NCEP Global Ensemble Forecast System (GEFS) - Review

Yuejian Zhu
Ensemble Team Leader
Environmental Modeling Center
NCEP/NWS/NOAA

March 2011

Highlights

- Familiar to EMC ensemble team and collaborators.
- Why do we need ensemble?
- NCEP ensemble (GEFS) milestones
- Ensemble initialization and cycling
- Tropical storm relocation
- Stochastic total tendency perturbation (STTP)
- Multi-model ensemble application – NAEFS
- Future plan

Team members

- Yuejian Zhu
  - Lead and over all planning
- Dingchen Hou
  - GEFS implementation
  - Model STTP (THORPEX proposal)
  - Post processing
  - River ensembles
- Mozheng Wei
  - Ensemble initialization
  - HVEDS (THORPEX proposal)
  - HFIP high resolution demonstration
- Richard Wobus
  - Code manager
  - GEFS implementation
  - TIGGE and NAEFS data exchange
- Malaquias Pena
  - Intraseasonal forecast calibration
  - Coupling GEFS and CFS ensembles (CTB proposal)
- Yucheng Song
  - WSR (winter storm reconnessass)
  - Targeting observation (THORPEX proposal)
- Bo Cui
  - Ensemble post processing (THORPEX proposal)
  - NAEFS/UNOPC and GEFS post process implementation
  - Forecast evaluations
- Yan Luo
  - Precipitation forecast calibration (THORPEX/HYDRO proposal)
  - Precipitation analysis (CCPA)
- Jiayi Peng
  - HFIP post processing
  - Track verifications, TIGGE cxml track data exchanges
- Juhui (Jessie) Ma
  - PhD student
  - Ensemble initialization and configuration
- Jun Du
  - SREF code manager
  - SREF leading implementation
- Bo Yang
  - SREF post processing
  - SREF initialization (ETR)
- Yuqiu (Julia) Zhu
  - Ensemble RFCs
  - Real time experiments setting up and run
- Zhan Zhang
  - HFIP high resolution demonstration
  - HWRF ensembles
- Weiyu Yang
  - ESMF for ensembles
  - MOM4 for GFS/CFS coupling
- George Vandenberghe
  - HFIP high resolution demonstration
  - NOAA HPC - research
- Mary Hart
  - Ensemble web master
- Shrinivas Moorthi
  - GFS model consulting

Collaborators

- International:
  - NAEFS
    - Meteorological Service of Canada (MSC)
    - National Meteorological Service of Mexico
  - ECMWF/UKMet
    - Ensemble development and application
  - CMA/KMA/JMA/Roshydromet
    - WMO/RDP – Beijing Olympic demonstration project
    - Exchange visitors for ensemble development
- National
  - THORPEX - Earth System Research Lab (ESRL)
  - NUOPC
    - FNMOC and NRL
    - AFWA
  - THORPEX-HYDRO - OHD
  - Ensemble post processing – OST/MDL
  - NCEP service centers
    - SPC – storm probabilistic guidance
    - OPC – wave probabilistic guidance
    - TPC – hurricane probabilistic guidance
    - HPC – 1-7 days probabilistic guidance
    - CPC – week-2 forecast (precipitation and temperature)
- Universities

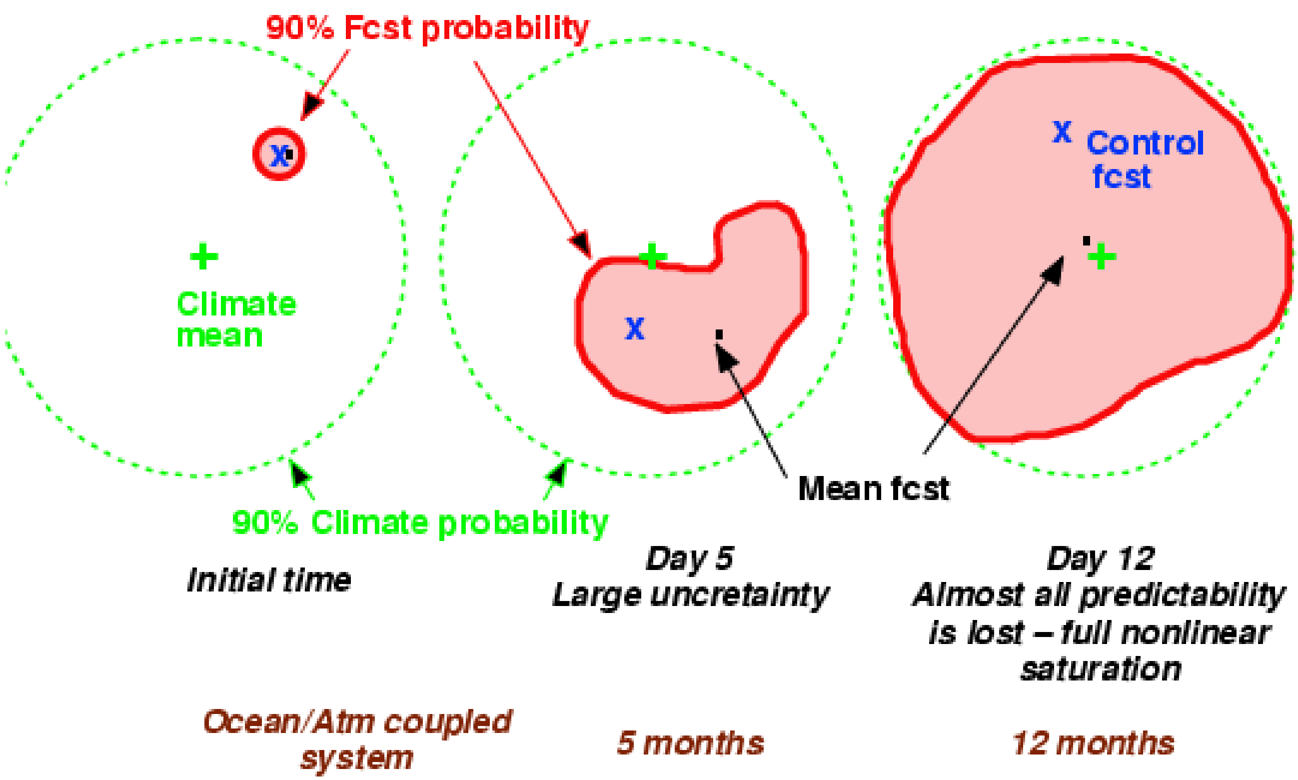
Scientific Needs – Ensemble forecast System
Describe Forecast Uncertainty Arising Due To Chaos

ORIGIN OF FORECAST UNCERTAINTY

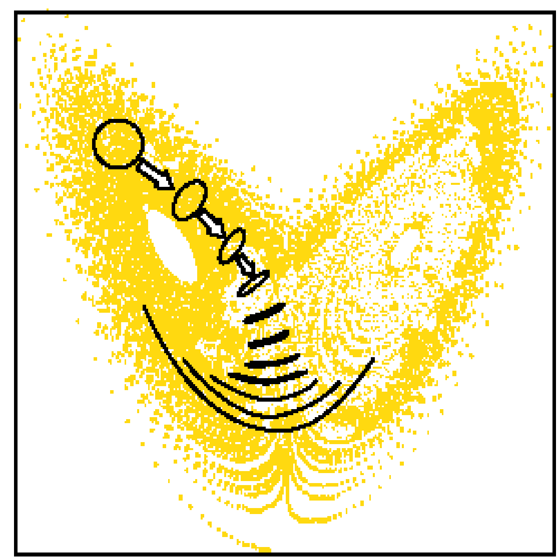
1) The atmosphere is a **deterministic system** *AND* has at least one direction in which **perturbations grow**

2) **Initial** state (and model) has **error** in it ==>

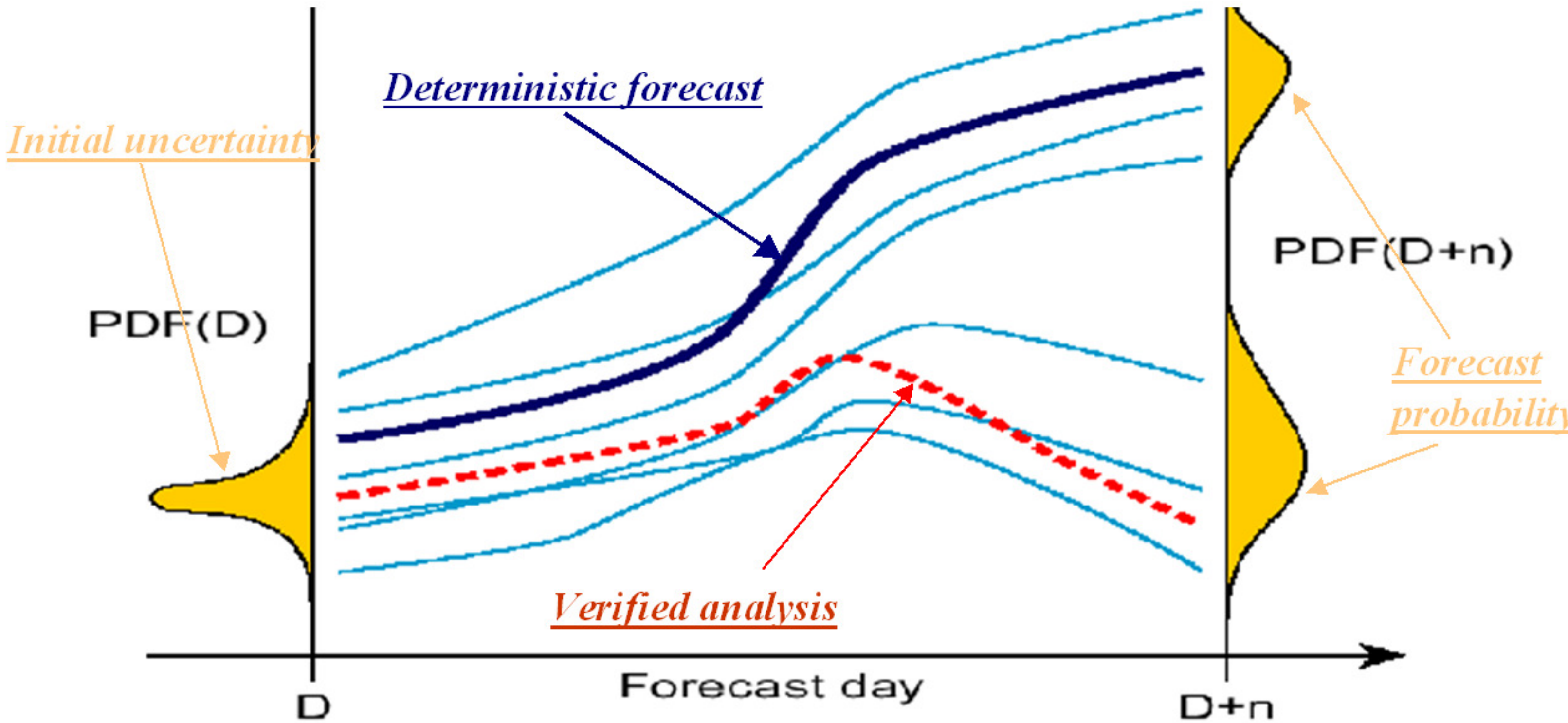
Chaotic system + Initial error =(Loss of) Predictability



Buizza 2002

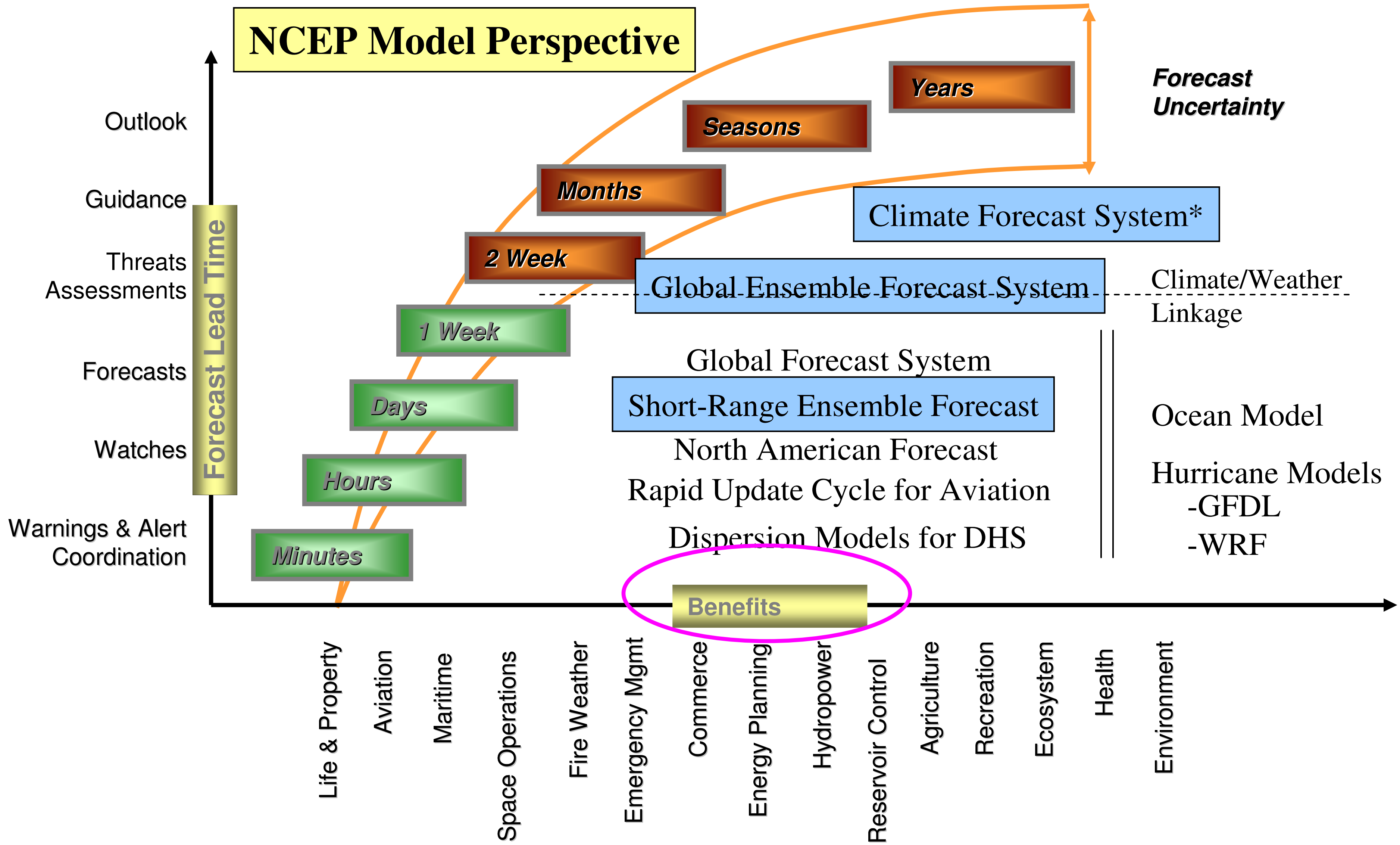


What is the difference of deterministic and ensemble forecast?



Schematic of Stochastic Prediction: The initial probability PDF(D) represents the initial uncertainties. From the best estimate of the initial state a single deterministic forecast (blue solid curve) is performed. This single deterministic forecast fails to predict correctly the future state (red dotted curve). An ensemble of perturbed forecasts (thin blue solid curves) starting from perturbed initial conditions designed to sample the initial un- certainties can be used to estimate the probability PDF(D+n) at future time, D+n.

