Ensemble Forecasting

Yuejian Zhu
Ensemble team leader
Environmental Modeling Center
NCEP/NWS/NOAA

Acknowledgements: EMC Ensemble team staffs
January 28 2013
Ensemble forecast is widely used in daily weather forecast.
December 2012 was 20 anniversary of both NCEP and ECMWF global ensemble operational implementation
Highlights

• Description of current operational global ensemble forecast systems
  – NCEP global ensemble forecast system
  – ECMWF global ensemble forecast system
  – CMC global ensemble forecast system

• Ensemble performance
  – Compare to deterministic forecast
  – Multi-model ensemble

• Interaction between data assimilation and ensemble forecast
  – Hybrid 3DV/EnKF implementation

• Future direction
  – EMC’s plan
  – Preliminary study/discussion
  – Challenges
Important terminology

- BV – breeding vector
- ETR – ensemble transform with rescaling
- TSR – tropical storm relocation
- STTP – stochastic total tendency perturbation

- SV - singular vector
- MME – multi-model ensemble
- SPPT - stochastically-perturbed physics tendencies
- GEM – global environmental multiscale model
- SEF – Canadian spectrum finite element model
Each ensemble member evolution is given by integrating the following equation

\[ e_j(T) = e_0(0) + \frac{d}{dt} e_j(0) + \int_{t=0}^{T} \left[ P_j(e_j, t) + dP_j(e_j, t) + A_j(e_j, t) \right] dt \]

where \( e_j(0) \) is the initial condition, \( P_j(e_j, t) \) represents the model tendency component due to parameterized physical processes (model uncertainty), \( dP_j(e_j, t) \) represents random model errors (e.g. due to parameterized physical processes or sub-grid scale processes – stochastic perturbation) and \( A_j(e_j, t) \) is the remaining tendency component (different physical parameterization or multi-model).


Reference:
"A Comparison of the ECMWF, MSC, and NCEP Global Ensemble Prediction Systems“  
# Evolution of NCEP GEFS configuration

<table>
<thead>
<tr>
<th>Year</th>
<th>Initial uncertainty</th>
<th>TS relocation</th>
<th>Model uncertainty</th>
<th>Resolution</th>
<th>Forecast length</th>
<th>Ensemble members</th>
<th>Daily frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992.12</td>
<td>BV</td>
<td>None</td>
<td>None</td>
<td>T62L18</td>
<td>12</td>
<td>2</td>
<td>00UTC</td>
</tr>
<tr>
<td>1994.3</td>
<td></td>
<td></td>
<td></td>
<td>T62L18</td>
<td>16</td>
<td>10(00UTC) 4(12UTC)</td>
<td>00,12UTC</td>
</tr>
<tr>
<td>2000.6</td>
<td></td>
<td></td>
<td></td>
<td>T126L28(0-2.5) T62L28(2.5-16)</td>
<td></td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>2001.1</td>
<td></td>
<td></td>
<td></td>
<td>T126(0-3.5) T62L28(3.5-16)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004.3</td>
<td></td>
<td></td>
<td></td>
<td>T126L28(0-7.5) T62L28(7.5-16)</td>
<td></td>
<td></td>
<td>00,06,12,18UTC</td>
</tr>
<tr>
<td>2005.8</td>
<td>TSR</td>
<td></td>
<td></td>
<td>T126L28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006.5</td>
<td>ETR</td>
<td></td>
<td></td>
<td>T190L28</td>
<td></td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>2007.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>2010.2</td>
<td>STTP</td>
<td></td>
<td></td>
<td>T190L42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012.2</td>
<td></td>
<td></td>
<td></td>
<td>T254L42 (0-8) T190L42 (8-16)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Estimating and Sampling Initial Errors:
The Breeding Method - 1992

- **DATA ASSIM:** Growing errors due to cycling through NWP forecasts
- **BREEDING:** - Simulate effect of obs by rescaling nonlinear perturbations
  - Sample subspace of most rapidly growing analysis errors
    - Extension of linear concept of Lyapunov Vectors into nonlinear environment
    - Fastest growing nonlinear perturbations
    - Not optimized for future growth –
      - Norm independent
      - Is non-modal behavior important?

