Development of the initialization system for seasonal-to-decadal climate prediction

Yoo-Geun Ham

Global Modeling and Assimilation Office, NASA/GSFC

Collaborators: Michele M. Rienecker, Data Assimilation team in GMAO, Ricardo Todling, Hyo-Jong Song, Jong-Seong Kug
Data assimilation system in NASA/GMAO

**Coupled A-O-S-L initialization of decadal predictions**

**Atmosphere:** 6-hour interval
- “replay” of GMAO atmospheric MERRA analysis

**Ocean:** 1-day interval
- Optimal Interpolation (OI)
  - Flow adaptive error-covariance localization following density
  - Adaptive estimation of representation error for SSH
- Ensemble Kalman Filter (EnKF)

**Sea-ice:** 1-day interval
- Ensemble Optimal Interpolation (EnOI)

---

**Ocean observations:**
- Topex/Jason SSH anomalies
- Argo, PIRATA, RAMA in situ T and S profiles
- In situ T from TAO, XBT, Pirata and Rama
- Reynolds SST
- Levitus surface salinity while waiting for Aquarius
Ocean reanalysis results

Time series of 7-month running means of HC300a (°C)

Climatological Fields over Atlantic

(a) Temperature (0–500m)

(b) AMOC
Need for Decadal Prediction

- Key Variable
  - Atmosphere
  - Mixed-layer, Thermocline
  - Deep Ocean

- Crucial Component
  - Initial value
  - Initial Value + Boundary Value
  - Boundary value (ex. CO₂)

- Time scale
  - 1 month
  - 1 yr
  - 10 yr
  - 100~yr

- Bridge the gap
  - Huge Gap!!
  - Crucial Component

- Trend

- Decadal Prediction
  - Bridge the gap
Decadal Climate Prediction in NASA/GMAO
- follows the CMIP5 experimental protocol for decadal predictions

- Forecast period: 10 years
- Number of hindcasts: 61 (1959 to 2009)
- Number of ensemble member: 3 (Control run + BV)
- Initial Condition: Reanalysis from NASA/GMAO
Generation of Ensemble Perturbation

- Method: **Two-sided breeding**
- Norm variable: **Averaged Temperature 0-500m (HC500)**
- Norm Regions: **Atlantic ocean (100W-20E, 20N-70N)**
- Initial BV magnitude: **Reduced to 10% of natural variability**
- Rescaling Interval: **1-, and 5- year**
- Initial perturbation: **6-hour atmospheric perturbation**
RMS of BV Temperature over Atlantic

Time-averaged

RMSE of BV TEMP

Depth (m)

RMSE of BV

Time

*Monthly mean
1st EOF of BV HC500

(a) 1YR BV : HC EOF 1st 14.36% (b) 1YR BV : HC EOF 2nd 13.09%

(c) 5YR BV : HC EOF 1st 34.32% (d) 5YR BV : HC EOF 2nd 20.07%
Anomaly Correlation Skill of SST, with trend

* 3-year mean
MSSS of SST, with trend

\[
\text{MSSS} = \frac{\text{MSE}_{\text{C20C}} - \text{MSE}_{\text{Post}}}{\text{MSE}_{\text{C20C}}}
\]
MSSS of HC500, with trend

\[
\text{MSSS} = \frac{\text{MSE}_{\text{C20C}} - \text{MSE}_{\text{Post}}}{\text{MSE}_{\text{C20C}}}
\]
Development of Prediction System

1. Non-stationary Incremental Analysis Updates (NIAU)
   → Prevent time inconsistency in IAU


2. Generation of Optimal Ensemble Perturbation
   → Generating fast-growing perturbation
Introduction of IAU

Arbitrary fluctuation due to the assimilation
IAU vs Non-stational IAU (NIAU)

**IAU:** \( x_n^b = M_{n-1}(x_{n-1}^b) + f, \ n = 1, \ldots, N \)

- Time-constant Forcing (f)

**NIAU:** \( x_n^b = M_{n-1}(x_{n-1}^b) + M_{n-1} \cdots M_1 M_0 f, \)

- Dynamically-constrained time-evolving Forcing

\[
\begin{align*}
M_{n-1} & : \text{Nonlinear Operator} \\
x_{n-1}^b & : \text{Background Vector at time } n-1 \\
f & = \frac{\Delta X_0}{N} : \text{IAU forcing} \\
M & : \text{TLM}
\end{align*}
\]
The solution at final time is equivalent between NIAU and intermittent method.
Time Tendency difference

RMS of Time Tendency Difference

Analysis Time in days

EKF—True
IAU—True
NIAU—True
Lorenz model results: RMS Error

![Graph showing RMS Error over Analysis Time in days for different models: EKF, IAU, NIAU.](image-url)
Sensitivity to the assimilation interval