NWS Seamless Suite of Forecast Products Spanning Climate and Weather



Milestones of GEFS development

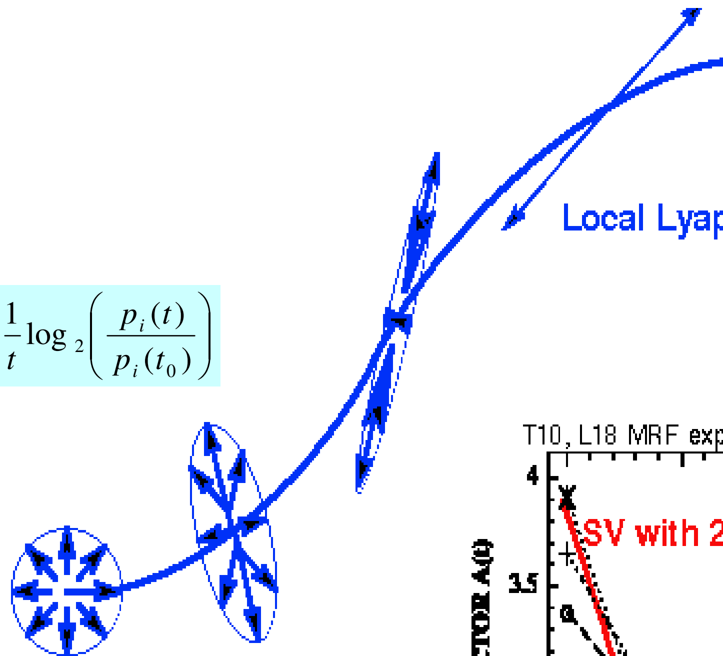
- 1992 – GEFS was in NCEP operation
  - 00UTC only, T62 (220km), 2+1 members, out to 12 days
- 1994 – GEFS run twice per day
  - 00UTC 10+1 members, 12UTC 4+1 members
  - Out to 16 days
- 2000 – GEFS increased resolution
  - T126L28 (110km) for first 60 hours
- 2004 – GEFS run four times per day
  - 00UTC, 06UTC, 12UTC and 18UTC; 40+4 members
- 2005 – Introduced TS relocation (TSR)
- 2006 – ETR replaced BV, 6-h cycling instead of 24-h
- 2007 – Full size GEFS, 80+4 members per day
  - 20 perturbed ensembles plus control
- 2010 – introduced STTP, increased resolution (70km)
  - T190L28 resolution
- 2012 – Increasing resolution (~50km)
  - T254L42 resolution, tuning TSR, ETR and STTP

Ensemble initializations and cycling

LYAPUNOV, SINGULAR, AND BRED VECTORS

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

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

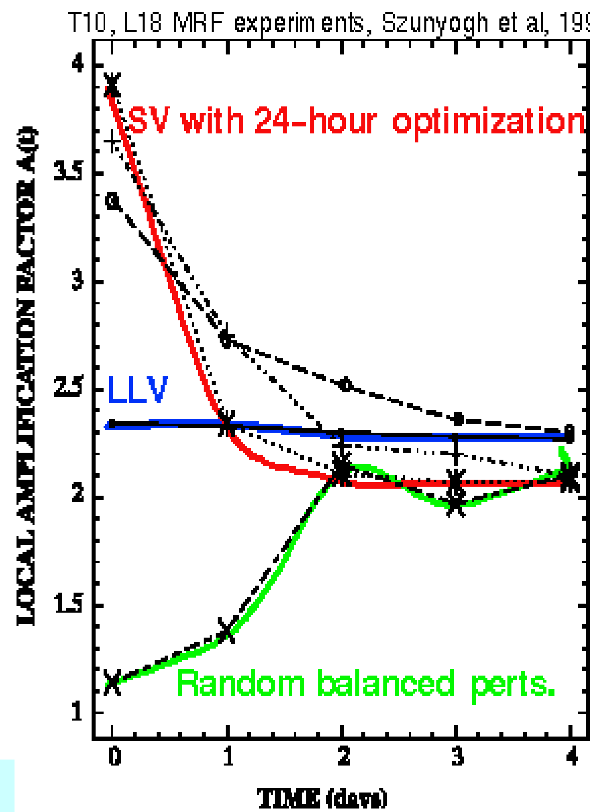


- **SINGULAR VECTORS (SV):**
  - Linear perturbation evolution
  - Fastest growth
  - Transitional (optimized)
  - Norm dependent
  - Spectrum of SVs

$$\|x(t)\|^2 = < L^* E\, L\, x_0 ; x_0 >$$

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

$$\frac{dv}{dt} = av\,(1 - v)$$



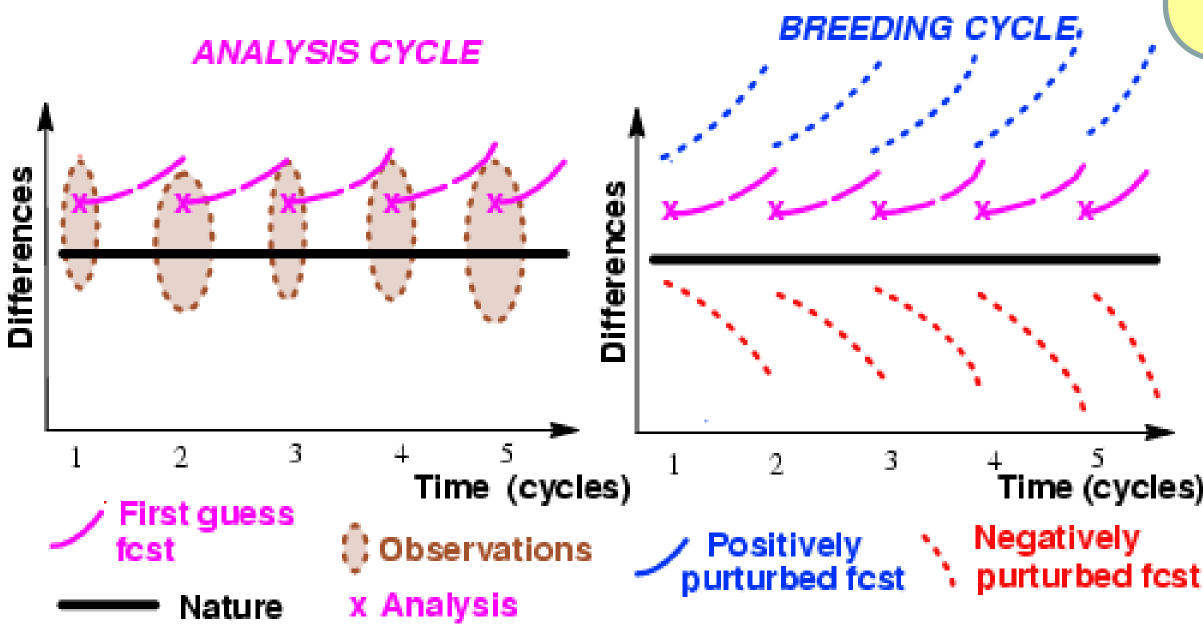
Courtesy of Zoltan Toth

ESTIMATING AND SAMPLING INITIAL ERRORS: THE BREEDING METHOD - 1992

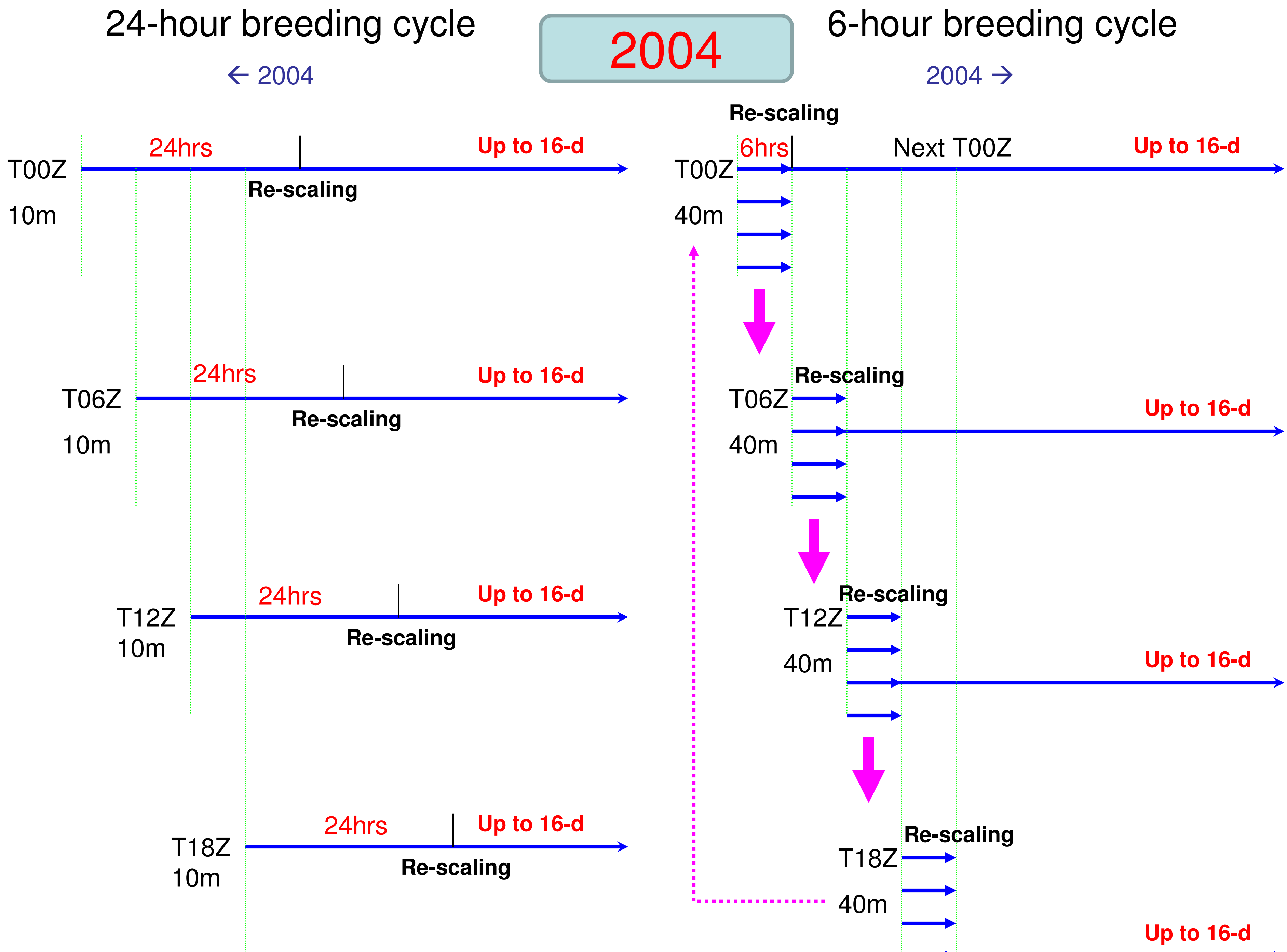
- **DATA ASSIM:** Growing errors due to cycling through NWP forecasts
- **BREEDING:** - Simulate effect of obs by rescaling nonlinear perturbations
  - Sample subspace of most rapidly growing analysis errors
    - Extension of linear concept of Lyapunov Vectors into nonlinear environment
    - Fastest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

References

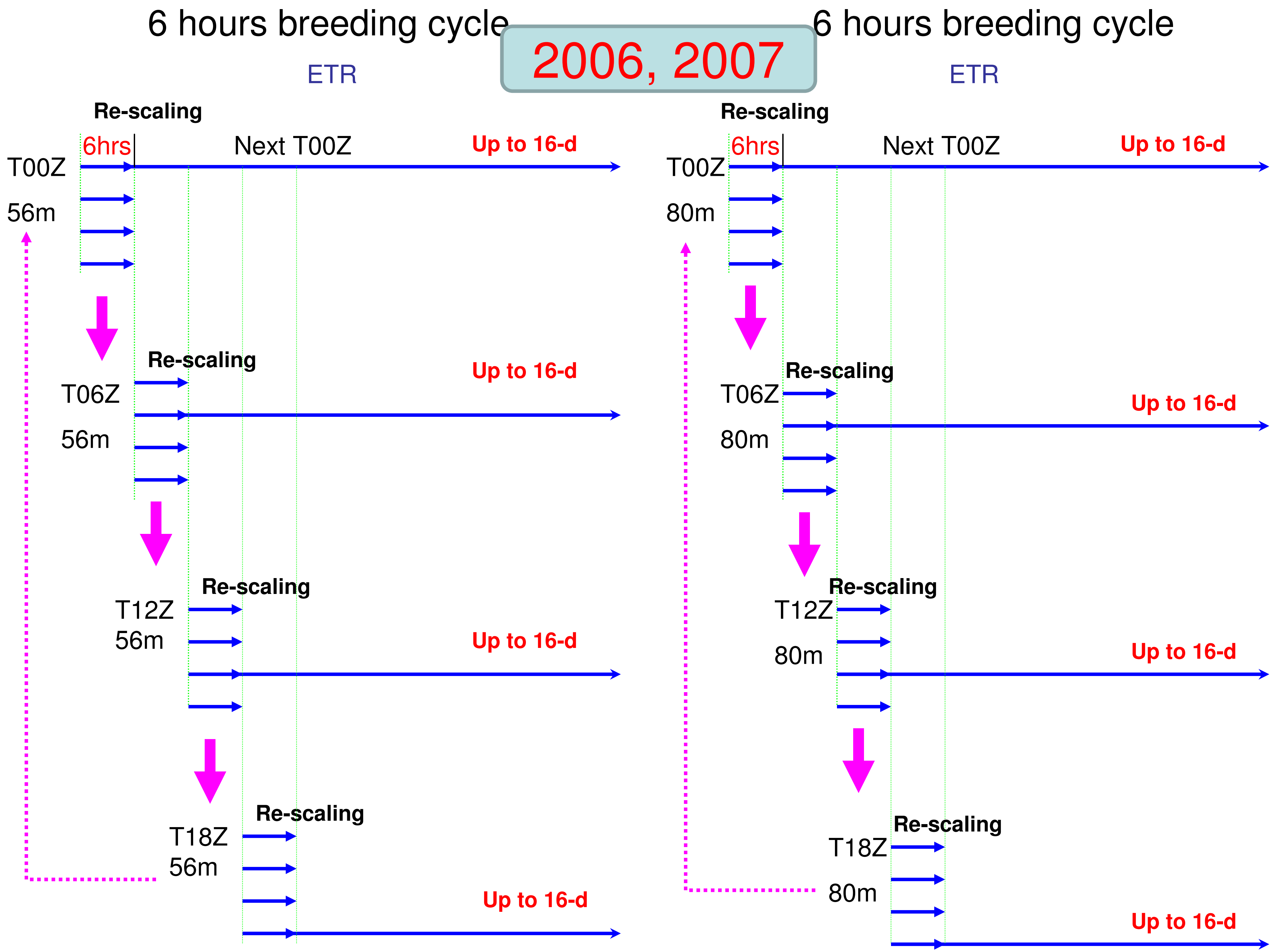
1. Toth and Kalnay: 1993 BAMS
2. Tracton and Kalnay: 1993 WAF



