.04.30 (Total : 180 cases)	

**ESV shows systematic improvement in MJO prediction**

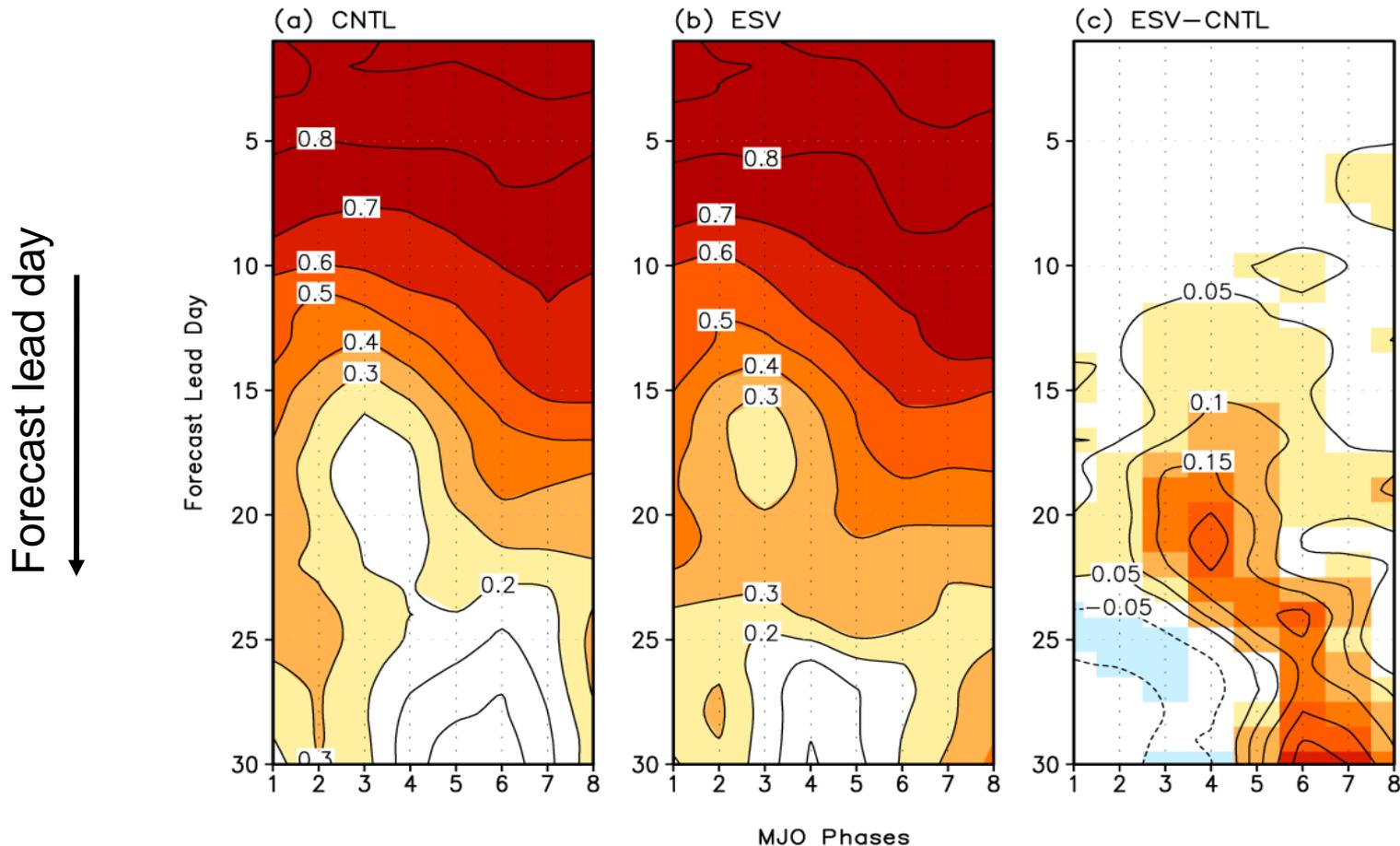
# Forecast skill of bi-variate RMM index