Time–Averaged (5–10dy) RMSE

- EKF
- IAU
- NIAU

Assimilation Interval (hour)
Ensemble Generation Methods

1. Empirical Singular Vector (ESV)

2. EnKF + ESV concept

3. Bred Vector (BV)
   • Ham, Y.-G., I.-S. Kang, and J.-S. Kug, 2012: Detection of Two Independent Coupled Bred Vectors in the Tropical Pacific and Their Application to ENSO prediction, Prog. Oceanogr.
1. Formulate the Empirical Operator \((L)\):

\[
\begin{align*}
X & \xrightarrow{L} Y \\
\text{-Initial-} & \quad \text{-Final-} \\
\text{Thermocline depth} & \quad \text{SST after} \\
\text{at initial time} & \quad \text{6-month} \\
\text{(snap shot)} & \quad \text{(monthly mean)}
\end{align*}
\]

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

\[
Y = L \cdot X \\
L = YX^T (XX^T)^{-1}
\]

\(X, Y\) are known from hindcast data.

2. Find fast growing perturbation using SVD:

\[
L_{\text{linear}} = \text{USV}^T
\]

- Fast growing perturbations:
  - Right singular vectors whose singular value is maxima
Empirical Singular Vector

(a) Initial Thermocline Depth

(b) Final SST
Empirical Singular Vector

Anomaly Correlation Skill 1980-2010, 2 ensemble member (Simple Method)

ESV also has advantages on probabilistic forecasts.
**MJO forecast**: Forecast skill of bi-variate RMM index

Correlation skill of RMM index ($\tau$)=

$$
\frac{\sum_{i=1}^{N} a_{ii}(t) \cdot b_{ii}(t) + a_{2i}(t) \cdot b_{2i}(t)}{\sqrt{\sum_{i=1}^{N} [a_{ii}^2(t) + a_{2i}^2(t)] \cdot \sum_{i=1}^{N} [b_{ii}^2(t) + b_{2i}^2(t)]}}
$$

$a_{ii}(t), a_{2i}(t)$: observed RMM1,2 at day $t$

$b_{ii}(t), b_{2i}(t)$: simulated RMM1,2 at day $t$

$\tau$: Forecast lead day

N: Number of forecasts

---

**Winter**

<table>
<thead>
<tr>
<th>Ensemble perturbation</th>
<th>Red</th>
<th>Black</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESV</td>
<td></td>
<td>LAF (1-day)</td>
</tr>
</tbody>
</table>

**Observation**

<table>
<thead>
<tr>
<th>Initialized method</th>
<th>MERRA replay with coupled GCM</th>
</tr>
</thead>
</table>

**Ensemble member**

<table>
<thead>
<tr>
<th>Model</th>
<th>GEOS5 CGCM</th>
</tr>
</thead>
</table>

**Prediction period**

1990.1.1-1995.12.30 (Total: 180 cases)

---

**ESV shows systematic improvement in MJO prediction**
Combine the ESV concept to the EnKF

- EnKF
  - Short assimilation interval
  - Flow-dependent

- ESV
  - Fast growing mode under climatological flow

Flow-dependent ESV (FESV)

Generation of FESV

Empirical Operator within analysis cycle

\[ L_1 \quad L_2 \quad L_3 \quad \cdots \quad L_9 \]

Empirical Operator for the FESV

\[ L_{\text{flow}} = L_9 \quad \cdots \quad L_3 \quad L_2 \quad L_1 \]

Flow-dependent ESV (FESV) vs Climatic ESV (CESV)
Statistical Forecast of SST

RMSE of SST

(a) STATIC OPERATOR \((L_{\text{clim}})\)

(b) FLOW OPERATOR \((L_{\text{flow}})\)

* CZ-SPEEDY model result
Dominant Patterns: 1st EOF

(a) CESV - Initial H (3-month interval)

(b) FESV - Initial H (3-month interval)

(c) EnKF - Initial H (10 day interval)

(d) BV - Initial H (3-month interval)

(e) CESV - Final SST

(f) FESV - Final SST

(g) EnKF - Final SST

(h) BV - Final SST
Forecast skill of NINO3.4 index
Seamless Climate Forecast System

Unified Initialization system
- EnKF + Breeding Concept
- Flow Dependent ESV (EnKF + ESV Concept)
- Non-stationary IAU in Ensemble Method

Unification

Conventional Framework

Ensemble Data Assimilation System
for OAGCM Initialization
- Development of Hybrid Method
  - EnKF + 4DVAR concept
- Incremental Updates for balanced I.C
  - Non-stationary IAU

Ensemble Generation Method
for Optimal Perturbations
- Development of new method
  - Empirical Singular Vector (ESV)
- Application of Conventional Method
  - BV, SV
Thank you