References
1. Toth and Kalnay: 1993 BAMS
2. Tracton and Kalnay: 1993 WAF
3. Toth and Kalnay: 1997 MWR

Courtesy of Zoltan Toth
**Bred Vector** (←2006)  

- \(P_#, N#\) are the pairs of positive and negative  
- \(P_1\) and \(P_2\) are independent vectors  
- Simple scaling down (no direction change)

**Ensemble Transform with Rescaling** (2006 ➔)

- \(P_1, P_2, P_3, P_4\) are orthogonal vectors  
- No pairs any more  
- To centralize all perturbed vectors (sum of all vectors are equal to zero)  
- Scaling down by applying mask,  
- The direction of vectors will be tuned by ET.

**References:**

1. Wei and et al: 2006 Tellus  
2. Wei and et al: 2008 Tellus
P#, N# are the pairs of positive and negative
P1 and P2 are independent vectors
Simple scaling down (no direction change)

P1, P2, P3, P4 are orthogonal vectors
No pairs any more
To centralize all perturbed vectors (sum of all vectors are equal to zero)
Scaling down by applying mask,
The direction of vectors will be tuned by ET.

References:
1. Wei and et al: 2006 Tellus
2. Wei and et al: 2008 Tellus
24-hour breeding cycle

- T00Z 10m
- T06Z 10m
- T12Z 10m
- T18Z 10m

24hrs | Re-scaling | Up to 16-d

6-hour breeding cycle

- T00Z 40m
- T06Z 40m
- T12Z 40m
- T18Z 40m

Re-scaling

Next T00Z | Up to 16-d

2004

24-hour breeding cycle

2004

2004

6-hour breeding cycle

2004 →

2004

2004
How do we tune ETR initial perturbations?

500hPa NH

850hPa NH

1000hPa NH

Rescaling mask and factors

Current operation

Future

Schematic of tuning initial perturbations
GFS TS relocation

Fcast/guess

3hrs  6hrs  9hrs

Use GFS Track information

Relocated TS to Observed position

GDAS (SANL)

Fcast

6hrs fcst

P

Use ens. Track information

C

Use GFS Track information

To separate into env. Flow (EF) And storm perturbation (SP)

Ens. Rescaling For EF (p+n)

Ens. Rescaling For SP (p+n)

Combined

Reference:
Liu and et al: 2006 AMS conference extended paper
Hurricane Track Plots (case 1)

Frances (08/28)

With relocation

Without relocation

Reduced initial spread

Large initial spread
Hurricane Tracks Plots (case 2)

Ivan (09/14)

Without relocation

With relocation
Track errors and spreads
2004 Atlantic Basin (8/23-10/1)

From Timothy Marchok (GFDL)

Reduced mean track errors and spreads
Tropical Storm Relocation scheme is still important for today’s ensemble forecast.
Stochastic Total Tendency Perturbation (STTP)  
(Hou, Toth and Zhu, 2006)  

NCEP operation – Feb. 2010

Formulation:  
\[ \frac{\partial X_i}{\partial t} = T_i(X_i; t) + \gamma \sum_{j=1}^{N} w_{i,j} T_j(X_j; t) \]

Simplification: Use finite difference form for the stochastic term

Modify the model state every 6 hours:  
\[ X_i' = X_i + \gamma \sum_{j=1}^{N} w_{i,j}(t) \left\{ (X_j)_t - (X_j)_{t-6h} \right\} - \left\{ (X_0)_t - (X_0)_{t-6h} \right\} \]

Where \( w \) is an evolving combination matrix, and \( \gamma \) is a rescaling factor.

Reference:  
2. Hou and et al: 2010 in review of Tellus
STTP Scheme Application

Generation of Stochastic Combination Coefficients:

- **Matrix Notation** (N forecasts at M points)
  \[ S(t) = P(t) W(t) \]
  \[ M \times N \]

- As \( P \) is quasi orthogonal, an orthonormal matrix \( W \) ensures orthogonality for \( S \).