**Strong case : RMM magnitude > 1**  
**Weak case : 1 > RMM magnitude**

**ESV shows robust improvement for weak MJO cases**

# ESV Correlation skill : phase dependency

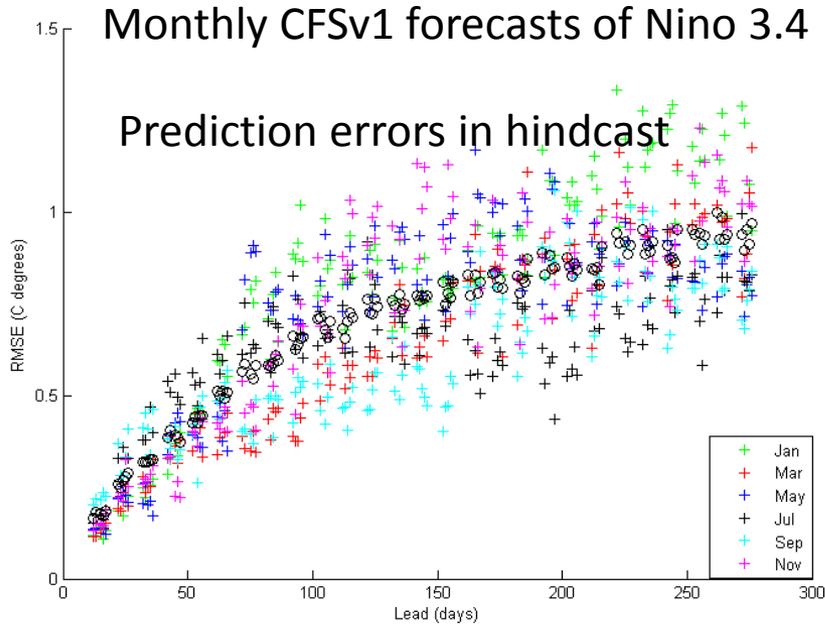


Correlation skill improvement of ESV prediction exists over unpredictable phase (phase 4-8)

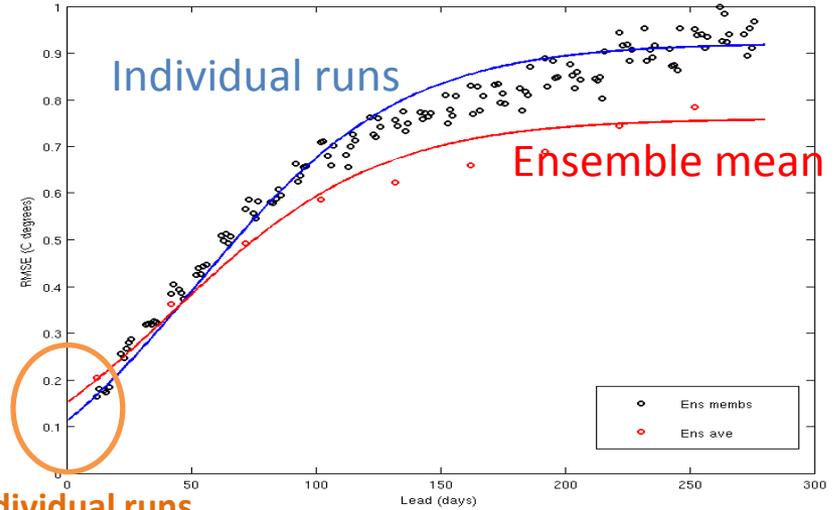
# Activities at EMC under this project

- Performance evaluation of the lagged versus the “all at once” ensemble generation approach
- Developing of a new strategy for analysis error estimation to improve ensemble perturbations
- Several attempts to compile and run an ESMF-based coupled model
- The lack of suitable coupled model to perform ET experiments led to i) the use of old CFSv1 as a testing model, and ii) carrying out and assess the performance of GEF5 on extended ranges (out to 35 days)

