WRF Variational Data Assimilation System (WRF-Var) Overview

WRF Tutorial Presentation

NCAR, Boulder, Colorado, USA

January 2007

Dale Barker (dmbarker@ucar.edu)

Acknowledge:
NCAR/MMM Division Staff
USWRP, NSF-OPP, NCAR Data Assimilation Initiative
US Air Force Weather Agency, Korean Meteorological Administration
Taiwanese Central Weather Bureau, Civil Aeronautics Administration

Outline of Talk

1) What is WRF-Var?

2) Practical Variational Data Assimilation.

3) Background Error Modeling.

4) Observational Issues.

5) Current Status and Future Plans.

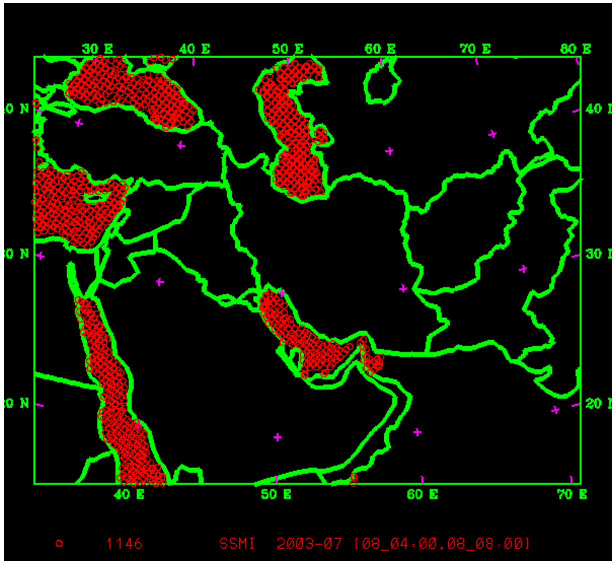
1. What is WRF-Var?

…WRF-Var is a **unified** variational data assimilation system built within the software framework of the Weather Research and Forecasting (WRF) model, used for application in both research and operational environments….

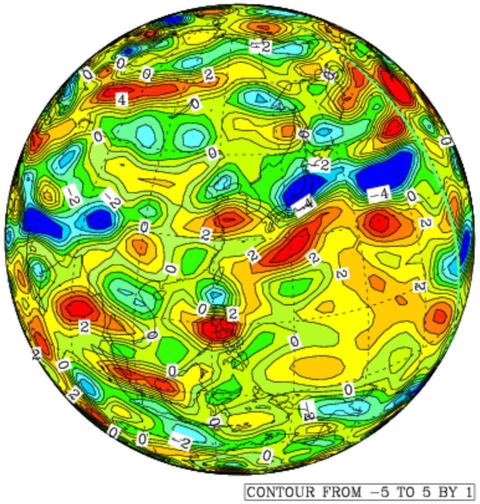
What Do We Mean By “Unified”

- Domains: Regional/global.
- Techniques: 3D-Var, 4D-Var, Hybrid Var/Ensemble DA.
- Code: Single code for research, development and release. Supported by NCAR/MMM.
- Software Engineering: WRF framework.
- Model: Runs with WRF, and also KMA global model.

AFWA 15km S-W Asia:

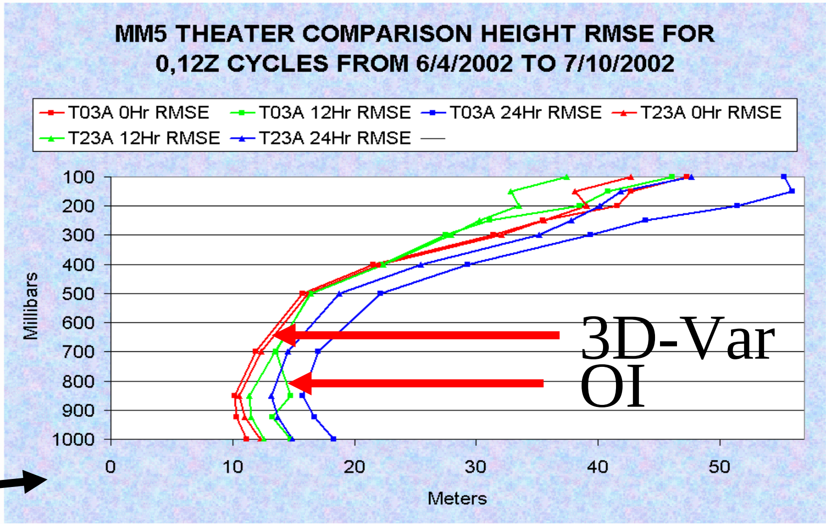


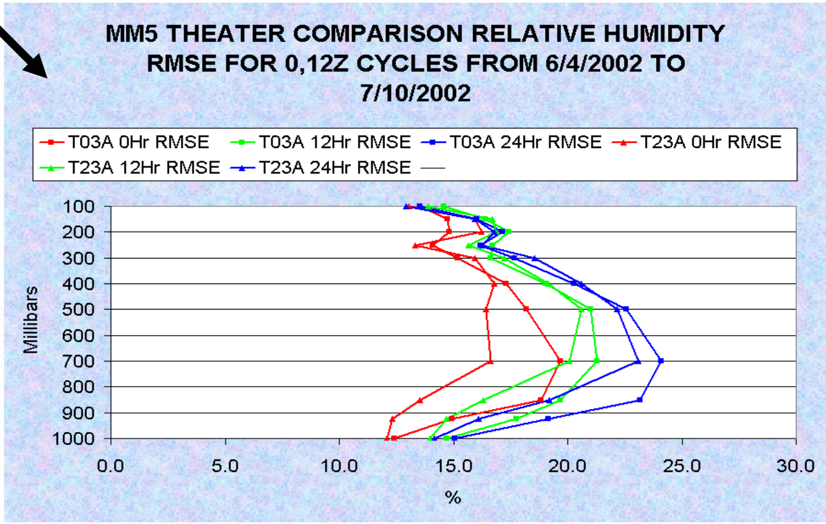
KMA T213 Global:



WRF Variational Data Assimilation (WRF-Var) History

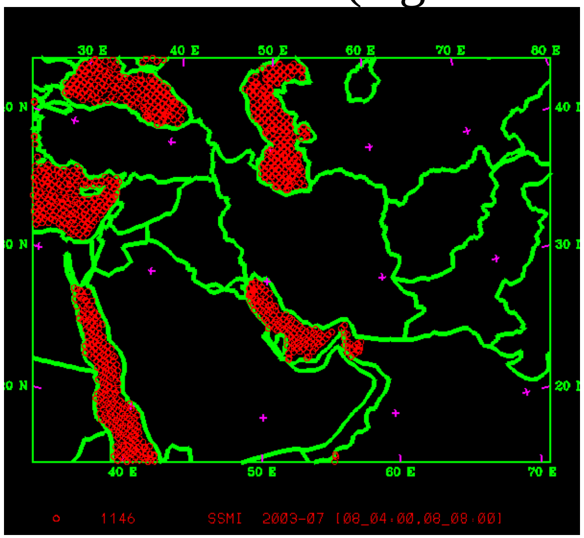
- **June 2001:** MM5-3DVar adopted as starting point for WRF 3D-Var.
- **May 2002:** MM5/WRF 3D-Var operational at Taiwanese CAA.
- **September 2002:** MM5/WRF 3D-Var operational in 45km domains at AFWA.
- **June 2003:** WRF 3D-Var V1.0 release.
- **May 2004:** WRF 3D-Var V2.0 release.
- **July 2005:** WRF-Var V2.1 release.
- **February 2007:** WRF-Var V2.2 release.



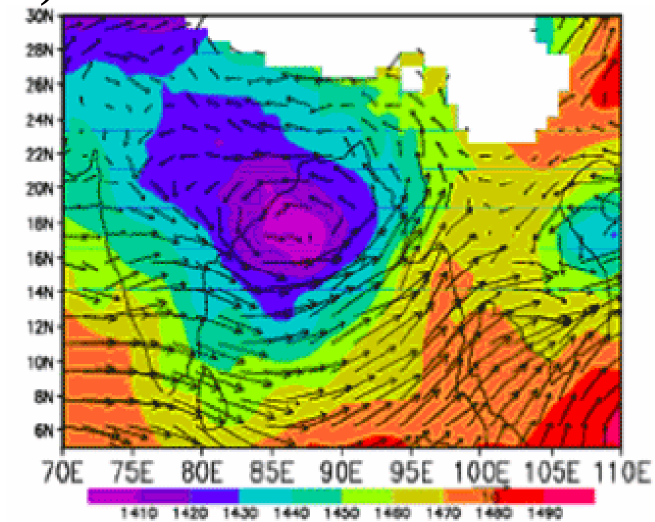


WRF-Var Operational Applications: June 2005

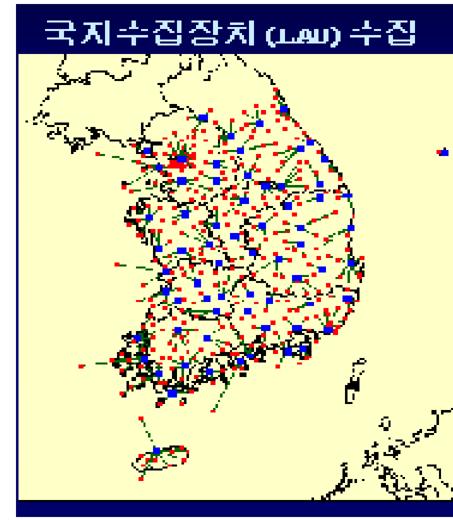
AFWA 15km (e.g. S-W Asia):



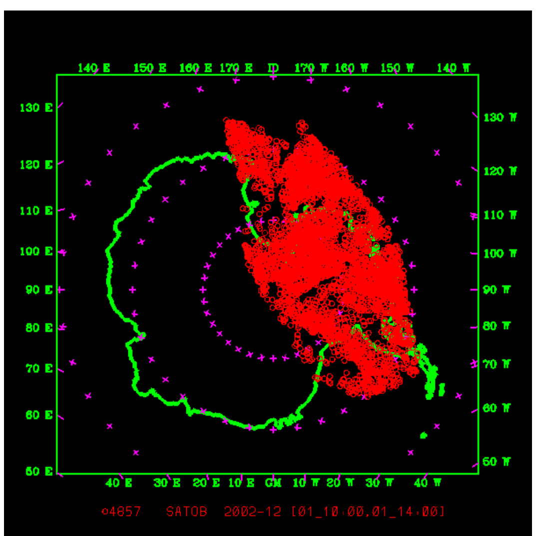
Indian NCMRWF 30km:



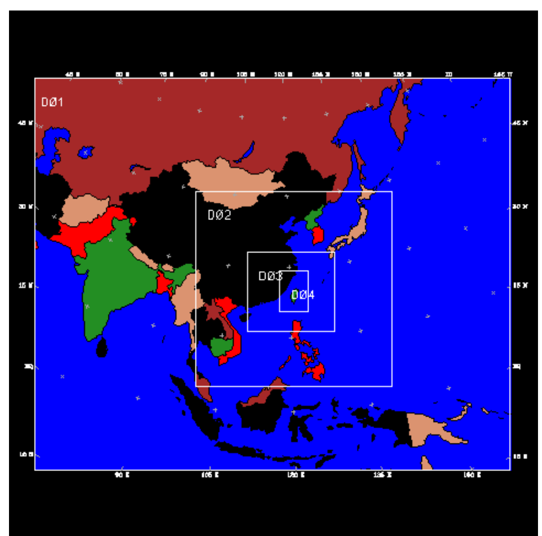
Korean 10km:



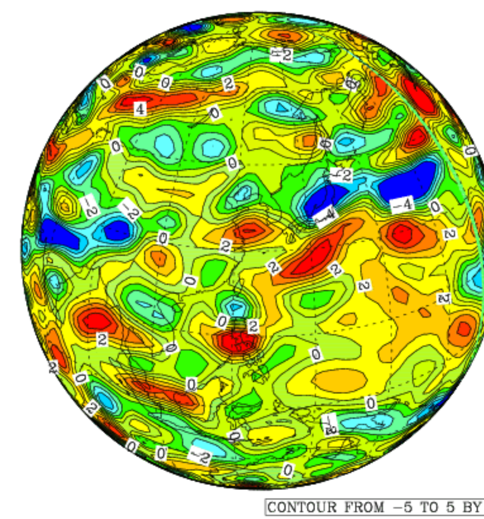
AMPS 30km:



Taiwanese CAA 135/45/15km:



Korean T213/T426:



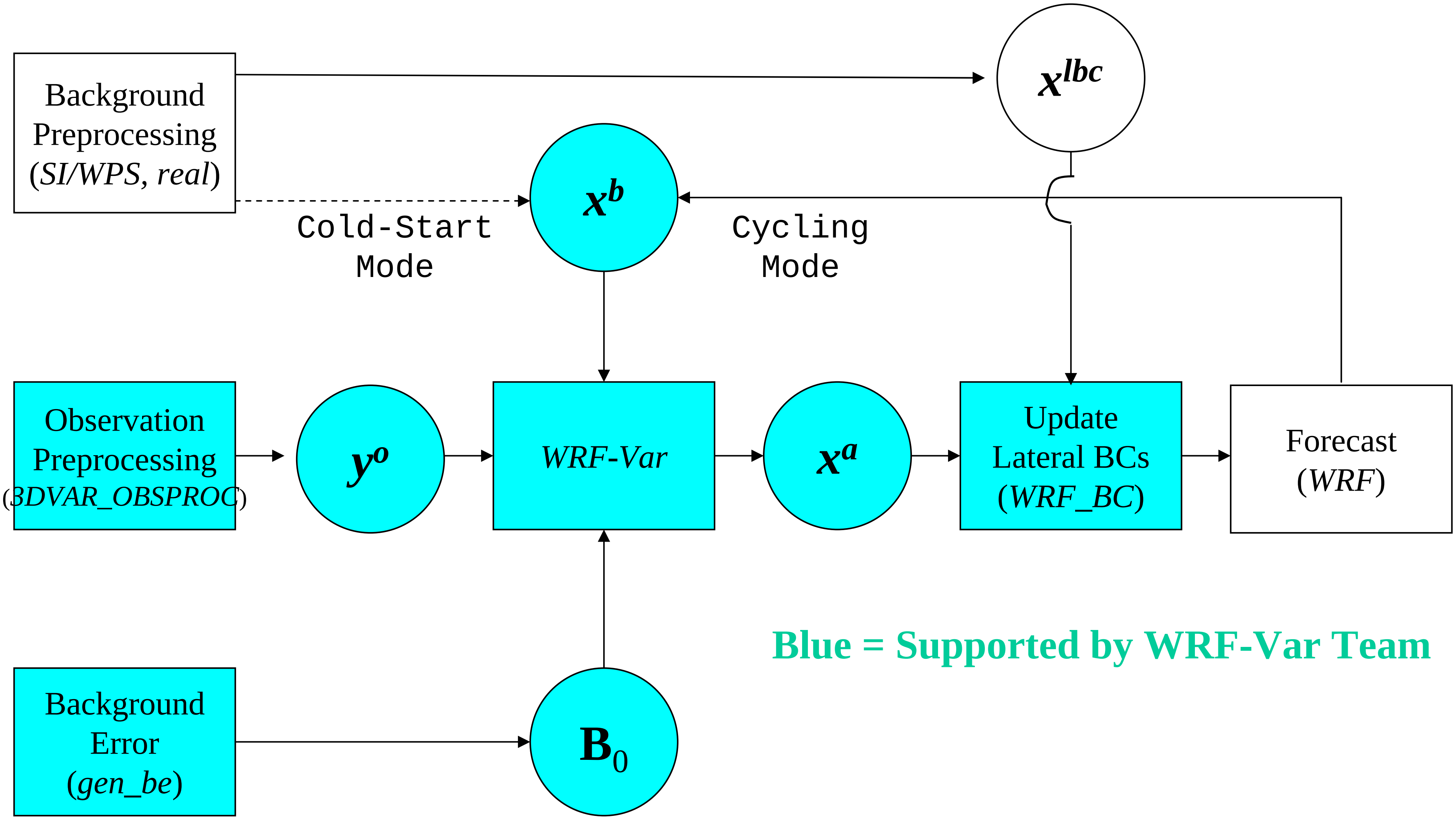
New Features Of WRF-Var Version 2.1 (Release July 2005)

- First Guess at Appropriate Time (FGAT).
- Radar reflectivity.
- Other new obs: GPS refractivity, MODIS AMVs.
- Platforms: IBM-SP, DEC, Linux, SGI, Cray X1, Apple G4/G5.
- Initial 4D-Var modifications.
- New utility *gen_be* to calculate local background error statistics.
- Global 3D-Var capability.