- **Generation of \( W \) matrix**: (Methodology and software provided by James Purser).
  - a) Start with a random but orthonormalized matrix \( W(t=0) \);
  - b) \( W(t) = W(t-1) R_0 R_1(t) \)

- \( R_0, R(t) \) represent random but slight rotation in N-Dimensional space

![Graph showing combination coefficients for Member 14](image)

- \( w_{ij}(t) \) for \( i=14, \) and \( j=1,14 \)

- Random walk \( (R_1) \) superimposed on a periodic Function \( (R_0) \)
Experiments for 2009 Operational Implementation
**T126L28 vs. T190L28 resolution**, Nov. 2007 Cases
SPS works with both resolutions

- T126L28
- T126L28 + SP
- T190L28
- T190L28 + SP

**Tropical 850hPa Temp.**
Continuous Ranked Probability Skill Scores
Average For 20071101 – 20071129

**CRPSS**

**ROC area (0–1)**
Average For 20071101 – 20071129

**ROC**
Random Model Error – Stochastic Schemes

- Stochastically-perturbed physics tendencies (SPPT) – operational ECMWF scheme.
- Stochastic total tendency perturbation (STTP) – operational NCEP scheme
- Vorticity confinement (VC) – under development at UKMet and ECMWF
- Stochastically-perturbed boundary-layer humidity (SHUM)
- Perturbed convective trigger on SAS
  - Testing on HWRF ensemble and global system
- Questions and issues
  - Will these schemes help to increase ensemble spread?
  - Will these schemes help to reduce missing forecast for extreme weather events?
  - What is the future of physical parameterizations?
    - Deterministic or probabilistic?
Initial Condition = Unperturbed Analysis + EDA-based perturbation + SV-based perturbation

Unperturbed Analysis - 4D-VAR (T₄L1279L91)

EDA-based perturbation - the difference between the perturbed (perturb all obs and sea-surface T and use SPPT to simulate random model error) and unperturbed first-guesses (T₄L399L91)

SV-based perturbation - initial singular vectors (T₄L62)

Current ECMWF ensemble configuration

EDA1 → SV1 → mem1
EDA1 → SV2 → mem2
EDA1 → SV3 → mem3
EDA2 → SV4
EDA2 → SV5
EDA2 → SV6
EDA2 → SV7
EDA2 → SV8
EDA2 → SV9
EDA2 → SV10
EDA10

... → mem50
Similar to EnKF, growing slowly, good CRPS scores.
Canadian Meteorological Center
Ensemble Prediction System (EPS)

Configuration – 2004
(Before July 10 2007)

- CMC EPS started with 8 members (SEF T95) in quasi-operational mode in March 1996.
- became operational in February 1998.
- 8 new models (GEM) were added in August 1999.
- products available on external Web page available in October 1999.
- increased horizontal resolution to T149 in June 2001.

Operational configuration before July 10 2007

- 16 members
- 10 day forecasts done once a day (00Z run)
- perturbed analyses obtained from perturbed assimilation cycles
- multi-model approach: SEF T150 and GEM 1.2º (~130km)
- different model options used for both models
Canadian Meteorological Center
Ensemble Prediction System (EPS)

Configuration - 2004

doubling of the number of analyses

perturbed trial fields
perturbed observations

perturbed analyses

models

8 models each producing A data assimilation cycle

16 forecast cycles
  • 8 SEF - T149 or ~150 km
  • 8 GEM - 1.2° or ~135 km
  • 10 days at 00 UTC

16 members
# Combination of model perturbations (2004)