Courtesy of Zoltan Toth

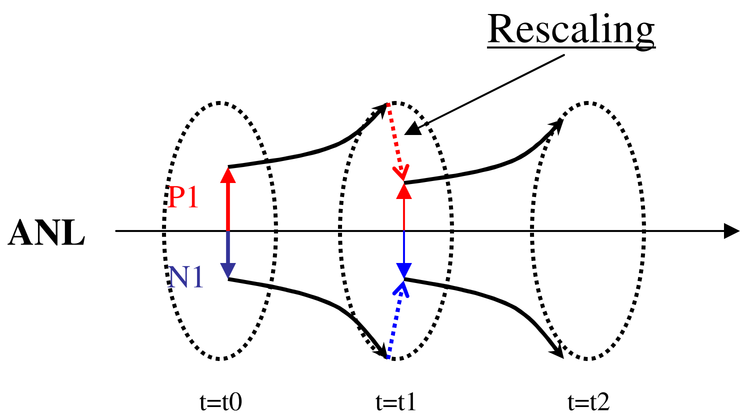


24-hour breeding cycle
← 2004
2004
6-hour breeding cycle
2004 →
T00Z
10m
24hrs
Up to 16-d
Re-scaling
T06Z
10m
24hrs
Up to 16-d
Re-scaling
T12Z
10m
24hrs
Up to 16-d
Re-scaling
T18Z
10m
24hrs
Up to 16-d
Re-scaling
Re-scaling
6hrs
Next T00Z
Up to 16-d
T00Z
40m
Re-scaling
T06Z
40m
Up to 16-d
Re-scaling
T12Z
40m
Up to 16-d
Re-scaling
T18Z
40m
Up to 16-d


6 hours breeding cycle
ETR
2006, 2007
6 hours breeding cycle
ETR
Re-scaling
6hrs
Next T00Z
Up to 16-d
T00Z
56m
Re-scaling
Up to 16-d
T06Z
56m
Re-scaling
T12Z
56m
Up to 16-d
Re-scaling
T18Z
56m
Up to 16-d
Re-scaling
6hrs
Next T00Z
Up to 16-d
T00Z
80m
Re-scaling
T06Z
Up to 16-d
80m
Re-scaling
T12Z
80m
Up to 16-d
Re-scaling
T18Z
80m
Up to 16-d

Bred Vector
(←2006)

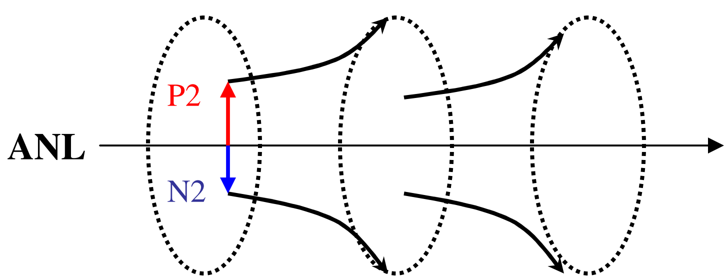
2006



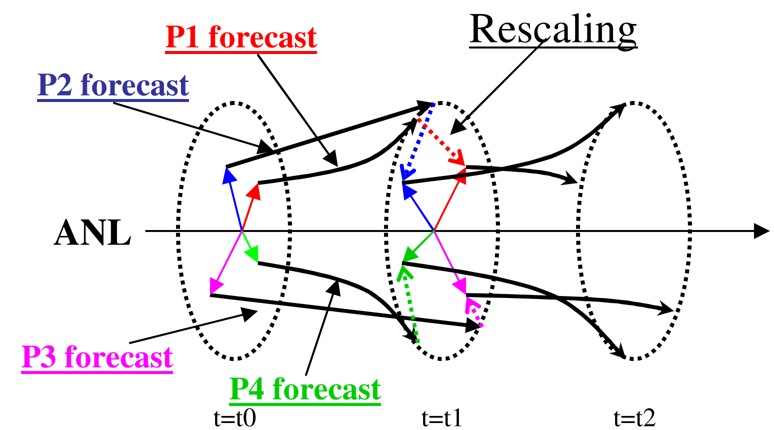
P#, N# are the pairs of positive and negative

P1 and P2 are independent vectors

Simple scaling down (no direction change)



Ensemble Transform with Rescaling
(2006 →)



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,

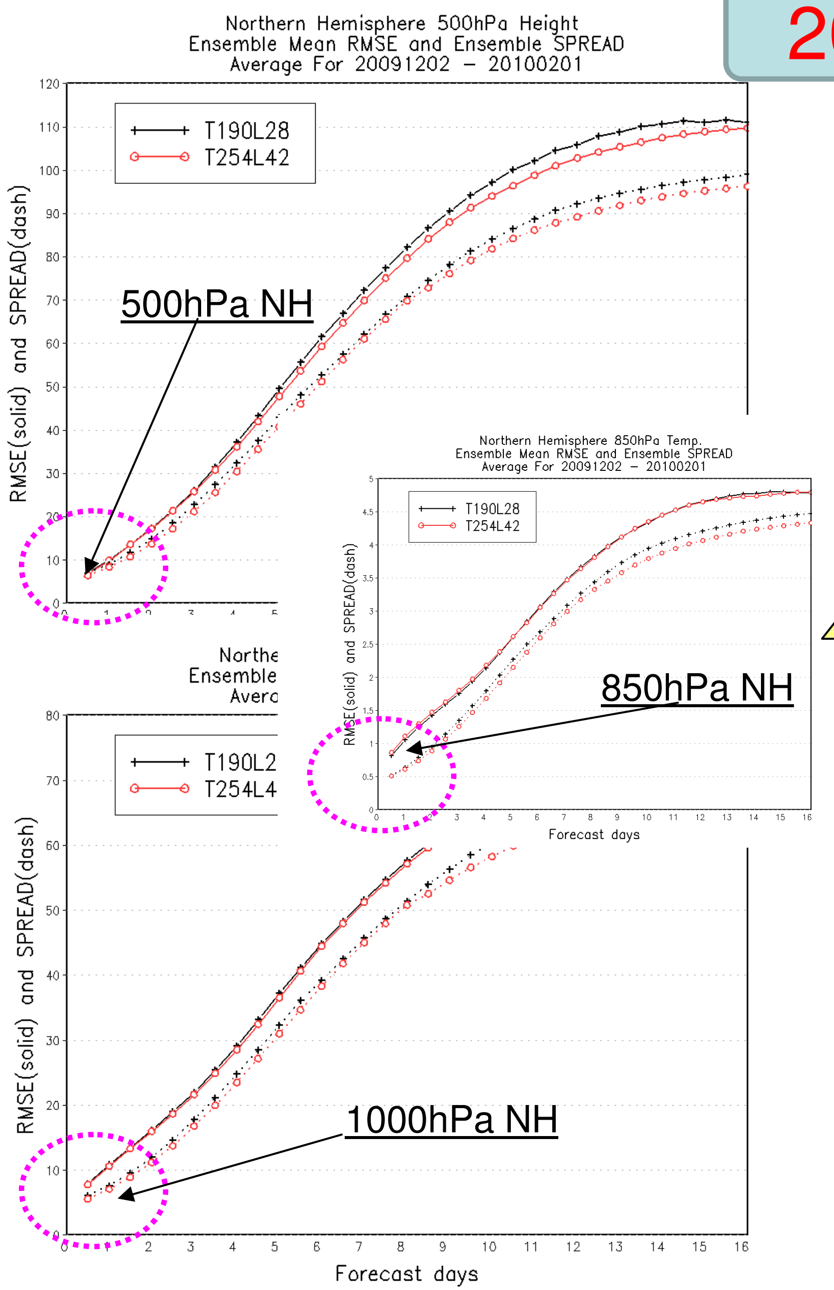
The direction of vectors will be tuned by ET.

References:

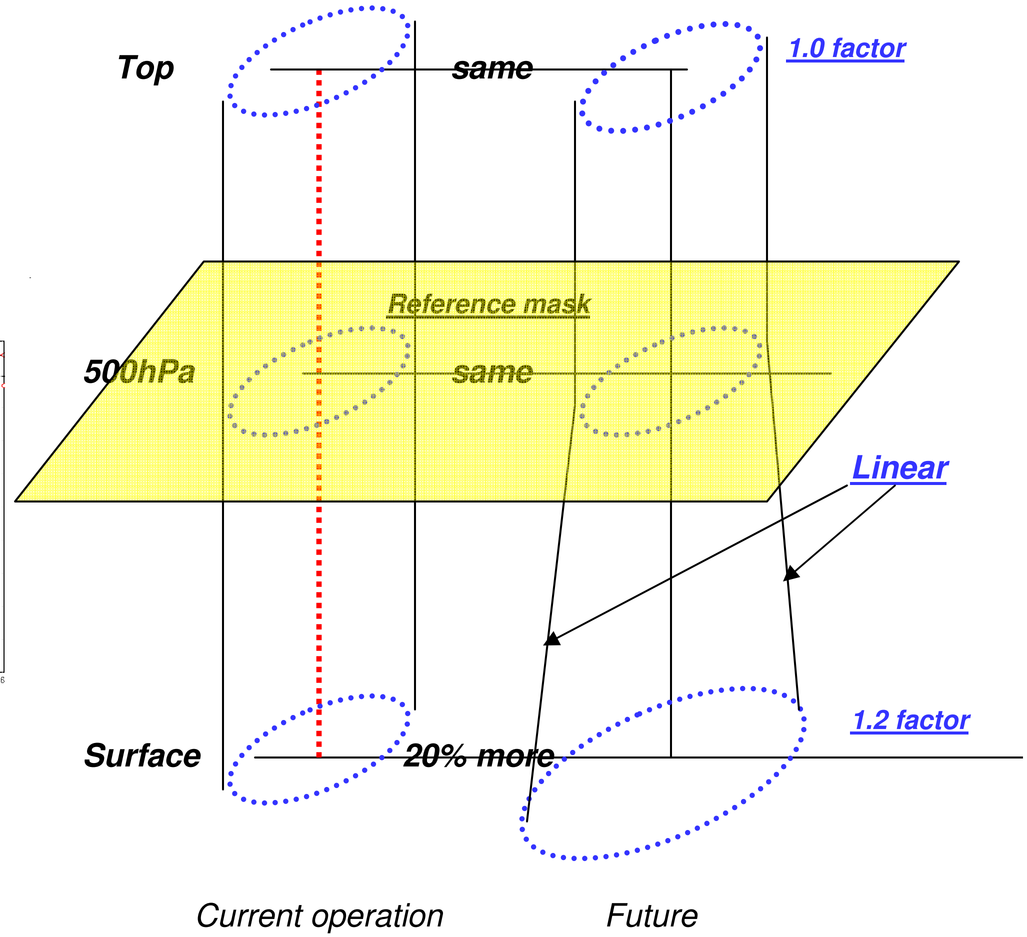
1. Wei and et al: 2006 Tellus
2. Wei and et al: 2008 Tellus

How do we tune ETR initial perturbations ?