# Ensemble generation approaches: Lagged versus “all at once”



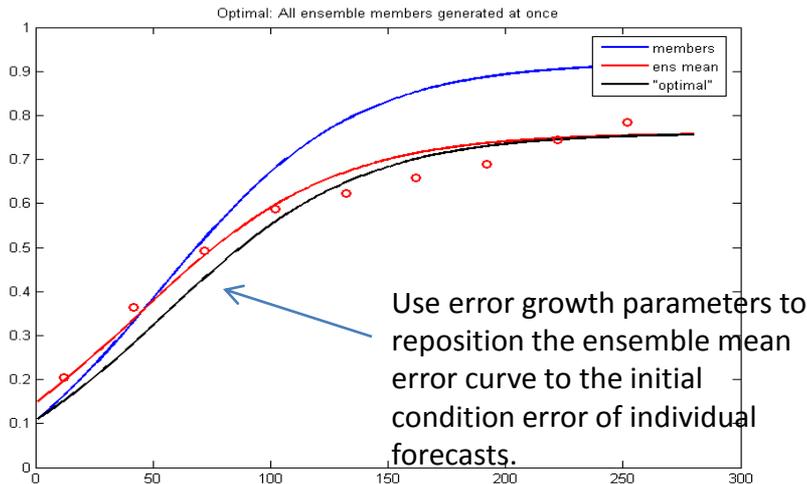
Error growth parameterization to estimate mean initial errors



Individual runs  
have smaller  
initial errors



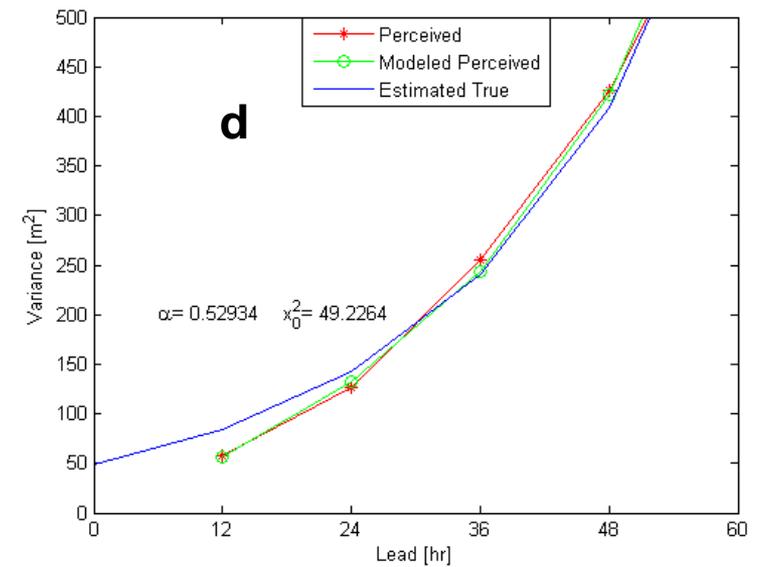
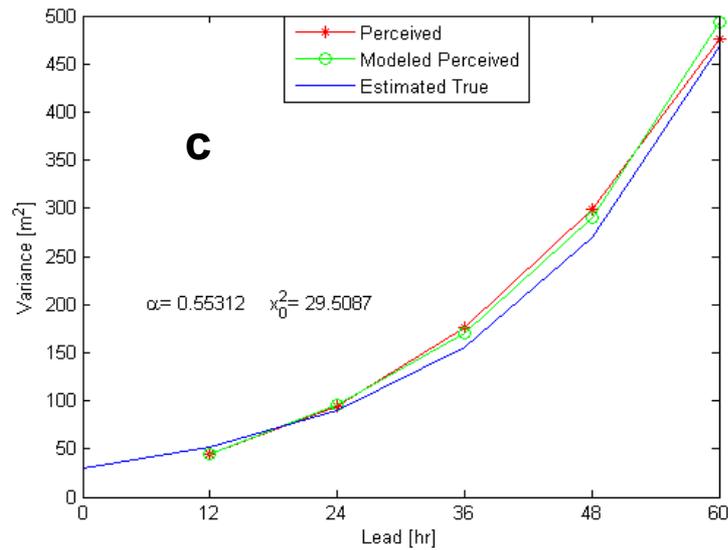
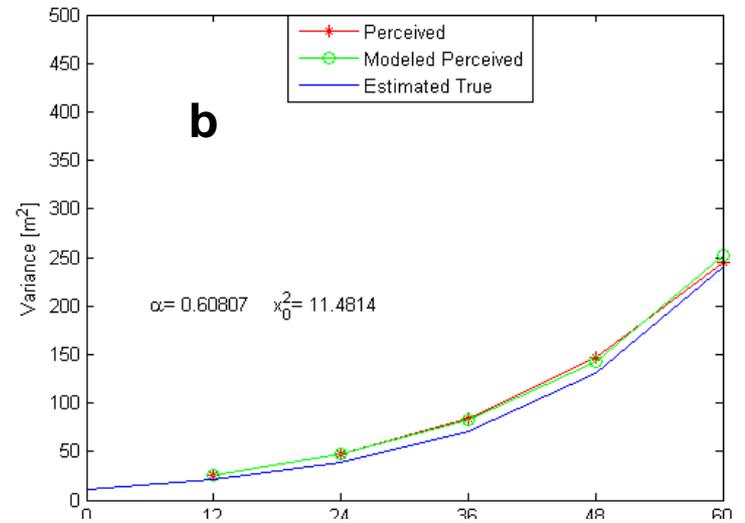
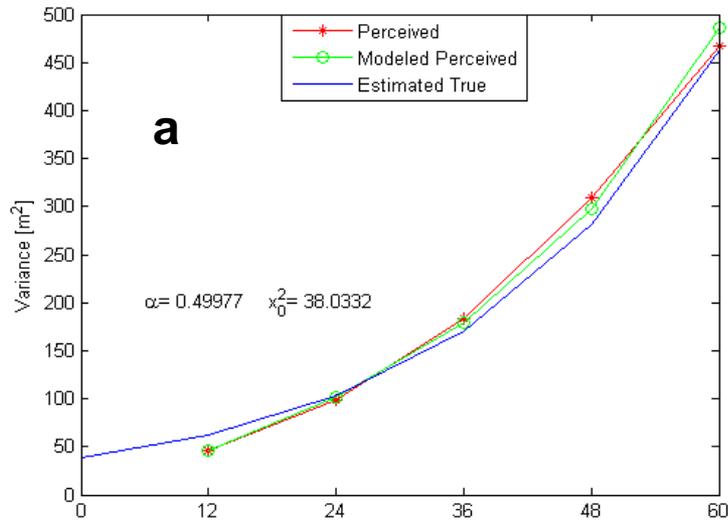
This is expected because old runs deteriorate initial conditions