<table>
<thead>
<tr>
<th>SEF (T149)</th>
<th>Add ops analysis</th>
<th>Convection/Radiation</th>
<th>GWD</th>
<th>GWD version</th>
<th>Orography</th>
<th>Number of levels</th>
<th>Time level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>yes</td>
<td>Kuo/ GarandStrong</td>
<td>High altitude</td>
<td>0.3</td>
<td>23</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>no</td>
<td>Manabe/ Sasamori</td>
<td>Strong</td>
<td>Low altitude</td>
<td>0.3</td>
<td>41</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>no</td>
<td>Kuo/ GarandWeak</td>
<td>Low altitude</td>
<td>Mean</td>
<td>23</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>yes</td>
<td>Manabe/ Sasamori</td>
<td>Weak</td>
<td>High altitude</td>
<td>Mean</td>
<td>41</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>yes</td>
<td>Manabe/ Sasamori</td>
<td>Strong</td>
<td>Low altitude</td>
<td>Mean</td>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>no</td>
<td>Kuo/ GarandStrong</td>
<td>High altitude</td>
<td>Mean</td>
<td>41</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>no</td>
<td>Manabe/ Sasamori</td>
<td>Weak</td>
<td>High altitude</td>
<td>0.3</td>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>yes</td>
<td>Kuo/ GarandWeak</td>
<td>Low altitude</td>
<td>0.3</td>
<td>41</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>control</td>
<td>mean</td>
<td>Kuo/ GarandMean</td>
<td>Low altitude</td>
<td>0.15</td>
<td>41</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GEM (1.2)</th>
<th>Add ops analysis</th>
<th>Deep convection</th>
<th>Shallow convection</th>
<th>Soil moisture</th>
<th>Sponge</th>
<th>Number of levels</th>
<th>Coriolis</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>no</td>
<td>Kuosym</td>
<td>new</td>
<td>Less 20%</td>
<td>global</td>
<td>28</td>
<td>Implicit</td>
</tr>
<tr>
<td>10</td>
<td>yes</td>
<td>RAS</td>
<td>old</td>
<td>Less 20%</td>
<td>equatorial</td>
<td>28</td>
<td>Implicit</td>
</tr>
<tr>
<td>11</td>
<td>yes</td>
<td>RAS</td>
<td>old</td>
<td>Less 20%</td>
<td>global</td>
<td>28</td>
<td>Implicit</td>
</tr>
<tr>
<td>12</td>
<td>no</td>
<td>Kuosym</td>
<td>old</td>
<td>More 20%</td>
<td>global</td>
<td>28</td>
<td>Implicit</td>
</tr>
<tr>
<td>13</td>
<td>no</td>
<td>Kuosym</td>
<td>new</td>
<td>More 20%</td>
<td>global</td>
<td>28</td>
<td>Implicit</td>
</tr>
<tr>
<td>14</td>
<td>yes</td>
<td>Kuosym</td>
<td>new</td>
<td>Less 20%</td>
<td>global</td>
<td>28</td>
<td>Implicit</td>
</tr>
<tr>
<td>15</td>
<td>yes</td>
<td>Kuosym</td>
<td>old</td>
<td>Less 20%</td>
<td>global</td>
<td>28</td>
<td>Implicit</td>
</tr>
<tr>
<td>16</td>
<td>no</td>
<td>OldKuo</td>
<td>new</td>
<td>More 20%</td>
<td>global</td>
<td>28</td>
<td>Implicit</td>
</tr>
</tbody>
</table>
## CMC’s Multi-model EPS for the assimilation (current-GEM)