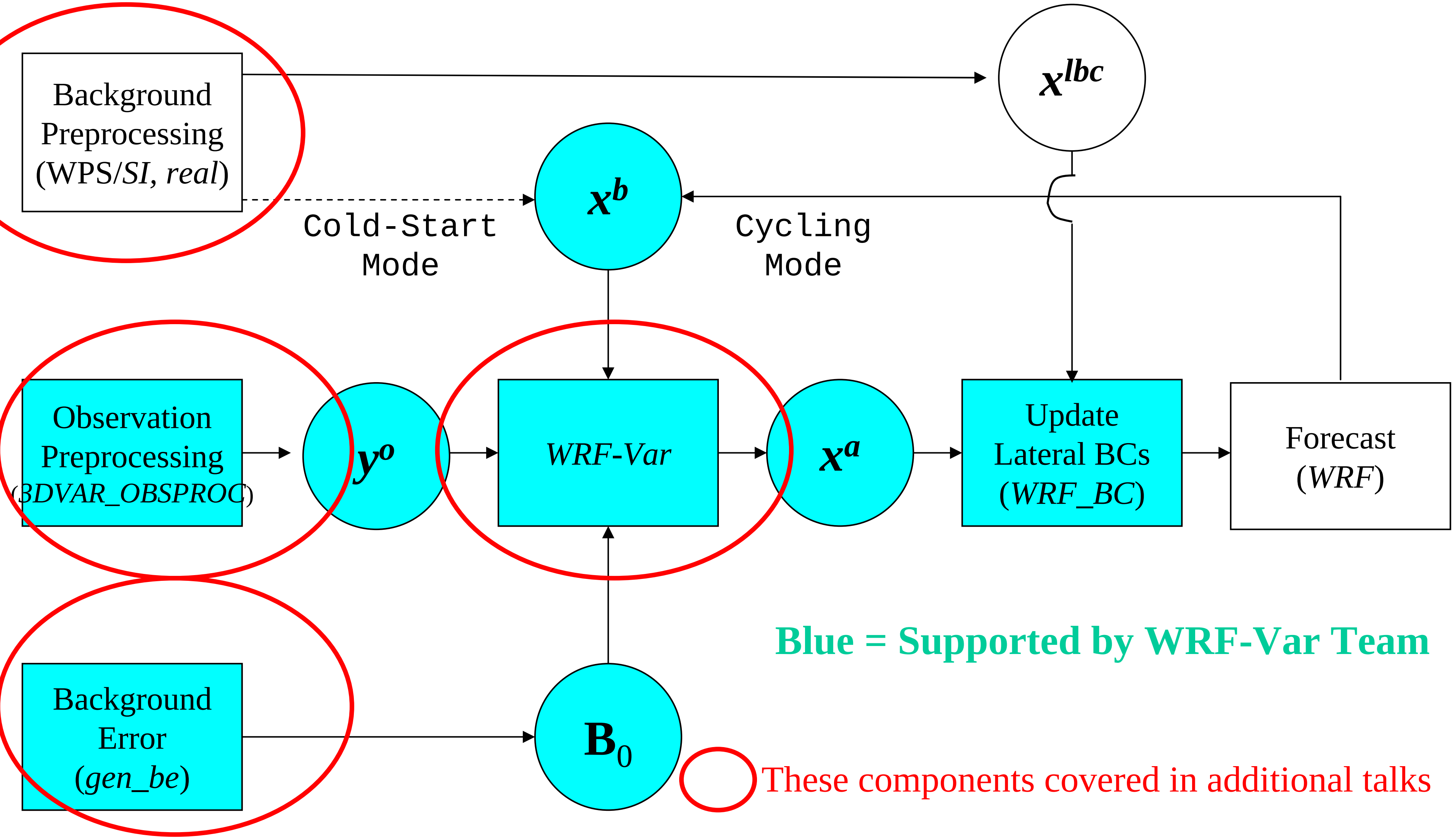
WRF-Var Version 2.2 (release February 2007?)

- Major new features/improvements (provisional):
  - Major software engineering reorganization.
  - Use of [PREP]BUFR for observation ingest.
  - Enhanced gen_be utility (EPS-based statistics, efficiency).
  - Flow-dependent forecast error covariances.
  - Remove obsolete features (e.g. MM5/GFS-based errors).
  - Radiance data assimilation.
- NOT included: 4D-Var (under development).

WRF-Var in the WRF Modeling System



WRF-Var in the WRF Modeling System



2. Practical Variational Data Assimilation

Need For Data Assimilation in NWP

Fact: There are never enough good observations!!

- Consider NWP model:
  - Typical global model – 425 * 325 * 50 = 6.9 million gridpoints.
  - Minimum number of prognostic variables = 6 *(u, v, w, T, p, q)*.
  - Number of degrees of freedom = 41.4 million.

- Typical number of observations = few x 10^6 but:
  - Inhomogeneous distribution of data.
  - Observations not always in sensitive areas.
  - Observations have errors.

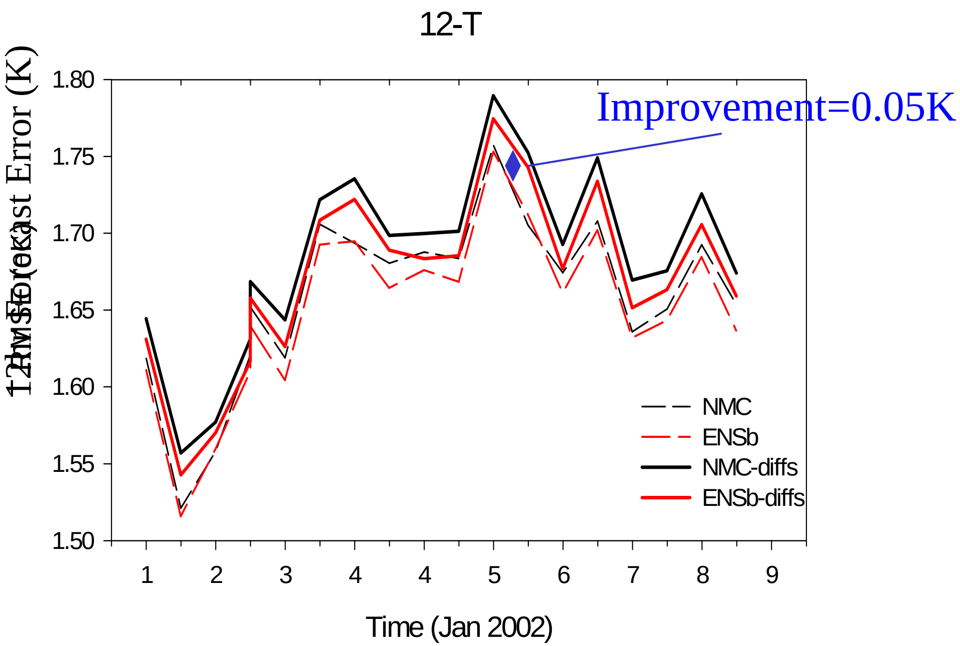
- Solutions:
  - Use sophisticated (variational/ensemble) techniques (can use “exotic” obs).
  - Use previous forecast to propagate obs. info from previous times.
  - Use approximate physical balance relationships.
  - More/better observations!