2011



Rescaling mask and factors

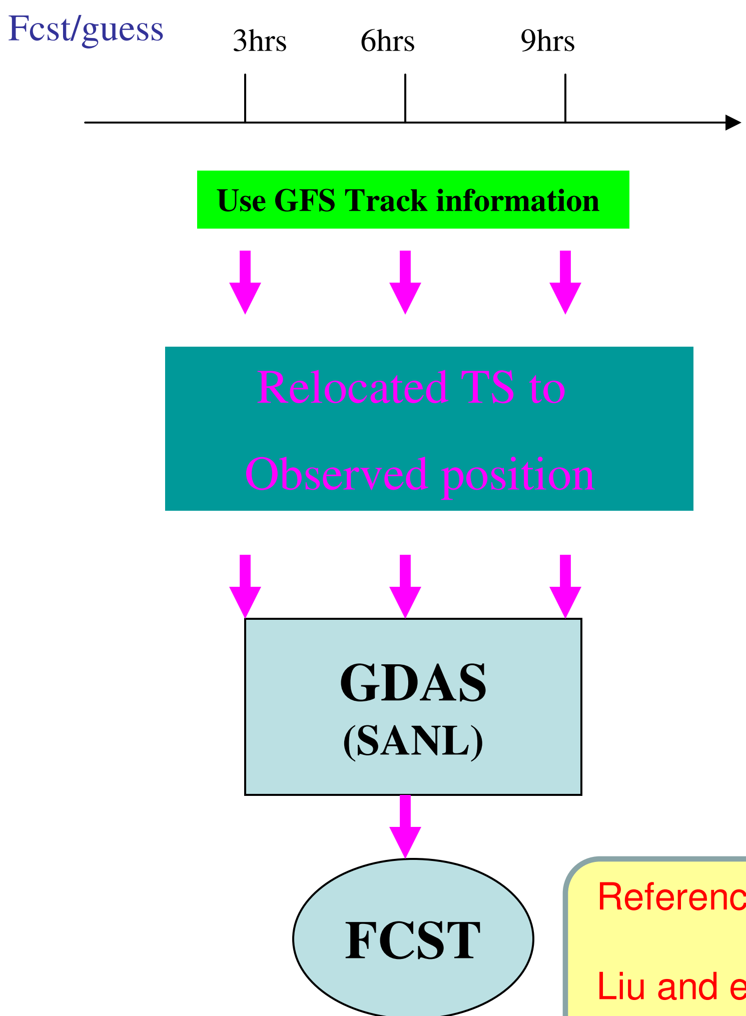


Schematic of tuning initial perturbations

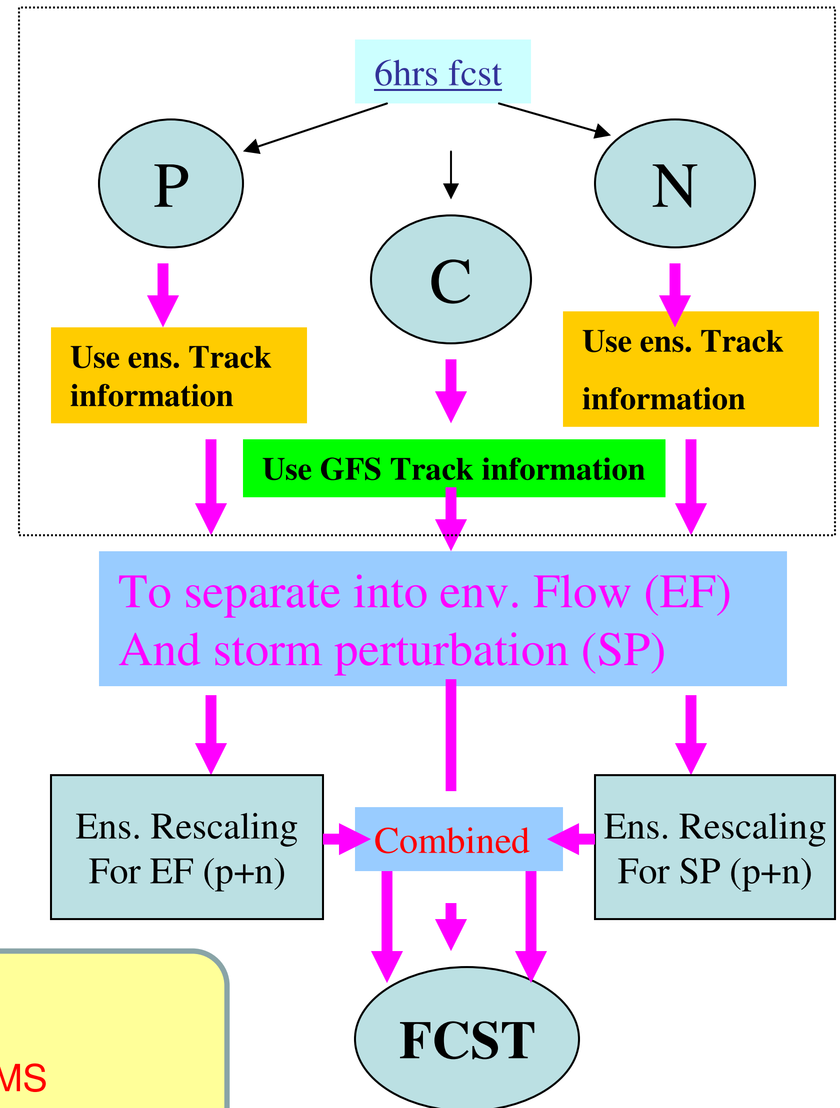
Ensemble tropical storm relocation

2005

GFS TS relocation



Ensemble TS relocation

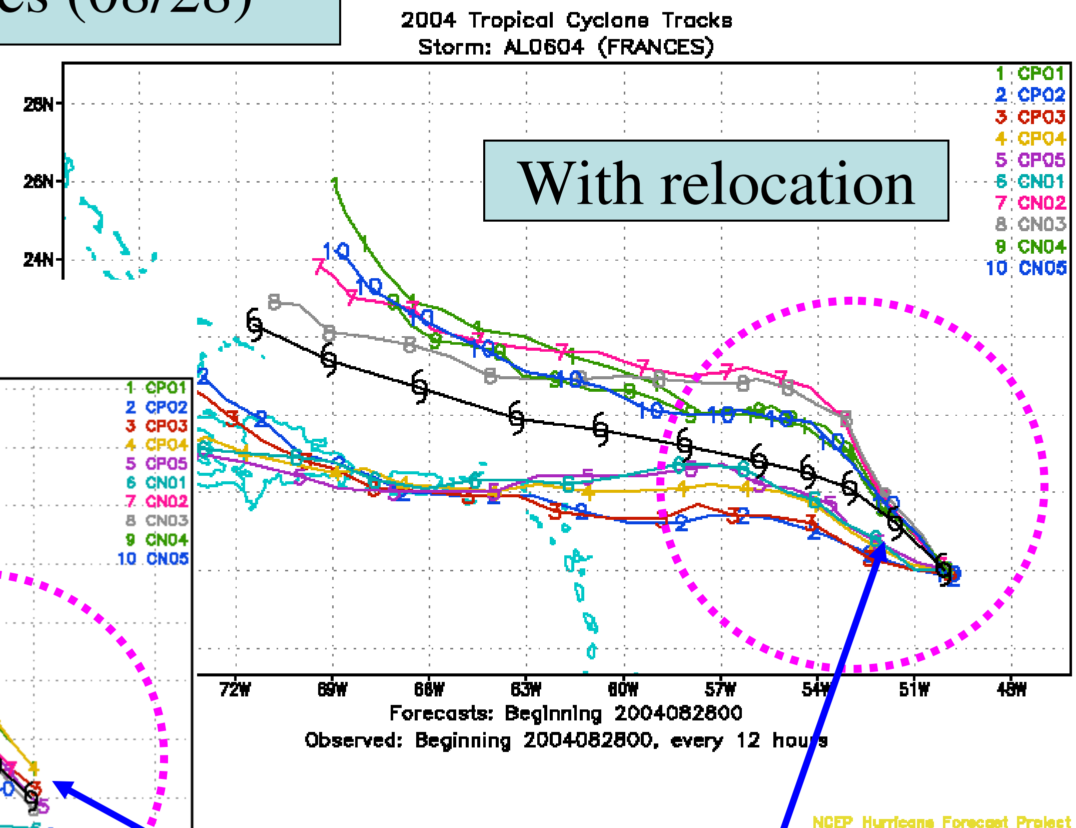


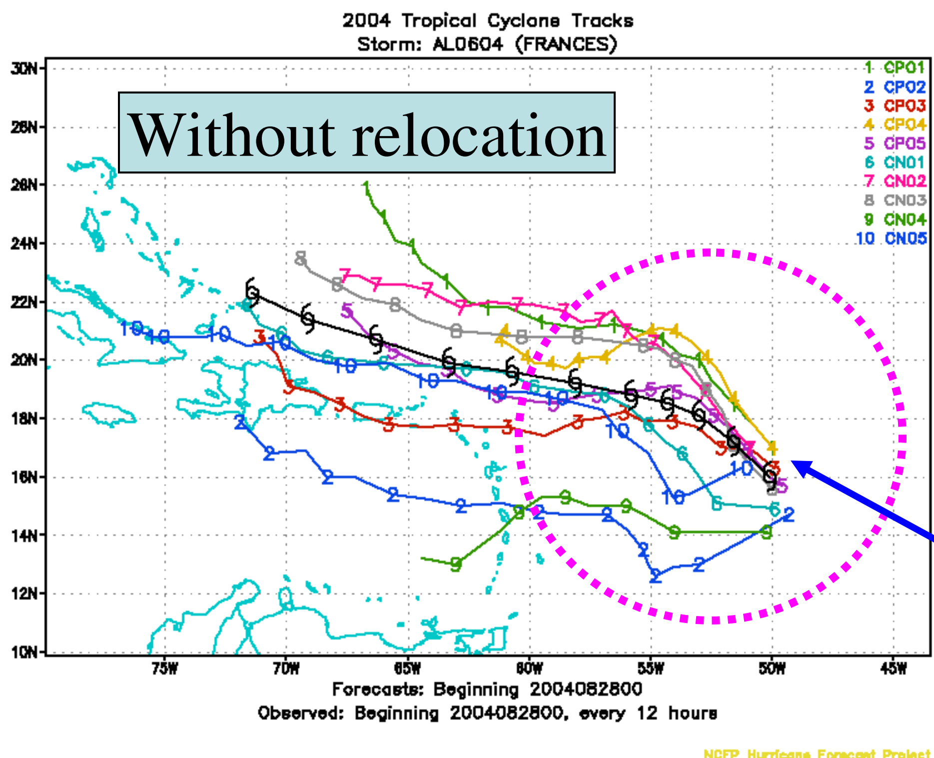
Reference:

Liu and et al: 2006 AMS conference extended paper

Hurricane Track Plots (case 1)

Frances (08/28)



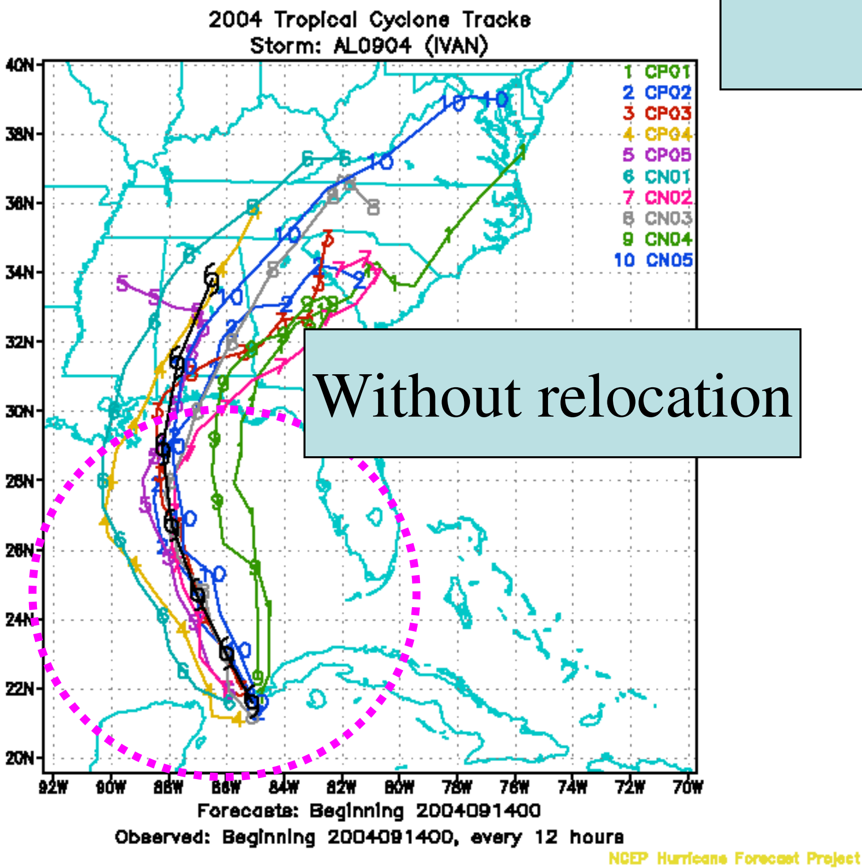


Reduced initial spread

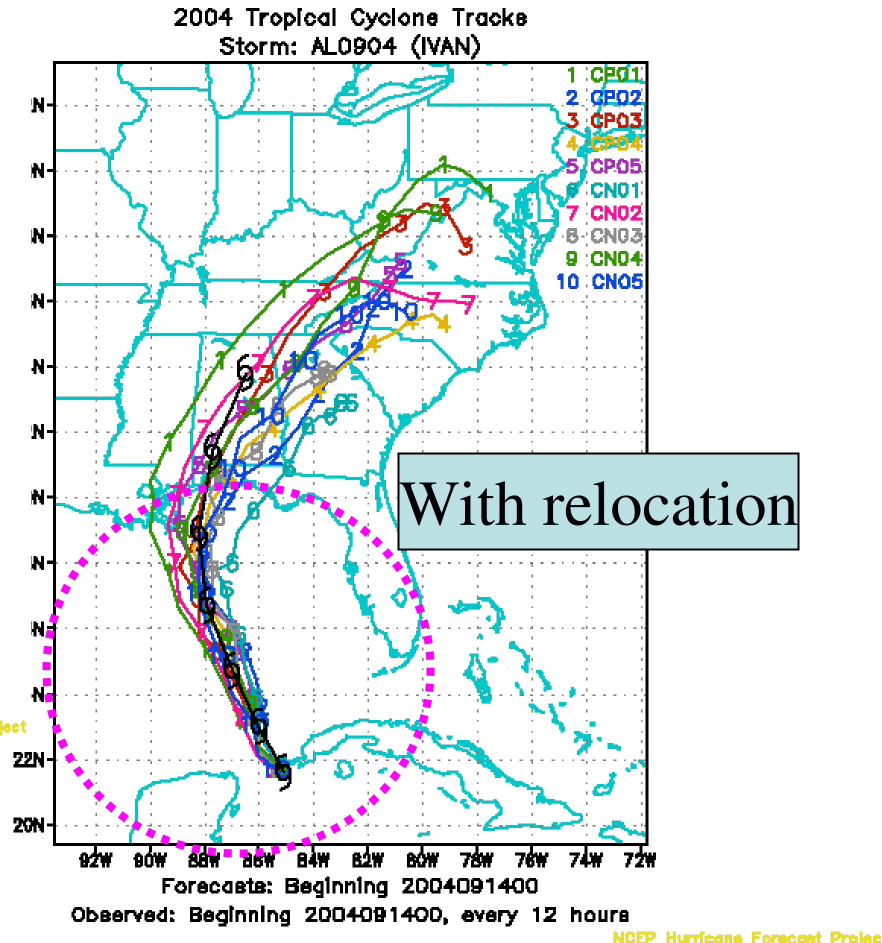
Large initial spread

Hurricane Tracks Plots (case 2)

Ivan (09/14)



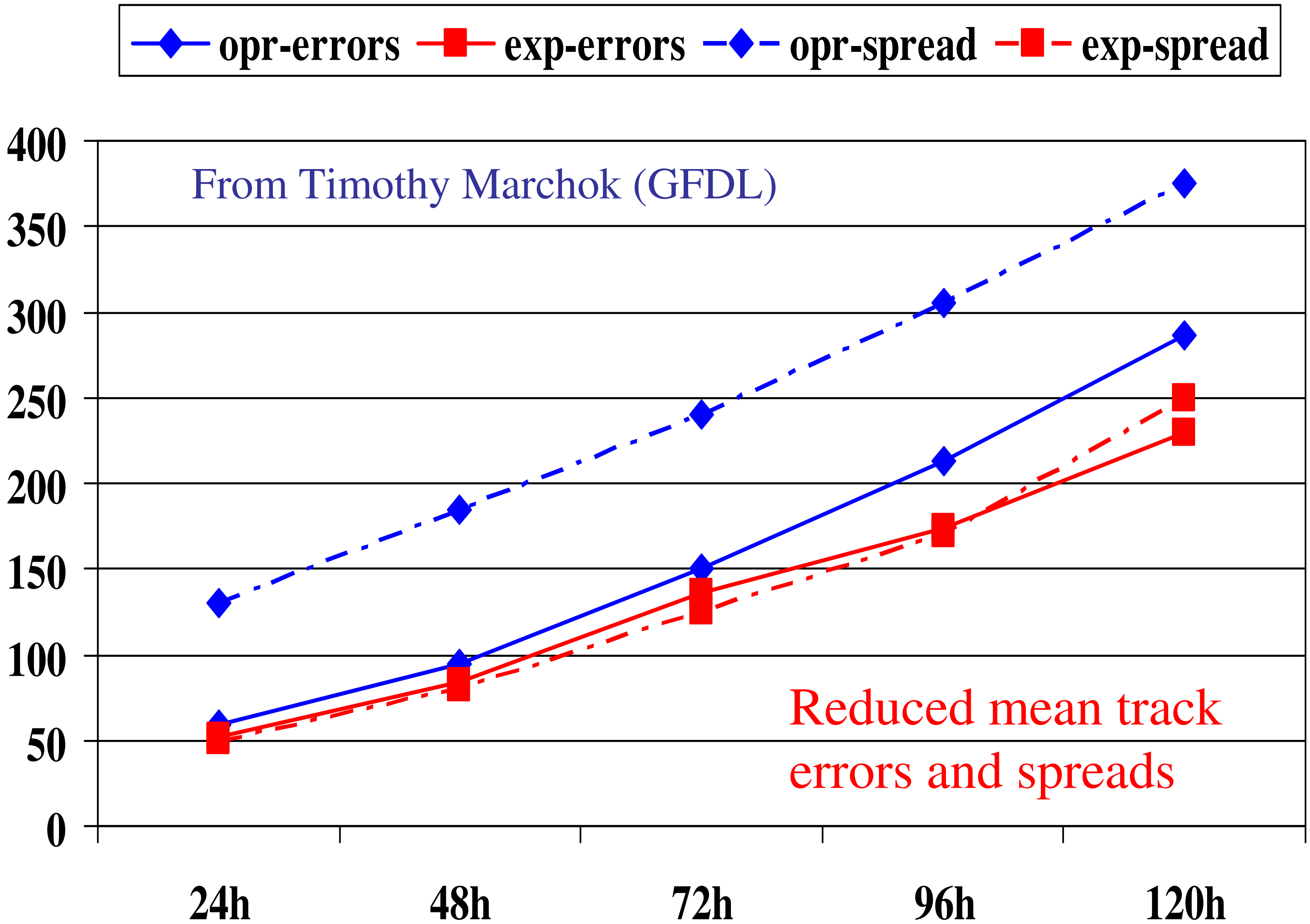
Without relocation



With relocation

Track errors and spreads

2004 Atlantic Basin (8/23-10/1)



Ensemble Stochastic Total Tendency Perturbation (STTP) Scheme

Ensemble Stochastic Total Tendency Perturbation (STTP) Scheme (*Hou, Toth and Zhu, 2006*)

NCEP operation – Feb. 2010

Formulation:

$$\frac{\partial X_i}{\partial t} = T_i(X_i;t) + \gamma \sum_{j=1,\ldots,N} w_{i,j} T_j(X_j;t)$$

Simplification: Use finite difference form for the stochastic term

Modify the model state every 6 hours:

$$X_i^{'} = X_i + \gamma \sum_{j=1}^{N} w_{i,j}(t) \left\{ \left[\left(X_j\right)_t - \left(X_j\right)_{t-6h} \right] - \left[\left(X_0\right)_t - \left(X_0\right)_{t-6h} \right] \right\}$$

Where w is an evolving combination matrix, and γ is a rescaling factor.