- ~11 days predictability gain in the CFSv1 15-members per month ensemble. Likely much less in the CFSv2.
- ~5 days using the latest 8 runs in the ensemble.
- No sophisticated ensemble generation scheme used
- For sub-seasonal predictions the gain would represent a significant improvement

# Improving ensemble perturbation approaches

- The amplitude of ensemble perturbations must reflect the analysis uncertainty of the ocean and the atmosphere variables in the coupled model.
- The problem is that analysis uncertainty is usually not known. In variational data assimilation systems it is computationally very expensive to estimate it. In ensemble-based data assimilation methods should produce this uncertainty but with limitations
- In the mean time we developed a technique to estimate analysis uncertainty based on knowledge of the error growth and the differences between analysis and forecasts.
- The ET method currently uses historical errors fixed in time, changing month by month, as a surrogate of analysis uncertainty. The new method is aimed at producing initial perturbations consistent with analysis uncertainty. We tested the method in a perfect-model environment and in historical archives of forecasts from global operational model (GFS, ECMWF, FNMOC, CMC)



Analysis and Forecast error variances as a function of lead time for four global models: (a) GFS, (b) CMC, (c) ECMWF, and (d) FNMOC for models operational in 2008.

# Plans for Final Year

- Continue evaluating ESV and ET approaches with coordinated experimentation (NASA/GMAO and EMC models)
- Implement NCEP physics into GEOS-5 model and re-assess ESV and other approaches
- Complete the development of ESMF-compatible coupled model at NCEP and test new initialization approaches

# VP200 Life cycle composite