Importance of Data Assimilation For General WRF Development/Testing

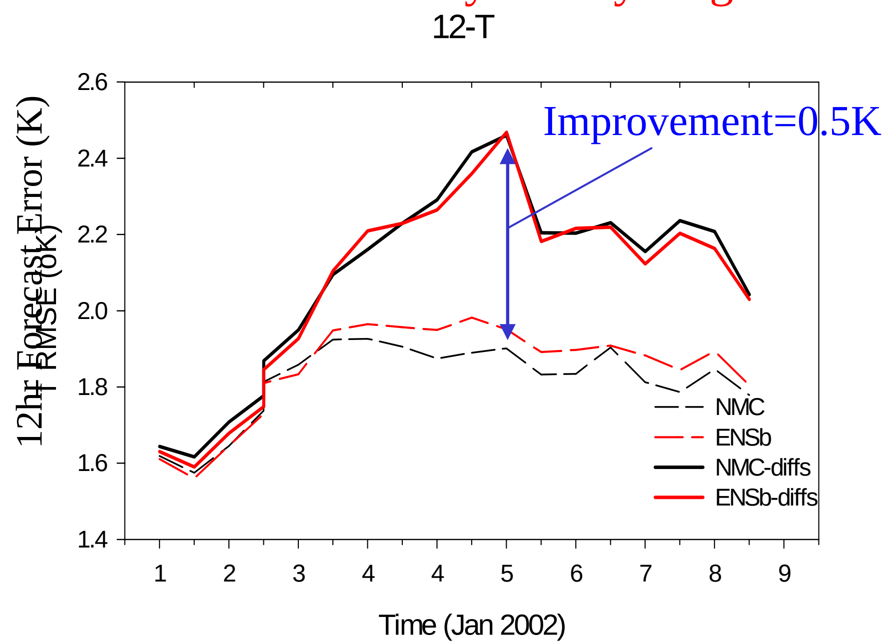
Experiment (Mi-Seon Lee, KMA):

- Test undisclosed change to WRF modeling system.
- 40km WRF CONUS application. Solid = Control, Dashed = Test.
- Use January 2002 conventional data for cycling.

Without DA Cycling

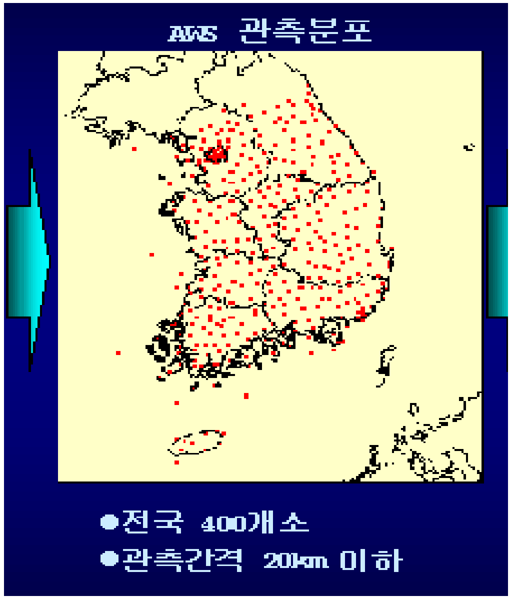


With 12hourly DA Cycling



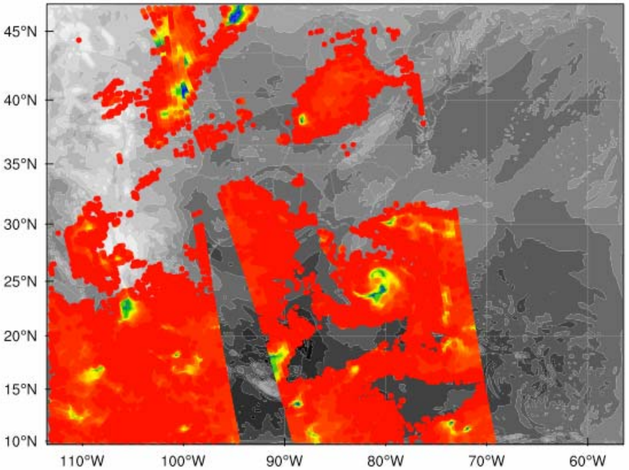
What Is Data Assimilation?

- Assimilation system combines:
  - Observations - $\boldsymbol{y^o}$
  - Previous forecast (“background field $\boldsymbol{x^b}$)
  - Estimate of errors in observations/background.
  - Laws of physics.



- Assimilation system outputs an “analysis”.

- Analysis used in a number of ways:
  - Initial conditions for numerical forecasts.
  - Climatology - reanalyses.
  - Observing system design (e.g. OSSEs).



Variational Data Assimilation

• Variational data assimilation systems attempt to find an analysis $\boldsymbol{x^a}$ that minimizes a cost-function