Reference:

1. Hou and et al: 2008 AMS conference extended paper
2. Hou and et al: 2010 in review of Tellus

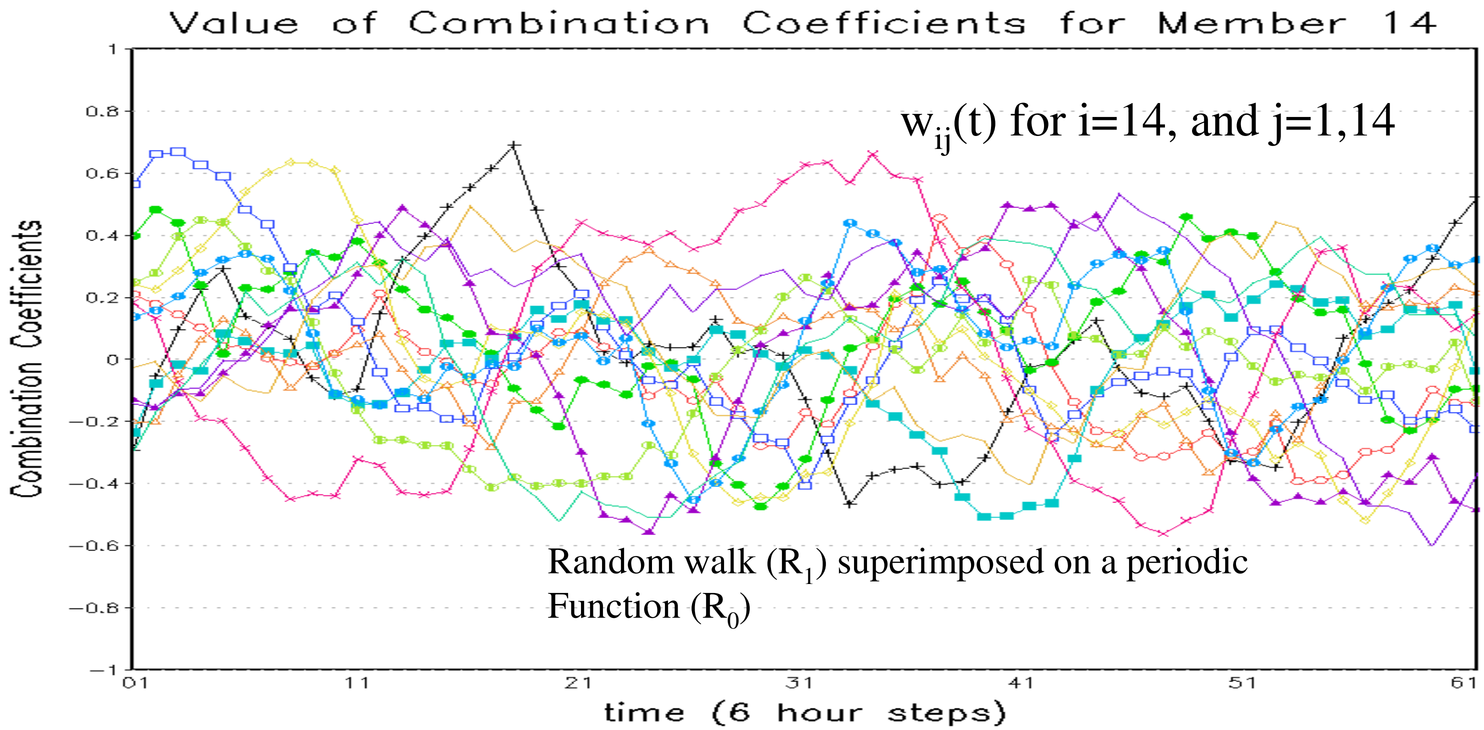
Stochastic Total Tendency Perturbation (STTP) Scheme Application

Generation of Stochastic combination coefficients:

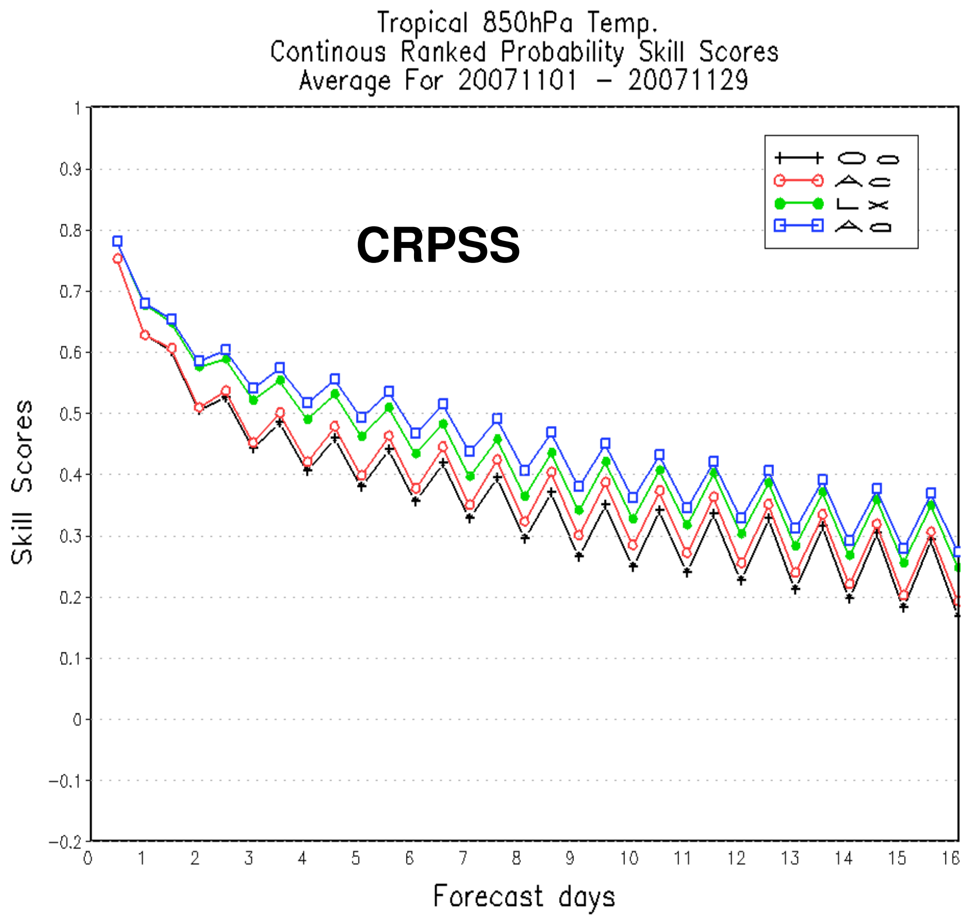
- **Matrix Notation** (N forecasts at M points)

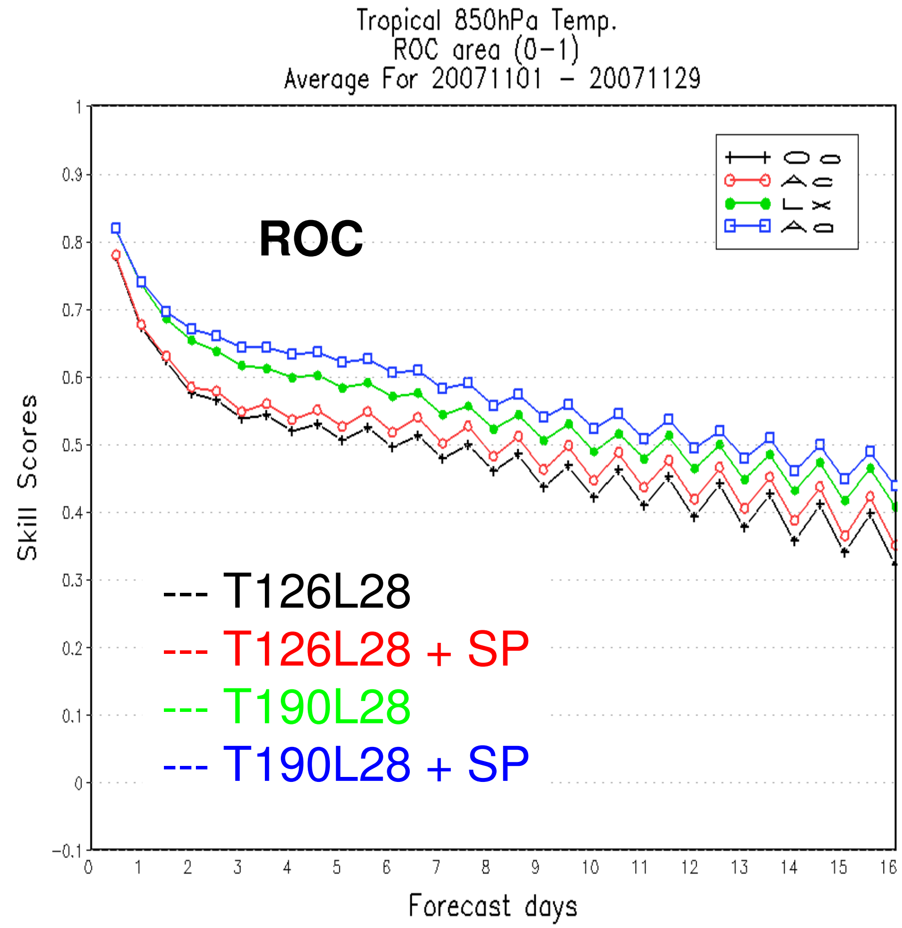
  S (t) = P(t) W(t)

  MxN MxN NxN
- 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



Experiments for 2009 Operational Implementation
T126L28 vs. T190L28 resolution, Nov. 2007 Cases
SPS works with both resolutions



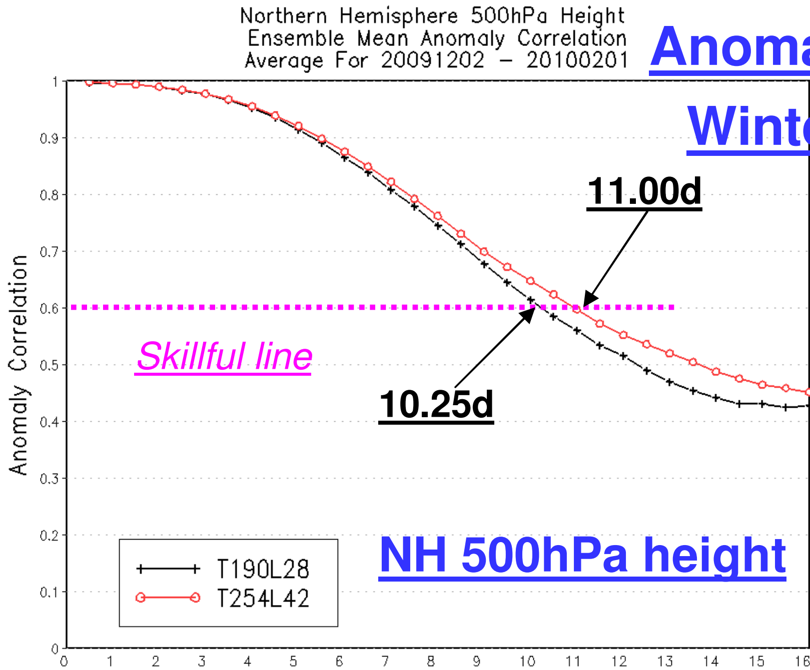


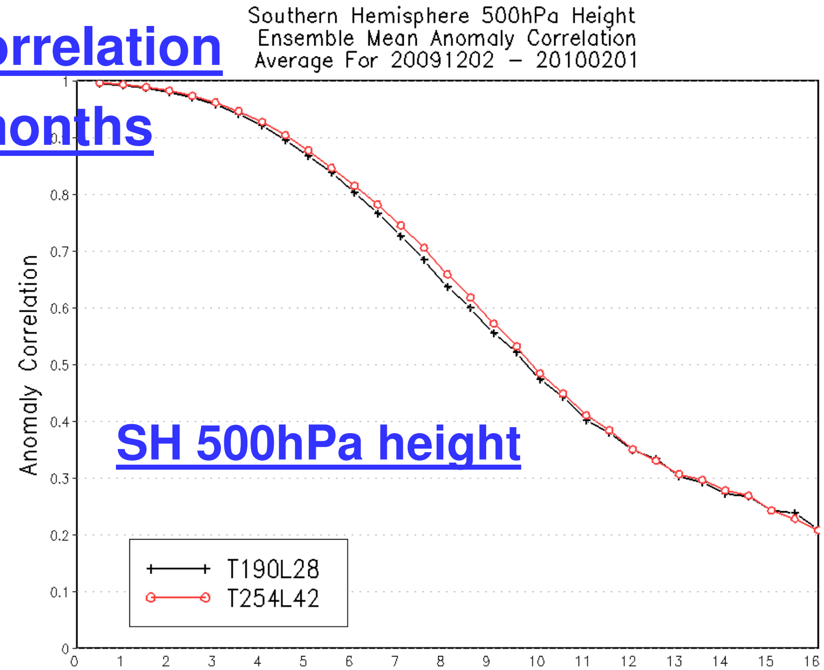
Next GEFS implementation (Q4FY2011)

- Model and initialization
  - Using GFS V9.01 instead of GFS V8.00
  - Improved Ensemble Transform with Rescaling (ETR) initialization
  - Improved Stochastic Total Tendency Perturbation (STTP)
- Configurations
  - T254 (55km) horizontal resolution for 0-192 hours (from T190 – 70km)
  - T190 (70km horizontal resolution for 192-384 hours (same as current opr)
  - L42 vertical levels for 0-384 hours (from L28)
- Part of products will be delayed by approximately 20 minutes
  - Due to limit CCS resources
  - 40 nodes for 70 minutes (start +4:35 end: +5:45)
- Unchanged:
  - 20+1 members per cycle, 4 cycles per day
  - pgrb file output at 1*1 degree every 6 hours
  - GEFS and NAEFS post process output data format
- Why do we make this configurations?
  - Considering the limited resources
  - Resolution makes difference (example of T126 .vs T190)
- What do we expect from this implementation?
  - Preliminary results (NH 500hPa and SH 500hPa height and tracks)