<table>
<thead>
<tr>
<th>#</th>
<th>Deep convection</th>
<th>Surface scheme</th>
<th>Mixing length</th>
<th>Vertical mixing parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Kain &amp; Fritsch</td>
<td>ISBA</td>
<td>Bougeault</td>
<td>1.0</td>
</tr>
<tr>
<td>2</td>
<td>Oldkuo</td>
<td>ISBA</td>
<td>Blackadar</td>
<td>0.85</td>
</tr>
<tr>
<td>3</td>
<td>Relaxed Arakawa Schubert</td>
<td>force-restore</td>
<td>Bougeault</td>
<td>0.85</td>
</tr>
<tr>
<td>4</td>
<td>Kuo Symétrique</td>
<td>force-restore</td>
<td>Blackadar</td>
<td>1.0</td>
</tr>
<tr>
<td>5</td>
<td>Oldkuo</td>
<td>force-restore</td>
<td>Bougeault</td>
<td>1.0</td>
</tr>
<tr>
<td>6</td>
<td>Kain &amp; Fritsch</td>
<td>force-restore</td>
<td>Blackadar</td>
<td>0.85</td>
</tr>
<tr>
<td>7</td>
<td>Kuo Symétrique</td>
<td>ISBA</td>
<td>Bougeault</td>
<td>0.85</td>
</tr>
<tr>
<td>8</td>
<td>Relaxed Arakawa Schubert</td>
<td>ISBA</td>
<td>Blackadar</td>
<td>1.0</td>
</tr>
<tr>
<td>9</td>
<td>Kain &amp; Fritsch</td>
<td>ISBA</td>
<td>Blackadar</td>
<td>0.85</td>
</tr>
<tr>
<td>10</td>
<td>Oldkuo</td>
<td>ISBA</td>
<td>Bougeault</td>
<td>1.0</td>
</tr>
<tr>
<td>11</td>
<td>Relaxed Arakawa Schubert</td>
<td>force-restore</td>
<td>Blackadar</td>
<td>1.0</td>
</tr>
<tr>
<td>12</td>
<td>Kuo Symétrique</td>
<td>force-restore</td>
<td>Bougeault</td>
<td>0.85</td>
</tr>
<tr>
<td>13</td>
<td>Oldkuo</td>
<td>force-restore</td>
<td>Blackadar</td>
<td>0.85</td>
</tr>
<tr>
<td>14</td>
<td>Kain &amp; Fritsch</td>
<td>force-restore</td>
<td>Bougeault</td>
<td>1.0</td>
</tr>
<tr>
<td>15</td>
<td>Kuo Symétrique</td>
<td>ISBA</td>
<td>Blackadar</td>
<td>1.0</td>
</tr>
<tr>
<td>16</td>
<td>Relaxed Arakawa Schubert</td>
<td>ISBA</td>
<td>Bougeault</td>
<td>0.85</td>
</tr>
<tr>
<td>17</td>
<td>Kuo Symmetric</td>
<td>force-restore</td>
<td>Bougeault</td>
<td>1.0</td>
</tr>
<tr>
<td>18</td>
<td>Kain &amp; Fritsch</td>
<td>ISBA</td>
<td>Blackadar</td>
<td>0.85</td>
</tr>
<tr>
<td>19</td>
<td>Oldkuo</td>
<td>ISBA</td>
<td>Bougeault</td>
<td>0.85</td>
</tr>
<tr>
<td>20</td>
<td>Relaxed Arakawa Schubert</td>
<td>force-restore</td>
<td>Blackadar</td>
<td>1.0</td>
</tr>
<tr>
<td>21</td>
<td>Relaxed Arakawa Schubert</td>
<td>ISBA</td>
<td>Blackadar</td>
<td>0.85</td>
</tr>
<tr>
<td>22</td>
<td>Oldkuo</td>
<td>force-restore</td>
<td>Bougeault</td>
<td>1.0</td>
</tr>
<tr>
<td>23</td>
<td>Kain &amp; Fritsch</td>
<td>force-restore</td>
<td>Blackadar</td>
<td>1.0</td>
</tr>
<tr>
<td>24</td>
<td>Kuo Symétrique</td>
<td>ISBA</td>
<td>Bougeault</td>
<td>0.85</td>
</tr>
</tbody>
</table>

P. Houtekamer, ARMA
Changes to the 16 day forecast system

P. Houtekamer, ARMA

• Like in the EnKF :
  – Use of a more recent version of the model and the model physics,
  – Removing an old surface scheme,
  – 20 minute time step,
  – Use of a topography filter,

• No perturbation of model physics when convection is active,

• No longer ramping down the stochastic physics in the tropics.

• **With these changes the system is a lot more robust** (on occasion with the currently operational system we have to rerun an integration).