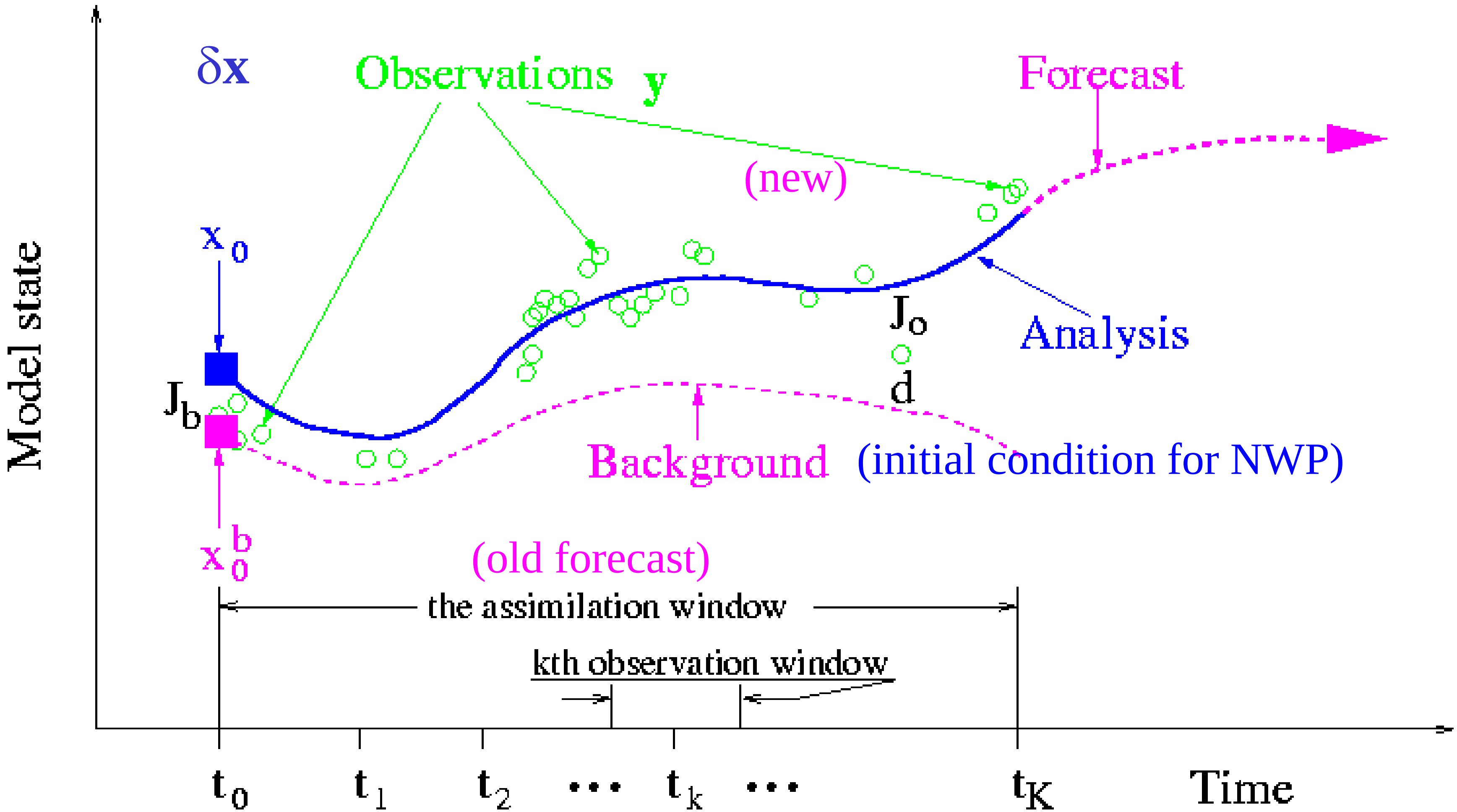
$$J = J_b + J_o$$

• Three-Dimensional Variational Data Assimilation = 3D-Var (first implemented at NCEP - Parrish and Derber 1992).

• Four-Dimensional Variational Data Assimilation = 4D-Var. First implemented at ECMWF - Rabier et al. 2000).

•4D-Var includes the time dimension by including the forecast model as part of the data assimilation system.

4D Variational Data Assimilation



Variational Data Assimilation

- The components J_b and J_o of the cost function are defined as

$$J_b\left[\mathbf{x}(t_0)\right]=\frac{1}{2}\left[\mathbf{x}(t_0)-\mathbf{x}^b(t_0)\right]^T\mathbf{B}_o^{-1}\left[\mathbf{x}(t_0)-\mathbf{x}^b(t_0)\right]$$

$$J_o\left[\mathbf{x}(t_0)\right]=\frac{1}{2}\sum_{i=0}^{n}\left[\mathbf{y}_i-\mathbf{y}_i^o\right]^T\mathbf{R}_i^{-1}\left[\mathbf{y}_i-\mathbf{y}_i^o\right]$$

- $\mathbf{B_0}$ is an *a priori* weight matrix estimating the error covariance of $\mathbf{x^b}$.

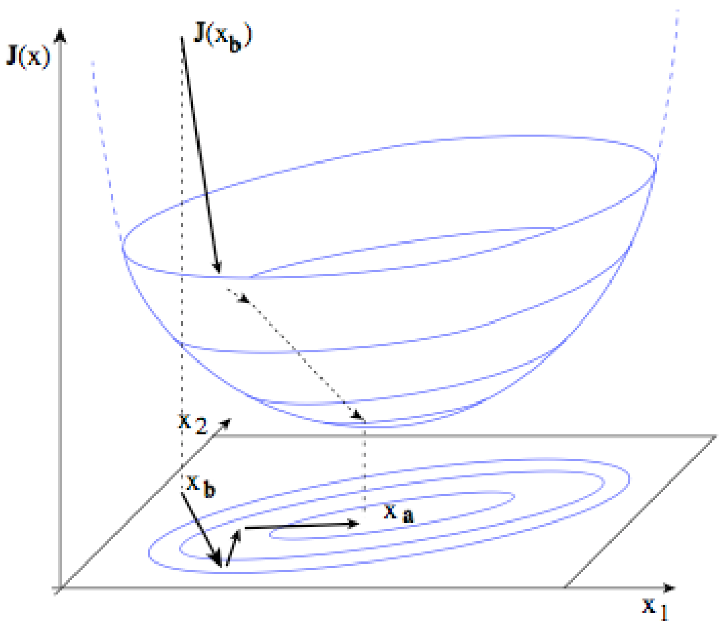
- The direct calculation of J_b and J_o is impossible for NWP problems ($\mathbf{B_0}$, $\mathbf{R}$ are matrices of dimension 10^7). Therefore many practical simplifications required for real-world models.

- *Incremental Var* produces analysis increments that are added back to a *first guess* field $\mathbf{x^g}$ to produce the analysis, i.e.