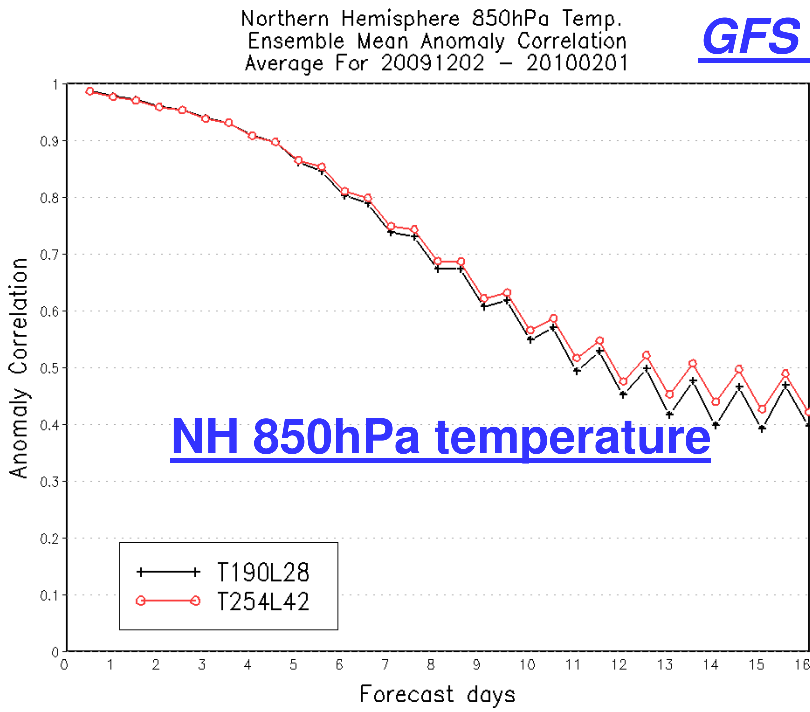
Anomaly Correlation

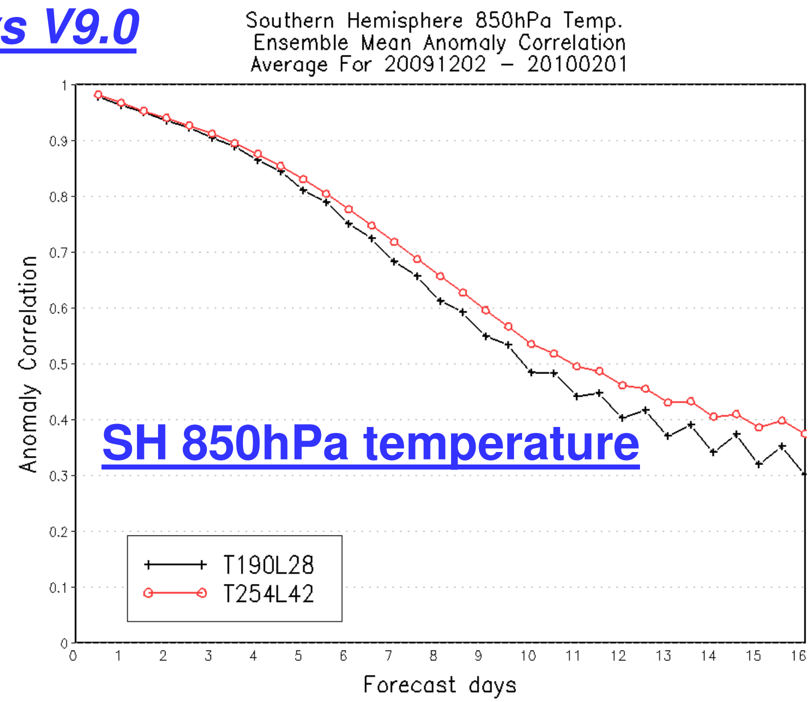
Winter 2 months





GFS V8.0 .vs V9.0

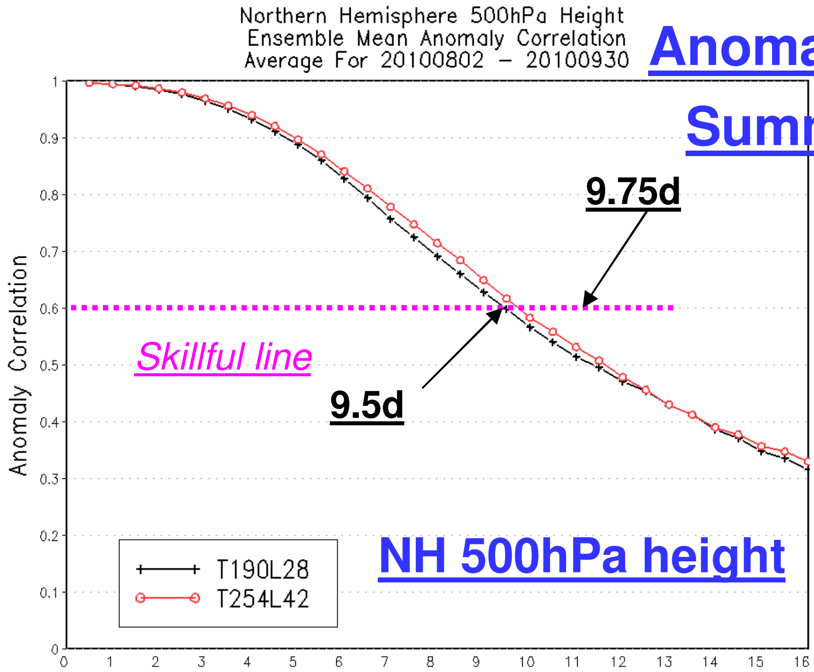


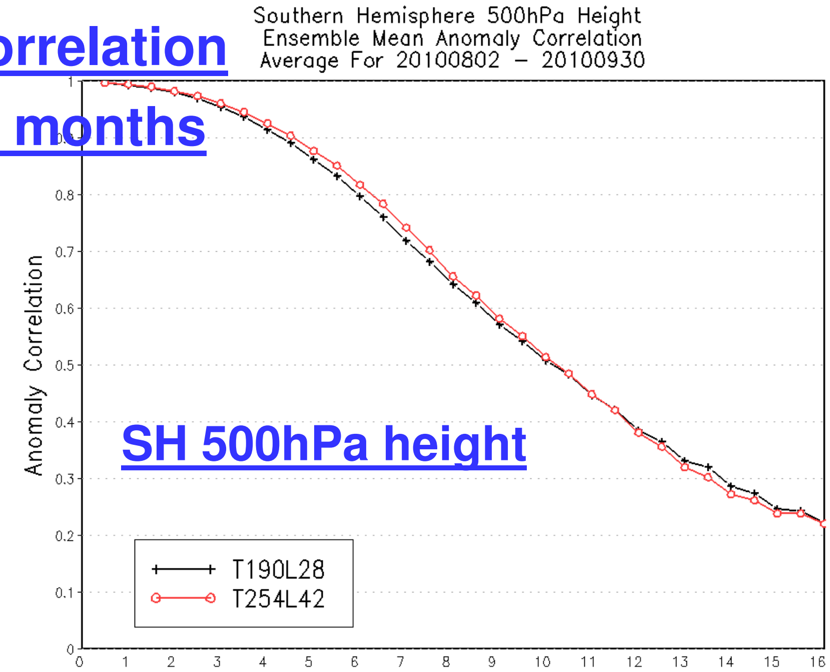


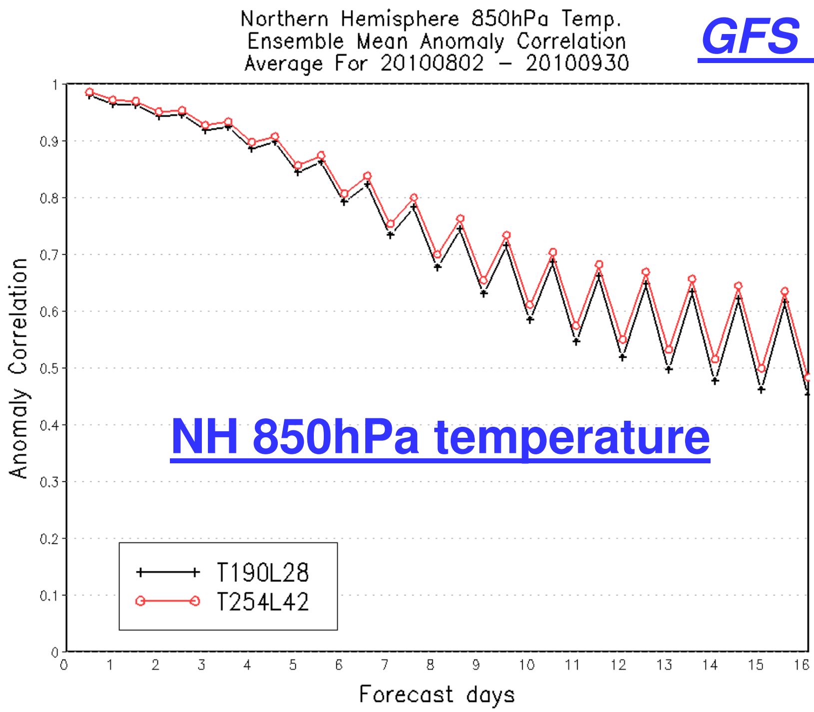
Anomaly Correlation

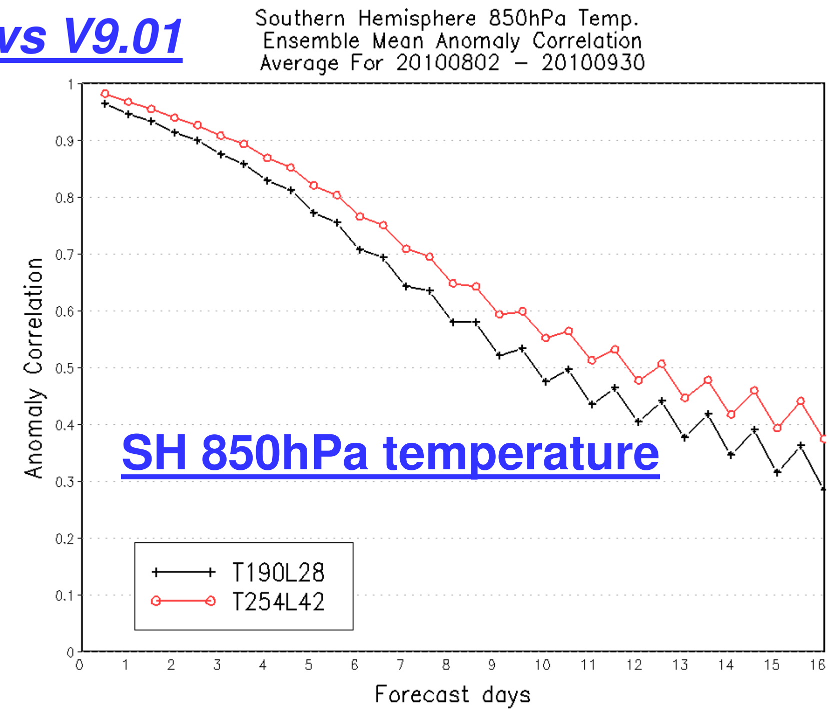
Summer 2 months

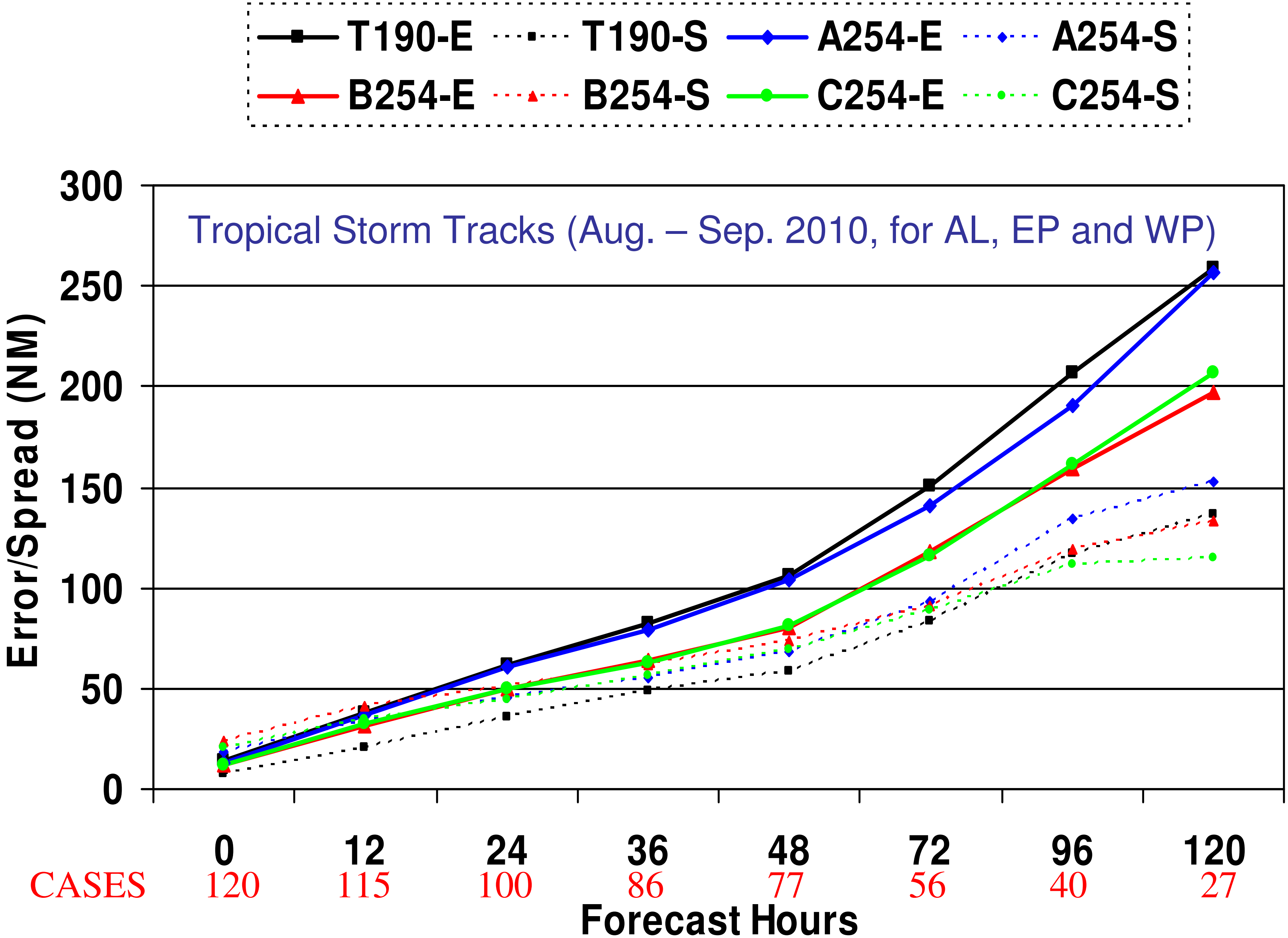
GFS V8.0 .vs V9.01









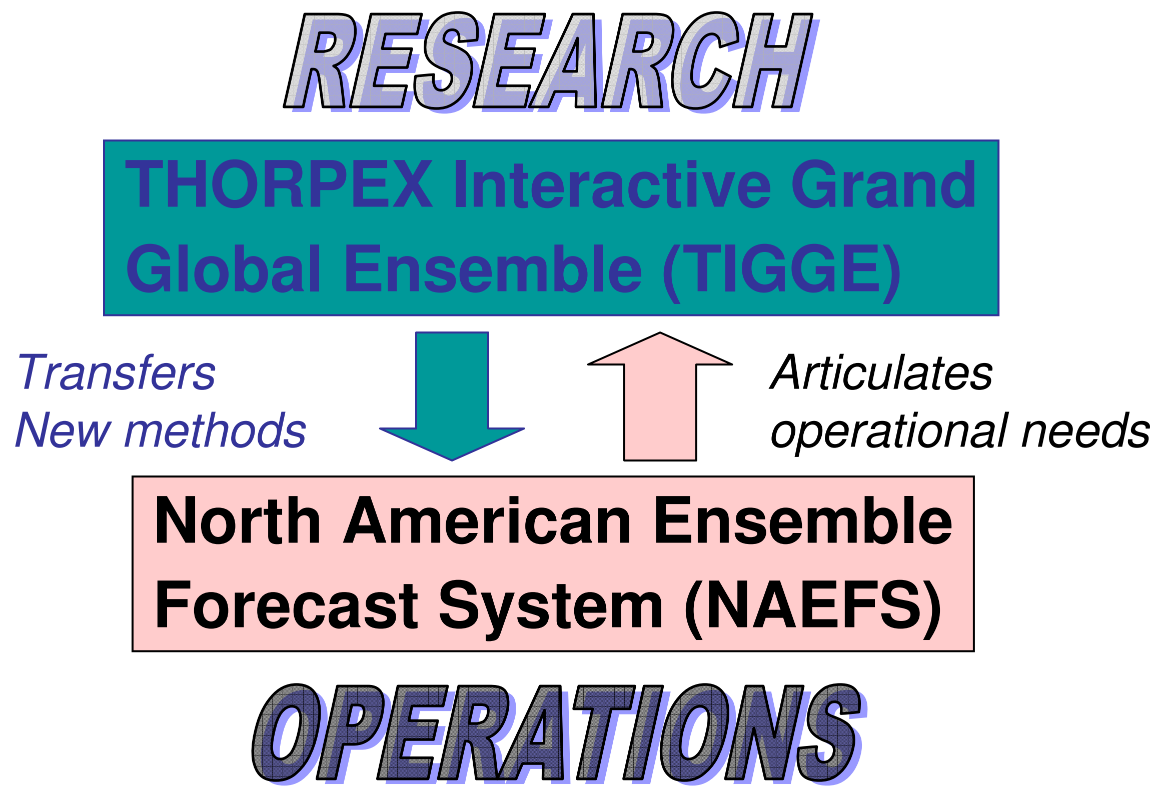

T190-E
T190-S
A254-E
A254-S
B254-E
B254-S
C254-E
C254-S
Tropical Storm Tracks (Aug. – Sep. 2010, for AL, EP and WP)
Error/Spread (NM)
300
250
200
150
100
50
0
0
12
24
36
48
72
96
120
CASES
120
115
100
86
77
56
40
27
Forecast Hours