• Implementation: Jan/Feb 2013
Highlights

• Description of current operational global ensemble forecast systems
  – NCEP global ensemble forecast system
  – ECMWF global ensemble forecast system
  – CMC global ensemble forecast system

• Ensemble performance
  – Compare to deterministic forecast
  – Multi-model ensemble

• Interaction between data assimilation and ensemble forecast
  – Hybrid 3DV/EnKF implementation

• Future direction
  – EMC’s plan
  – Preliminary study/discussion
  – Challenges
NH Anomaly Correlation for 500hPa Height

Period: January 1st – December 31st 2012

Anomaly Correlation vs. Forecast (days)

- GFS
- GEFS
- NAEFS

Skillful forecast

Ensemble mean
Day at which forecast loses useful skill (AC=0.6)
N. Hemisphere 500hPa height calendar year means

<table>
<thead>
<tr>
<th>Year</th>
<th>GFS</th>
<th>GEFS</th>
<th>NAEFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>8</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>2008</td>
<td>8</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>2009</td>
<td>8</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>2010</td>
<td>8</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>2011</td>
<td>8</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>2012</td>
<td>8</td>
<td>8</td>
<td>10</td>
</tr>
</tbody>
</table>
Continuous Ranked Probability Skill Scores

NH 500hPa Height 30−d Running Mean

5 days forecast

10 days forecast
Atlantic, AL01~19 (06/01~12/31/2011)
Atlantic, AL01~19 (05/01~11/27/2012)

**Track error (NM)**

**Forecast hours**

**CASES - 2011:**
- 0: 393
- 12: 356
- 24: 314
- 36: 274
- 48: 238
- 72: 182
- 96: 141
- 120: 98
- 144: 71
- 168: 49

**CASES - 2012:**
- 0: 439
- 12: 399
- 24: 357
- 36: 315
- 48: 277
- 72: 219
- 96: 178
- 120: 143
- 144: 119
- 168: 94

**GEFS-11**
- GEFS T190L28 (operational run)

**GEFS-12**
- GEFS T254L42 (operational run)
Track Forecast Error for Atlantic 2012 Season

GFS – NCEP deterministic forecast
GEFS – NCEP ensemble forecast
EC_det – ECWMF deterministic forecast
EC_ens – ECMWF ensemble forecast

NCEP made better forecast
ECMWF made better forecast

FCST HOURS Cases
0 170
12 151
24 137
36 123
48 110
72 90
96 73
120 57
Highlights

• Description of current operational global ensemble forecast systems
  – NCEP global ensemble forecast system
  – ECMWF global ensemble forecast system
  – CMC global ensemble forecast system

• Ensemble performance
  – Compare to deterministic forecast
  – Multi-model ensemble

• Interaction between data assimilation and ensemble forecast
  – Hybrid 3DV/EnKF implementation

• Future direction
  – EMC’s plan
  – Preliminary study/discussion
  – Challenges
Interactions between DA and EPS

- Ideally, EPS and DA systems should be consistent for best performance of both.

- DA provides best estimates of initial uncertainties, i.e. analysis error covariance for EPS.

- EPS produces accurate flow dependent forecast (background) covariance for DA.

\[
\begin{align*}
\text{DA} & \quad \text{Best analysis error variances (} P^a \text{)} \quad \text{EPS} \\
\text{} \quad \text{Accurate forecast error covariance } P^f \quad \text{}
\end{align*}
\]
NCEP Dual-Res Coupled Hybrid DA System

Generate new ensemble perturbations given the latest set of observations and first-guess ensemble

Ensemble contribution to background error covariance

Replace the EnKF ensemble mean analysis

Generate new ensemble perturbations given the latest set of observations and first-guess ensemble

Ensemble contribution to background error covariance

T254L64

member 1 forecast

member 2 forecast

member 3 forecast

T574L64

high res forecast

GSI Hybrid Ens/Var

high res analysis

Previous Cycle

Current Update Cycle

Congratulations of Daryl Kleist
GFS – T572L64
GEFS – T254L42
Highlights

• Description of current operational global ensemble forecast systems
  – NCEP global ensemble forecast system
  – ECMWF global ensemble forecast system
  – CMC global ensemble forecast system