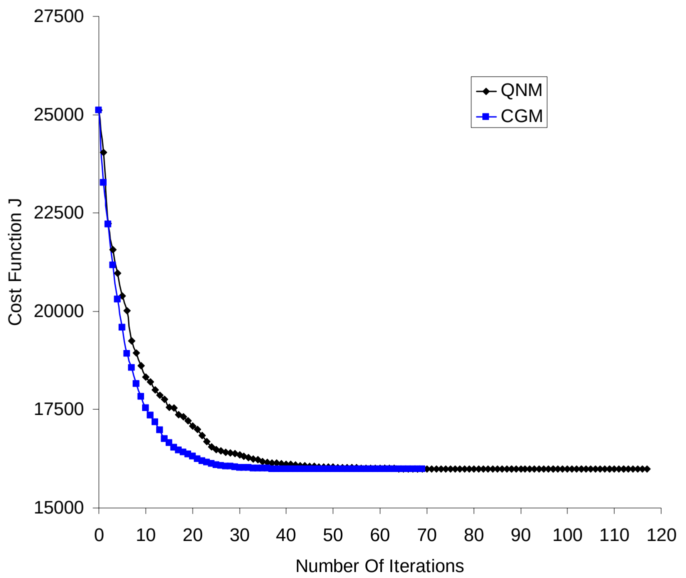
$$\mathbf{x}^a(t_0)\equiv\mathbf{x}^g(t_0)+\delta\mathbf{x}(t_0)$$

Minimization Of The Cost Function

- Minimization of the cost function proceeds iteratively.



From Bouttier and Courtier (1999)



From WRF-Var tutorial

- "Convergence" achieved when either 1) Maximum iterations reached, 2) Ratio final/initial gradient hits a specified criterion.

3. Background Error Modeling

Incremental WRF-Var J_b Preconditioning

$$J_b\left[\delta\mathbf{x}(t_0)\right] = \frac{1}{2}\left\{\delta\mathbf{x}(t_0) - \left[\mathbf{x}^b(t_0) - \mathbf{x}^g(t_0)\right]\right\}^T \mathbf{B}_o^{-1}\left\{\delta\mathbf{x}(t_0) - \left[\mathbf{x}^b(t_0) - \mathbf{x}^g(t_0)\right]\right\}$$

- *Define* ***preconditioned control variable v*** space transform

$$\delta\mathbf{x}(t_0) = \mathbf{U}\mathbf{v}$$

where **U** transform CAREFULLY chosen to satisfy $\mathbf{B_o} = \mathbf{U}\mathbf{U}^T$.

- Choose (at least assume) control variable components with uncorrelated errors:

$$J_b\left[\delta\mathbf{x}(t_0)\right] = \frac{1}{2}\sum_n v_n^2$$

- where n~number pieces of independent information.

WRF-Var Background Error Modeling

cv_options		2 (original MM5)	3(GSI)	4 (Global)	5(regional)
Analysis increments	$\mathbf{x}'$	$u',v',T',q',p_s'(i,j,k)$			
Change of Variable	U_p	$\psi',\chi',p_u',q'(i,j,k)$	$\psi',\chi_u',T_u',\tilde{r}',p_{su}'(i,j,k)$		
Vertical Covariances	U_v	$\mathbf{B}=\mathbf{E}\Lambda\mathbf{E}^{\mathrm{T}}$	RF	$\mathbf{B}=\mathbf{E}\Lambda\mathbf{E}^{\mathrm{T}}$	
Horizontal Correlations	U_h	RF		Spectral	RF
Control Variables	$\mathbf{v}$	$\mathbf{v}(i,j,m)$		$\mathbf{v}(l,n,m)$	$\mathbf{v}(i,j,m)$

$$\delta\mathbf{x}(t_0)=\mathbf{U}\mathbf{v}=\mathbf{U}_p\mathbf{U}_v\mathbf{U}_h\mathbf{v}$$

Up: Change of variable, impose balance.

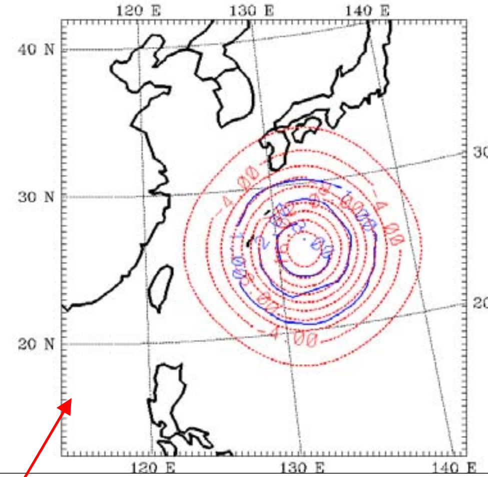
Uv: Vertical correlations EOF Decomposition

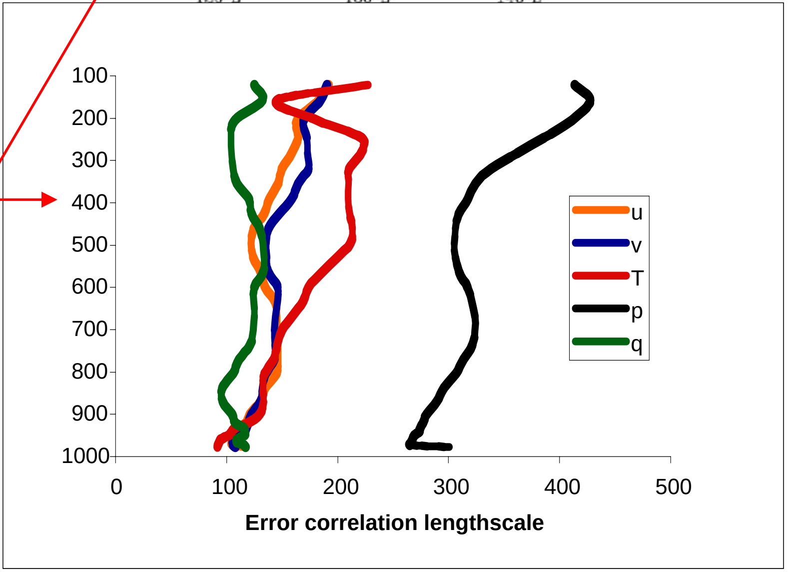
RF = Recursive Filter, e.g. Purser et al 2003