NAEFS and post process

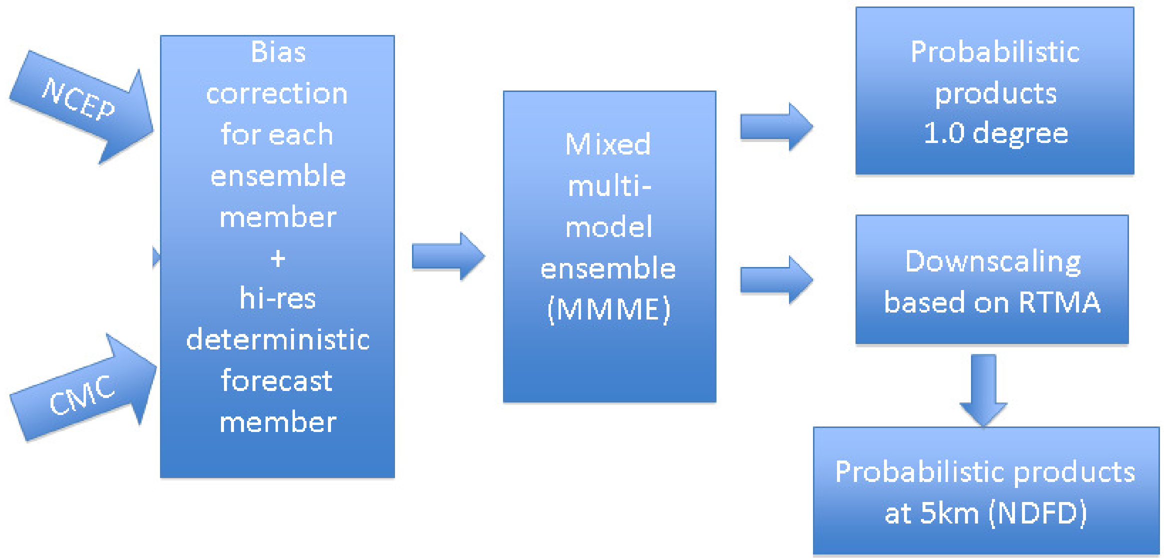
Multi-model ensembles

NAEFS & THORPEX

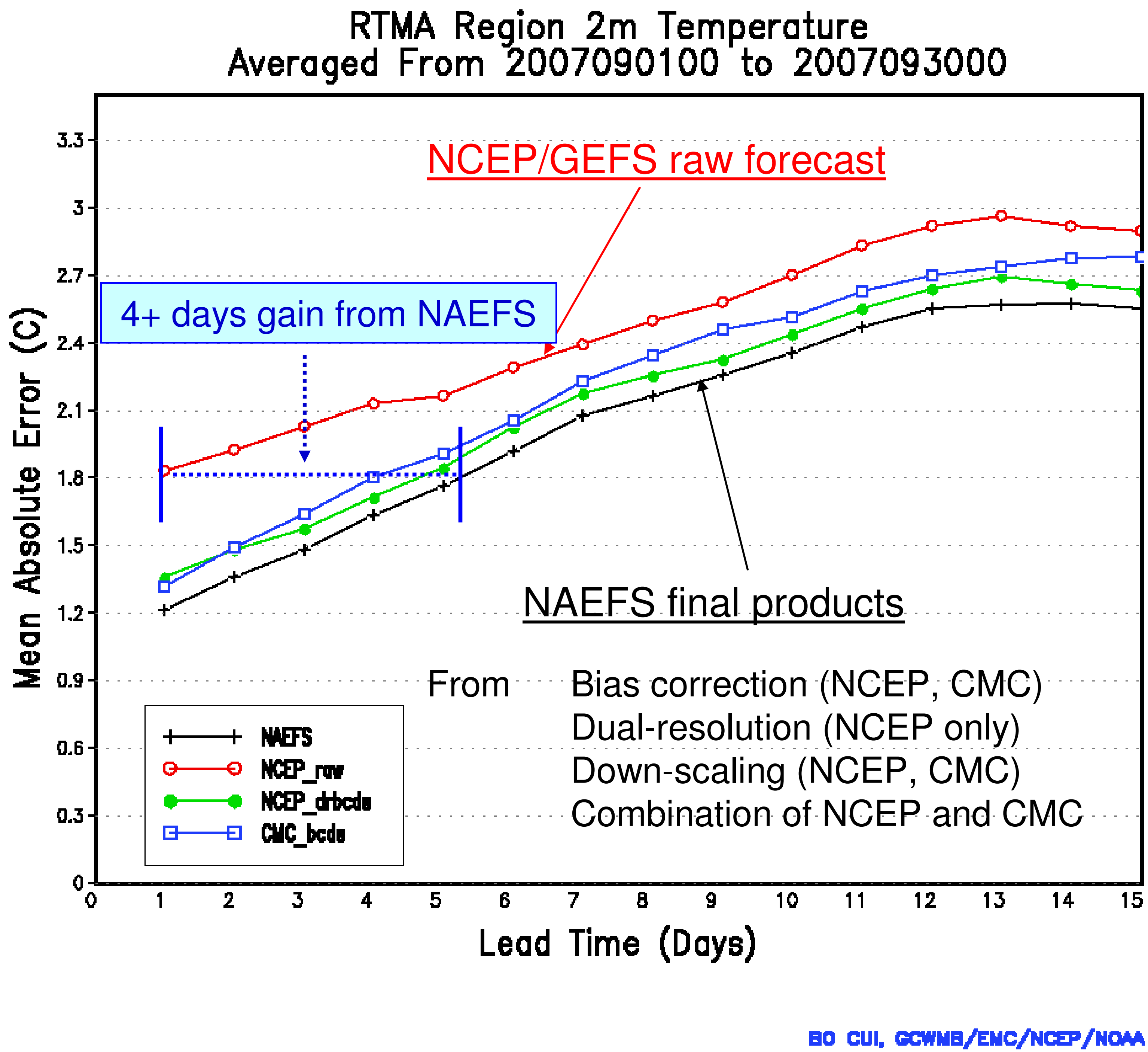
- Expands international collaboration
  - Mexico joined in November 2004
  - FNMOC to join in 2009
  - UK Met Office may join in 2009
- Provides framework for transitioning research into operations
  - Prototype for ensemble component of THORPEX legacy forecast system: ***Global Interactive Forecast System (GIFS)***

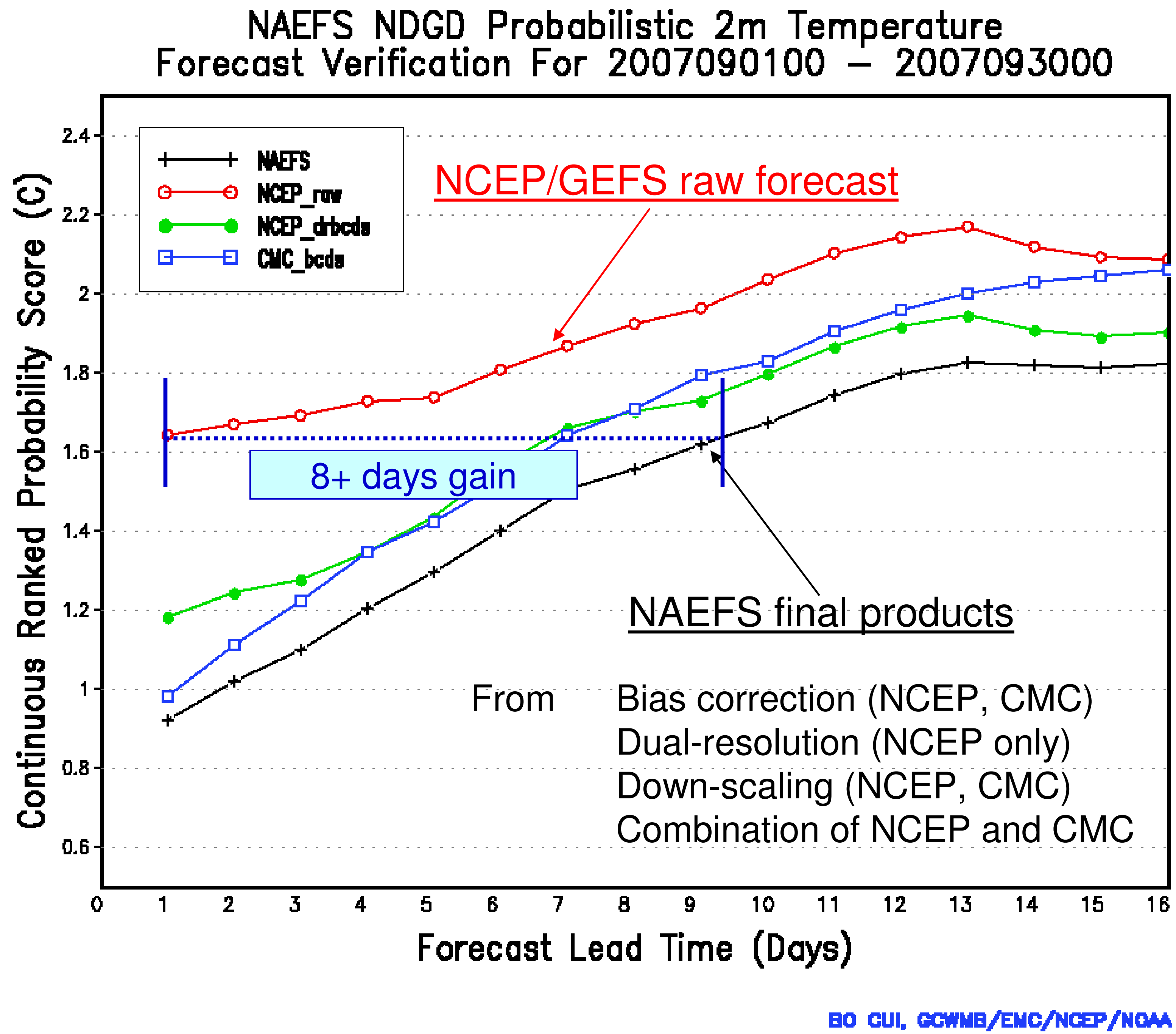


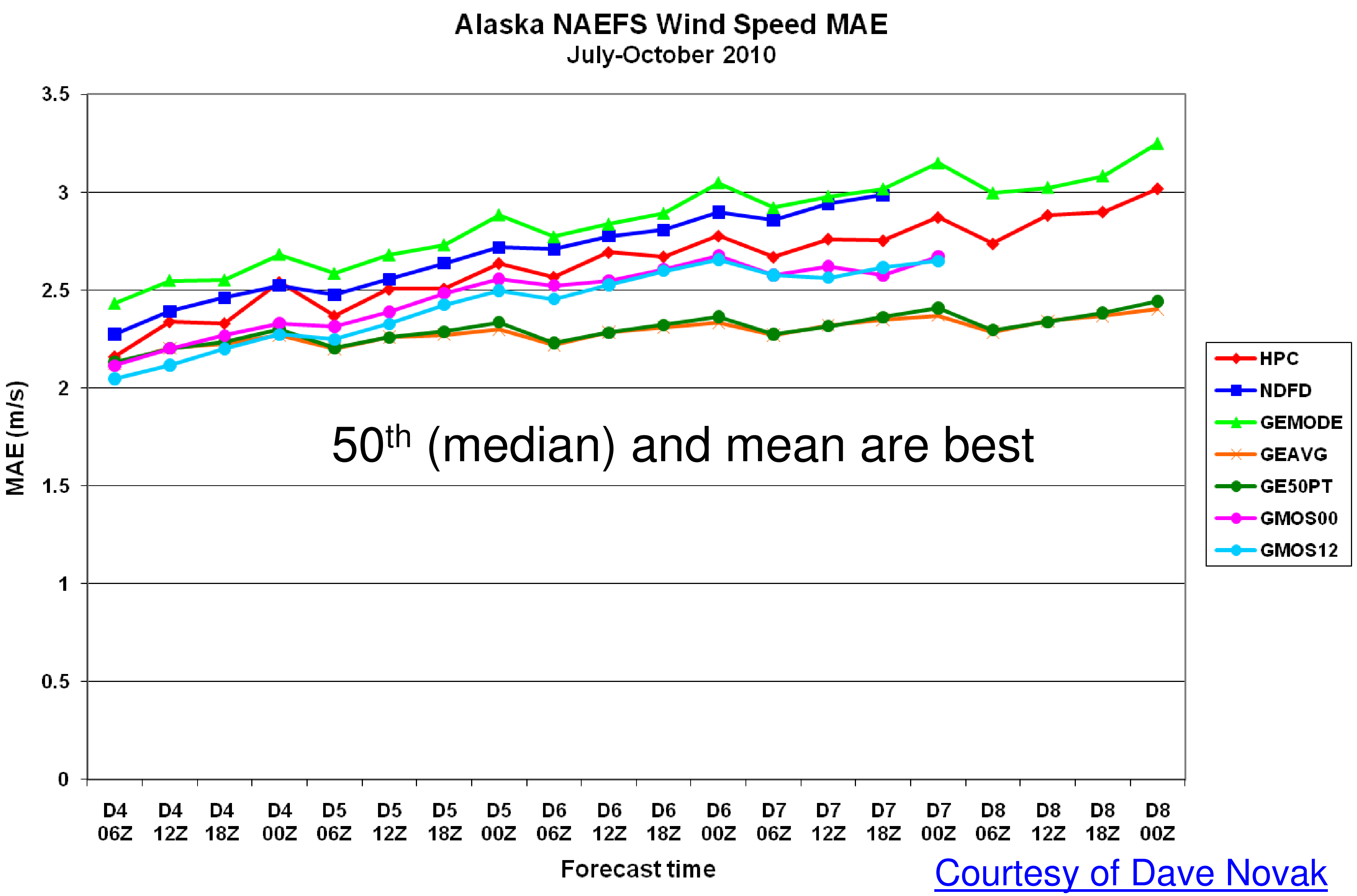
Current NCEP/EMC Statistical Post-Processing System



- Bias corrected NCEP/CMC GEFS and GFS forecast (up to 180 hrs), same bias correction algorithm
  - Combine bias corrected GFS and NCEP GEFS ensemble forecasts
  - Dual resolution ensemble approach for short lead time
  - GFS has higher weights at short lead time
- NAEFS products
  - Combine NCEP/GEFS (20m) and CMC/GEFS (20m), FNMOC ens. will be in soon
  - Produce Ensemble mean, spread, mode, 10% 50%(median) and 90% probability forecast at 1*1 degree resolution
  - Climate anomaly (percentile) forecasts also generated for ens. mean
- Statistical downscaling
  - Use RTMA as reference - NDGD resolution (5km), CONUS only
  - Generate mean, mode, 10%, 50%(median) and 90% probability forecasts