• Ensemble performance
  – Compare to deterministic forecast
  – Multi-model ensemble

• Interaction between data assimilation and ensemble forecast
  – Hybrid 3DV/EnKF implementation

• Future direction
  – EMC’s plan
  – Preliminary study/discussion
  – Challenges
EMC’s Plan to Reframe GEFS Initializations

• Background:
  – **BV-ETR**: It is NCEP Global Ensemble Forecast System (GEFS) initialization since 1992 – which is dynamically breading orthogonal, fast growing perturbations in region of high baroclinicity.
  – **HVEDAS**: Hybrid Variational Ensemble Data Assimilation System has been implemented on May 22\(^{nd}\) to deliver better quality of analysis (or initial condition of forecast) through improved background error covariance from EnKF 6hr forecasts.
  – **EnKF**: Ensemble Kalman Filter data assimilation has been implemented on May 22\(^{nd}\) for HVEDAS, which evolves an ensemble over data assimilation, updated at successive observation times.
  – **Evaluations**: To assess the difference by comparing BV-ETR and EnKF (F06) initialized ensemble forecast (show statistics – next slide).

• Motivations:
  – To reduce computational cost of double cycling of 80 members ensemble short forecasts.
  – To enhance our Global Ensemble Forecast System (GEFS).
  – To take the best of BV-ETR and EnKF, improving ensemble initial perturbations and forecast (in studying).

• In reality:
  – In daily operation, EnKF and HVEDAS run later (final - +6hrs) than GEFS (+4.5hrs). Therefore, EnKF (F06) from previous cycle could be only one for possible GEFS initial perturbations. In fact, EnKF (F06) perturbations are not ideally (optimum) representing analysis uncertainties. Additional processes, such as rescaling, adjustment and et al. are necessary.
Comparison of ETR vs EnKF (f06) initialized ensemble forecasts

Summary of comparison:
(Three seasons – Summer 2011, Winter 2011/2012 and Summer 2012)

1. For Northern Hemisphere ensemble mean and probabilistic forecasts – they are very similar to each other; the differences are insignificant mostly. EnKF has a little better probabilistic forecast for very short lead time due to larger spread.

2. For Southern Hemisphere mean and probabilistic forecast – ETR is better than EnKF, especially for the two summer seasons. EnKF is over dispersion for SH in generally.

3. Tropical Storm track forecast, ETR and EnKF have similar forecast errors and ensemble spread for two summer seasons. The differences are insignificant.

4. For summer 2012, ETR is better than EnKF’s performance over all.
EnKF06

Artificial inflation and centralization

EnKF update with observations, perturbations will be reduced usually, inflation is needed.

Perturbations are decayed for first 3-9 hours because most of inflated perturbations are non-growing, white noise.

ETR 6 hr fcst

18Z

00Z

EnKF 6hr fcst

3DETR

Schematic diagram of 00Z initial perturbations

Apply ET and 3D mask

3D mask is generated from accumulated EnKF analysis variance. In general, it masks out very fast growing and saturated modes.

ET is making initial perturbations orthogonally.

18Z

00Z

Advantage:
1. Selective potential growing mode;
2. All perturbations are orthogonal;
3. Continuation of vector (member) from cycle to cycle.
In general, breeding method is more conception, and SV is more practical.

Since we don’t know the size of initial uncertainties, we believe that smaller initial perturbations will be better (if it grows faster and catch up forecast errors).

Fig. 6. Schematic of the time evolution of the rms amplitude of high-energy baroclinic modes and low-energy convective modes. Note that although initially growing much faster than the baroclinic modes, convective modes saturate at a substantially lower level.