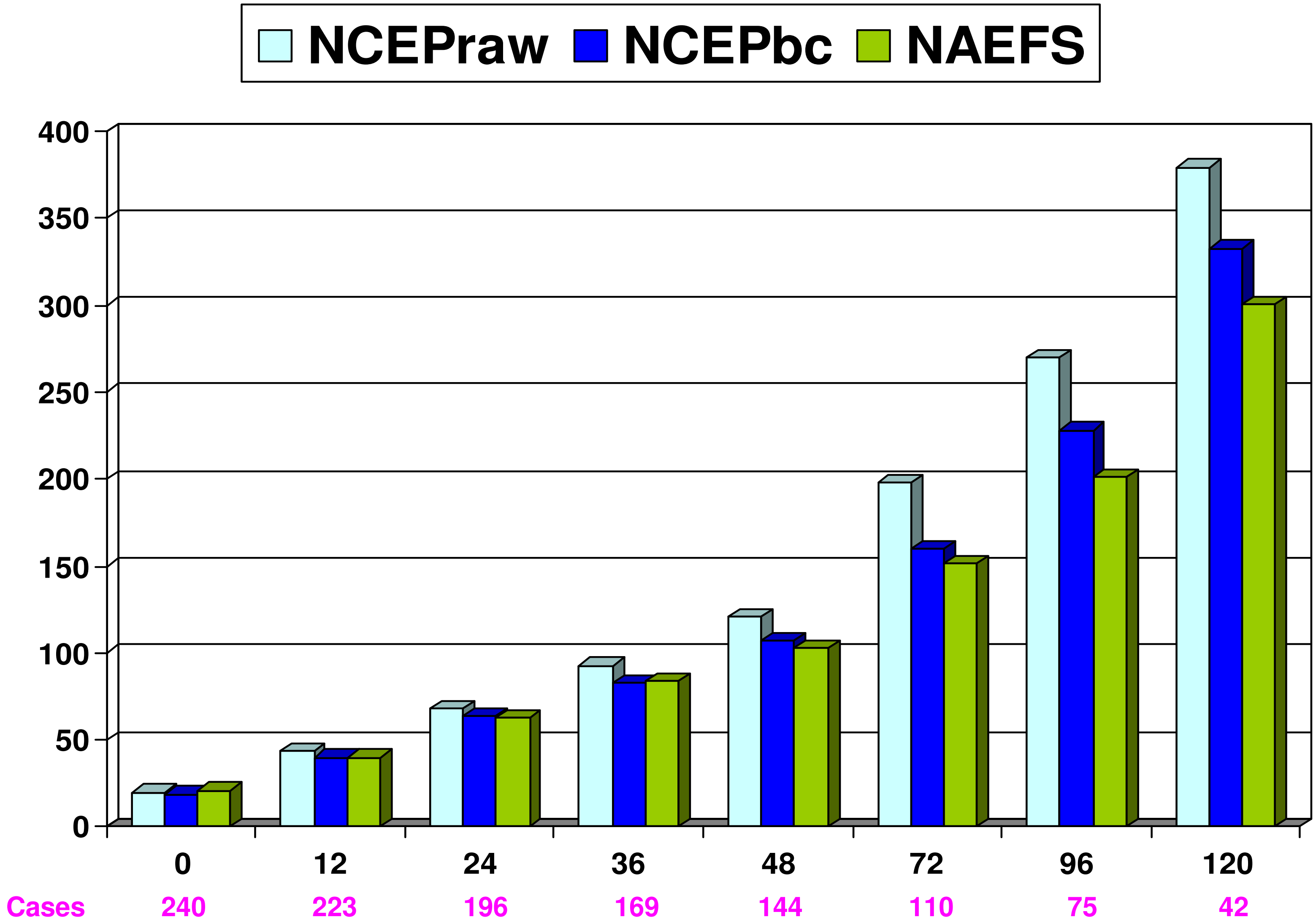

RTMA Region 2m Temperature
Averaged From 2007090100 to 2007093000
NCEP/GEFS raw forecast
4+ days gain from NAEFS
NAEFS final products
From
Bias correction (NCEP, CMC)
Dual-resolution (NCEP only)
Down-scaling (NCEP, CMC)
Combination of NCEP and CMC
NAEFS
NCEP_raw
NCEP_drbcds
CMC_bcds
Mean Absolute Error (C)
Lead Time (Days)
0
0.3
0.6
0.9
1.2
1.5
1.8
2.1
2.4
2.7
3
3.3
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
BO CUI, GCWMB/EMC/NCEP/NOAA


NAEFS NDGD Probabilistic 2m Temperature
Forecast Verification For 2007090100 – 2007093000
NAEFS
NCEP_raw
NCEP_drbcds
CMC_bcds
NCEP/GEFS raw forecast
8+ days gain
NAEFS final products
From
Bias correction (NCEP, CMC)
Dual-resolution (NCEP only)
Down-scaling (NCEP, CMC)
Combination of NCEP and CMC
Continuous Ranked Probability Score (C)
Forecast Lead Time (Days)
BO CUI, GCWMB/EMC/NCEP/NOAA



Courtesy of Dave Novak

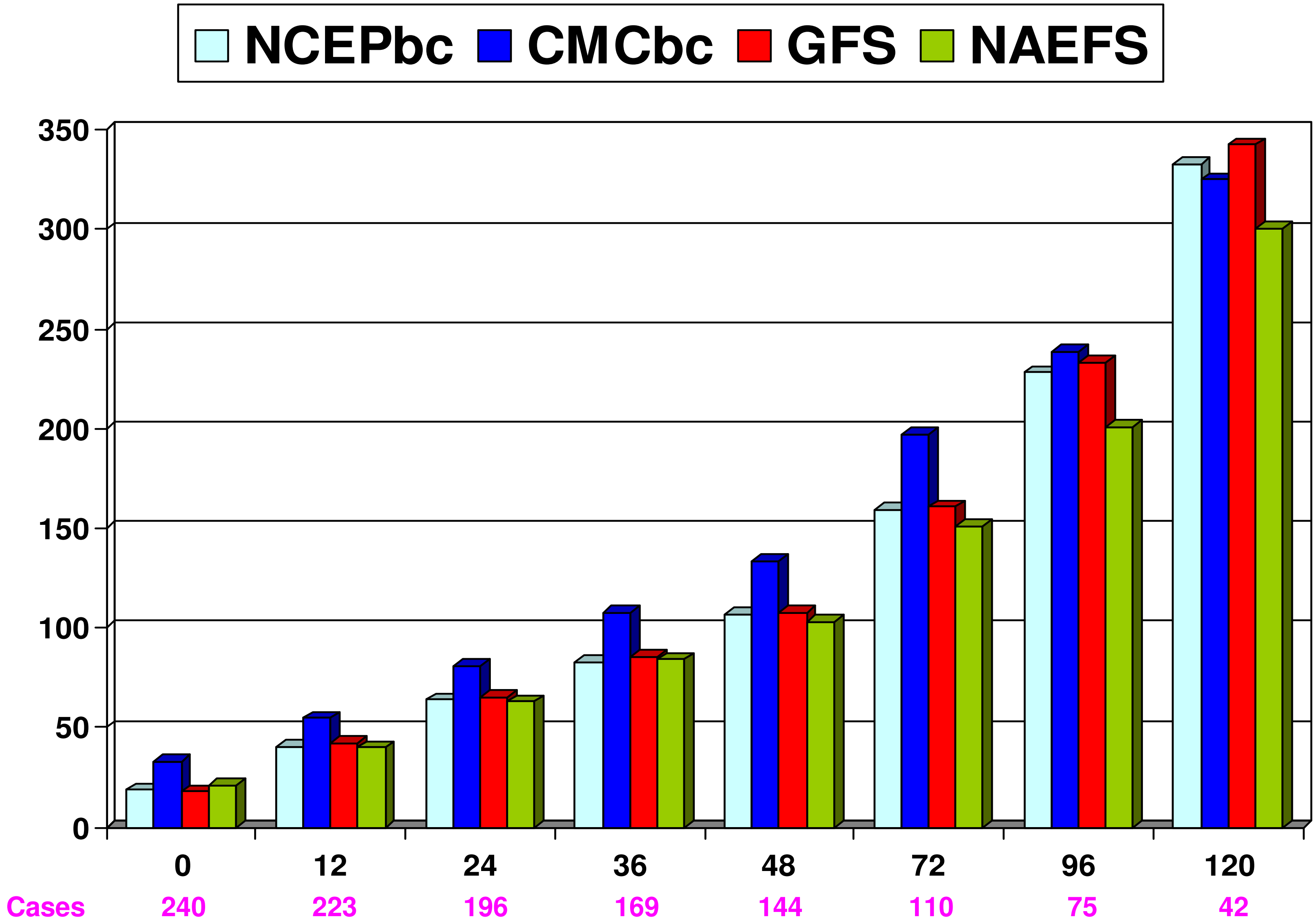
Track forecast error for 2009 season (AL+EP+WP)



NAEFS is combined NCEP (NCEPbc) and CMC's (CMCbc) bias corrected ensemble and bias corrected GFS

Contributed by Dr. Jiayi Peng (EMC/NCEP)

Track forecast error for 2009 season (AL+EP+WP)

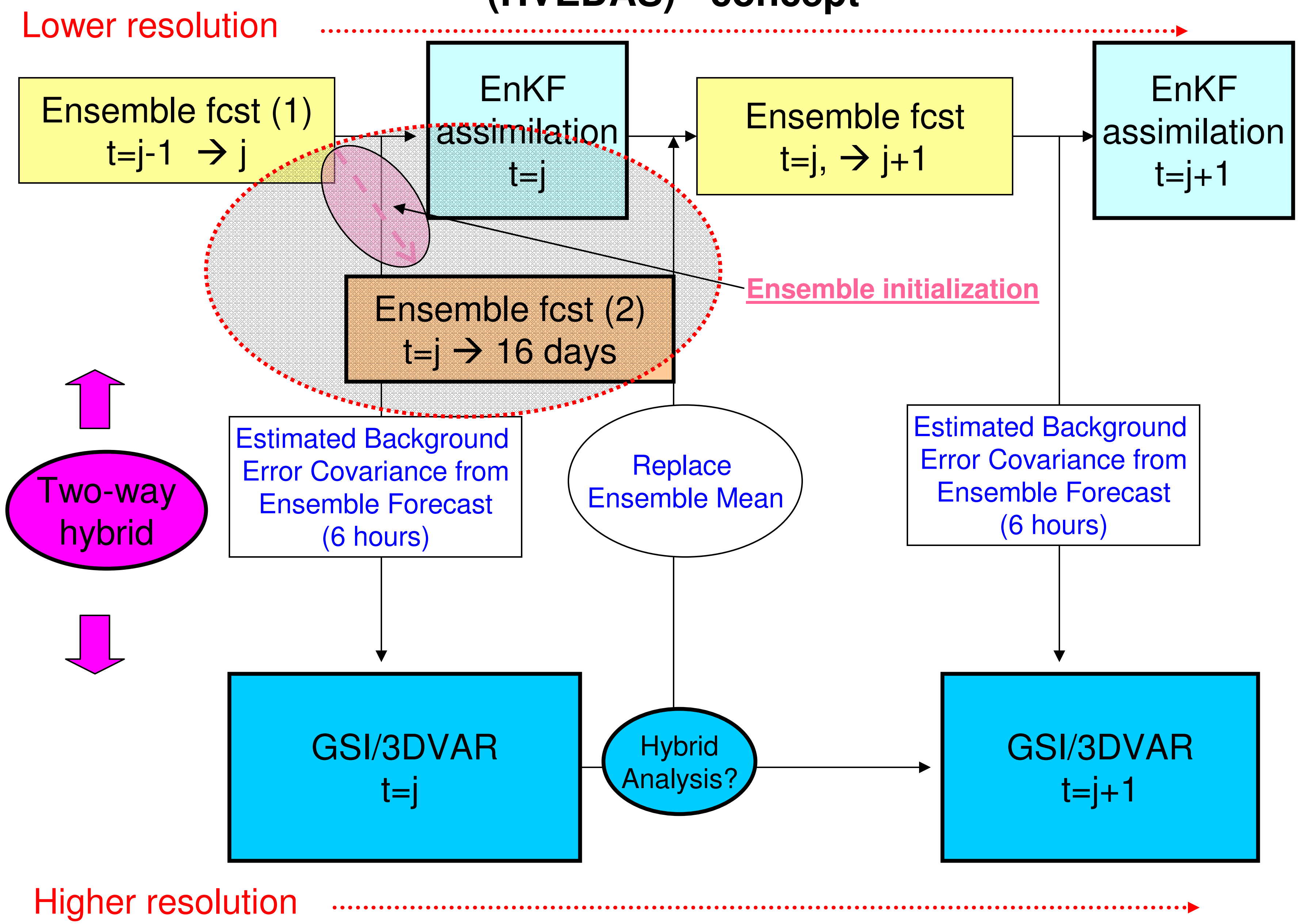


NAEFS is combined NCEP (NCEPbc) and CMC's (CMCbc) bias corrected ensemble and bias corrected GFS

Contributed by Dr. Jiayi Peng (EMC/NCEP)

Future plans

Flow Chart for Hybrid Variation and Ensemble Data Assimilation System (HVEDAS) - concept



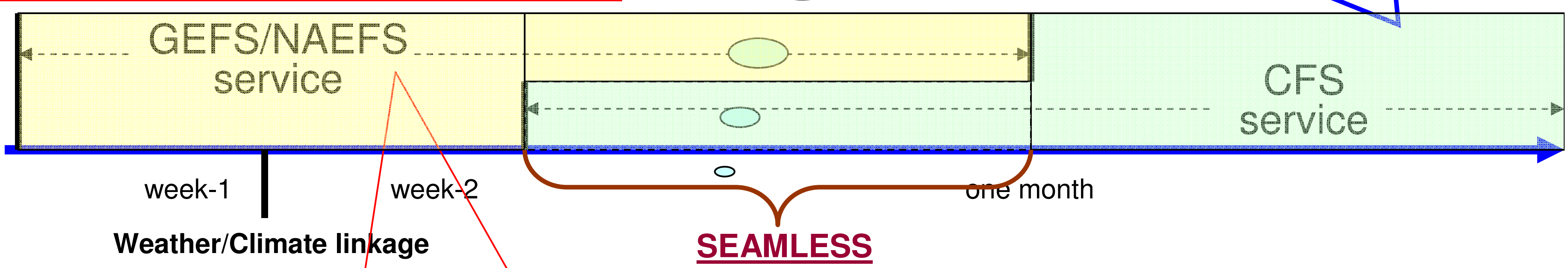
Future seamless forecast system

NCEP/GEFS will plan for T254L42 (2010 GFS version) resolution with tuned ETR initial perturbations and adjusted STTP scheme for 21 ensemble members, forecast out to 16 days and 4 cycles per day. **Extended to 45 days at T126L28/42 resolution, 00UTC only (coupling is still a issue?)**
NAEFS will include FNMOC ensemble in 2011, with improving post process which include bias correction, dual resolution and down scaling

Main event
MJO

Main products:

ENSO predictions???
Seasonal forecast???



Main products:

1. Probabilistic forecasts for every 6-hr out to 16 days, 4 times per day: 10%, 50%, 90%, ensemble mean, mode and spread.
2. D6-10, week-2 temperature and precipitation probabilistic mean forecasts for above, below normal and normal forecast
3. **MJO forecast (week 3 & 4 …)**

Next Operational CFS will plan to be implemented by Q2FY2011 with T126L64 atmospheric model resolution (CFSv2, 2010version) which is fully coupled with land, ocean and atmosphere (GFS+MOM4+NOAH), 4 members per day (using CFS reanalysis as initial conditions, one day older?), integrate out to 9 months.

Future: initial perturbed CFS

ENSEMBLES AND THE RESEARCH COMMUNITY

LINKED THROUGH THORPEX – MAJOR INTERNATIONAL RESEARCH PROGRAM

GOAL: Accelerate improvements of high impact weather forecasts