Early study from Zoltan Toth: BAMS 1992
2% decaying accumulation of 80 EnKF analysis errors (EnKF-siganl)

MASK in NH (left), SH (middle) and Tropics (right)

\[ TE' = \left\langle \frac{1}{80} \sum_{i=1}^{80} \sqrt{\frac{1}{2} \left( u_i'^2 + v_i'^2 + \frac{c_p}{T_r} T_i'^2 \right)} \right\rangle \]
Investigation/understanding current EnKF analysis and inflated analysis

Initial time: 2012090100

Black-sanl, original EnKF analysis perturbations

Red-siganl, inflated EnKF analysis (approximately 36% additive) and centerization

Green-f03, 3-hour forecast from inflated analysis

Blue-f06, 6-hour forecast

Orange-f09, 9-hour forecast

Purple- siganl/sanl, inflation rate for different region and different variables

Courtesy of Dr. Jessie Ma
Preliminary experiment

- Using inflated perturbations as statistical reference (3D mask) applying to ET

The skills should be very similar to EnKF’s at first try.

Note: 3D mask is decaying average (w=0.02) of total energy norm from EnKF analysis (without inflation)
Perturbations before/after 3DETR

Black – before

Red – after

2011081700

We should be able to tune this
The experiments will show the difference between ETR_2D (current operation) and ETR_3D (using EnKF analysis variance) and EnKF (f06).
ETR_3D is smaller, but catch up after 24 hours. It means ETR_3D has fast growth rate.
ETR_3D and EnKF have the same size of initial perturbation, but ETR is growing faster.

We should be able to reduce ETR perturbations easily.
ETR_3D initial perturbations are much smaller than EnKF, but getting closer after 12 hours.
Tropical temperature: the perturbations for ETR_3D is much smaller than EnKF.
For Discussion

• Do we really know the analysis errors?
  – Assume we know, can we use it directly as ensemble initial perturbation?
  – If we don’t, what is a good initial perturbation for global ensemble forecast system?
  – How to explain/understand Ricardo’s recent work? – Hybrid DA system without filtering analysis

• What is a best approach for imperfect numerical model?
  – Multi-model system
  – Multi-parameterization
  – Stochastic perturbation (physics) schemes

• Feedback from users
  – Under-forecast for extreme weather

• Future collaboration
Thanks!!!
LYAPUNOV, SINGULAR, AND BRED VECTORS

• **LYAPUNOV VECTORS (LLV):**
  - Linear perturbation evolution
  - Fast growth
  - Sustainable
  - Norm independent
  - Spectrum of LLVs

\[ \lambda_i = \lim_{t \to \infty} \frac{1}{t} \log_2 \left( \frac{p_i(t)}{p_i(t_0)} \right) \]

• **SINGULAR VECTORS (SV):**
  - Linear perturbation evolution
  - Fastest growth
  - Transitional (optimized)
  - Norm dependent
  - Spectrum of SVs

\[ \|x(t)\|^2 = \langle \mathcal{L}E L x_0; x_0 \rangle \]

• **BRED VECTORS (BV):**
  - Nonlinear perturbation evolution
  - Fast growth
  - Sustainable
  - Norm independent
  - Can orthogonal (Boffeta et al)

\[ \frac{dv}{dt} = av(1-v) \]

Local Lyapunov Vector (LLV)

T10, L18 MRF experiments, Szunyogh et al, 1998

SV with 24-hour optimization

Random balanced perturbation

Courtesy of Zoltan Toth
In applied mathematics and dynamical system theory, Lyapunov vectors, named after Aleksandr Lyapunov, describe characteristic expanding and contracting directions of a dynamical system. They have been used in predictability analysis and as initial perturbations for ensemble forecasting in numerical weather prediction.[1] In modern practice they are often replaced by bred vectors for this purpose.[2]

Analysis-error covariance
singular vector review

• Let $S$ be a symmetric, non-negative matrix defining a norm for the forecast errors. $S = DD^T$ (total energy)

• $M$ is tangent-linear operator, $x'^f$ is forecast perturbation, $x'^a$ is analysis perturbation, $P^a$ is analysis-error covariance.

• AEC SVs seek to maximize the forecast perturbation magnitude in this norm, subject to the constraint that initial perturbation is consistent with analysis error statistics.

\[
\max \frac{x'^f x'^f}{x'^a P^a x'^a} \quad \text{or} \quad \max \frac{x'^a M^T S M x'^a}{x'^a P^a x'^a}
\]