WRF-Var Background Error Modeling

$$\delta \mathbf{x}(t_0) = \mathbf{U}\mathbf{v} = \mathbf{U}_p \mathbf{U}_v \mathbf{U}_h \mathbf{v}$$

cv_options		2 (original MM5)	3(GSI)	4 (Global)	5(regional)
Analysis increments	$\mathbf{x}'$	$u',v',T',q',p_s'(i,j,k)$			
Change of Variable	U_p	$\psi',\chi',p_u',q'(i,j,k)$	$\psi',\chi_u',T_u',\tilde{r}',p_{su}'(i,j,k)$		
Vertical Covariances	U_v	$\mathbf{B} = \mathbf{E}\Lambda\mathbf{E}^T$	RF	$\mathbf{B} = \mathbf{E}\Lambda\mathbf{E}^T$	
Horizontal Correlations	U_h	RF		Spectral	RF
Control Variables	$\mathbf{v}$	$\mathbf{v}(i,j,m)$		$\mathbf{v}(l,n,m)$	$\mathbf{v}(i,j,m)$

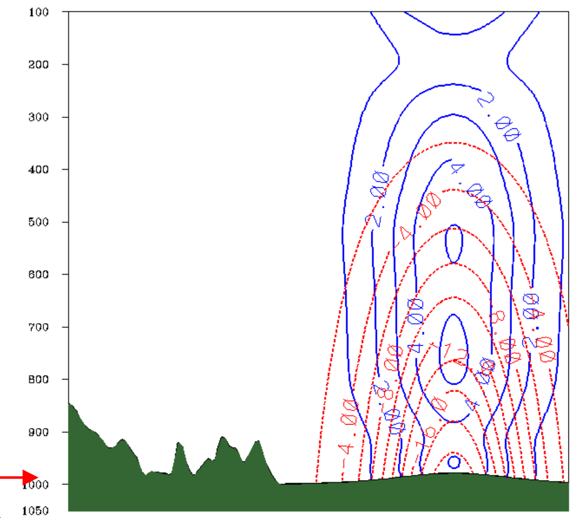


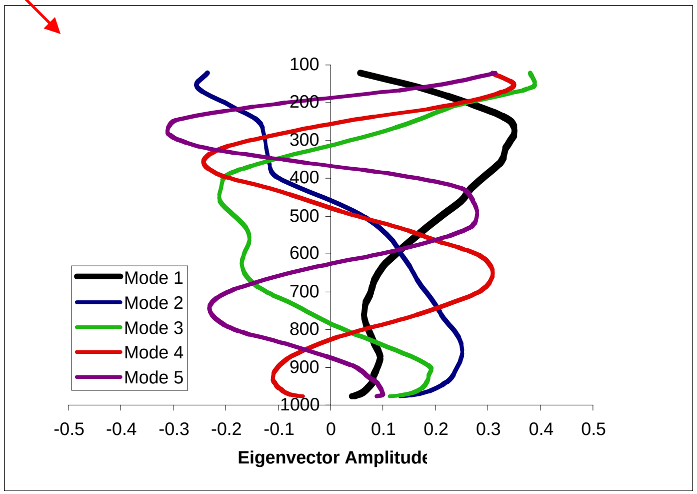


WRF-Var Background Error Modeling

cv_options		2 (original MM5)	3(GSI)	4 (Global)	5(regional)
Analysis increments	x'	$u',v',T',q',p_s'(i,j,k)$			
Change of Variable	U_p	$\psi',\chi',p_u',q'(i,j,k)$	$\psi',\chi_u',T_u',\tilde{r}',p_{su}'(i,j,k)$		
Vertical Covariances	U_v	$\mathbf{B}=\mathbf{E}\Lambda\mathbf{E}^T$	RF	$\mathbf{B}=\mathbf{E}\Lambda\mathbf{E}^T$	
Horizontal Correlations	U_h	RF		Spectral	RF
Control Variables	v	$\mathbf{v}(i,j,m)$		$\mathbf{v}(l,n,m)$	$\mathbf{v}(i,j,m)$

$$\delta\mathbf{x}(t_0) = \mathbf{U}\mathbf{v} = \mathbf{U}_p\mathbf{U}_v\mathbf{U}_h\mathbf{v}$$





WRF-Var Background Error Modeling

cv_options		2 (original MM5)	3(GSI)	4 (Global)	5(regional)
Analysis increments	$\mathbf{x}'$	$u',v',T',q',p_s'(i,j,k)$			
Change of Variable	U_p	$\psi',\chi',p_u',q'(i,j,k)$	$\psi',\chi_u',T_u',\tilde{r}',p_{su}'(i,j,k)$		
Vertical Covariances	U_v	$\mathbf{B}=\mathbf{E}\Lambda\mathbf{E}^{\mathrm{T}}$	RF	$\mathbf{B}=\mathbf{E}\Lambda\mathbf{E}^{\mathrm{T}}$	
Horizontal Correlations	U_h	RF		Spectral	RF
Control Variables	$\mathbf{v}$	$\mathbf{v}(i,j,m)$		$\mathbf{v}(l,n,m)$	$\mathbf{v}(i,j,m)$

$$\delta\mathbf{x}(t_0)=\mathbf{U}\mathbf{v}=\mathbf{U}_p\mathbf{U}_v\mathbf{U}_h\mathbf{v}$$

Define control variables:

$$\psi'$$

$$r'=q'/q_s(T_b,q_b,p_b)$$

$$\chi'=\chi_u'+\chi_b'(\psi')$$

$$T'=T_u'+T_b'(\psi')$$

$$p_s'=p_{su}'+p_{sb}'(\psi')$$

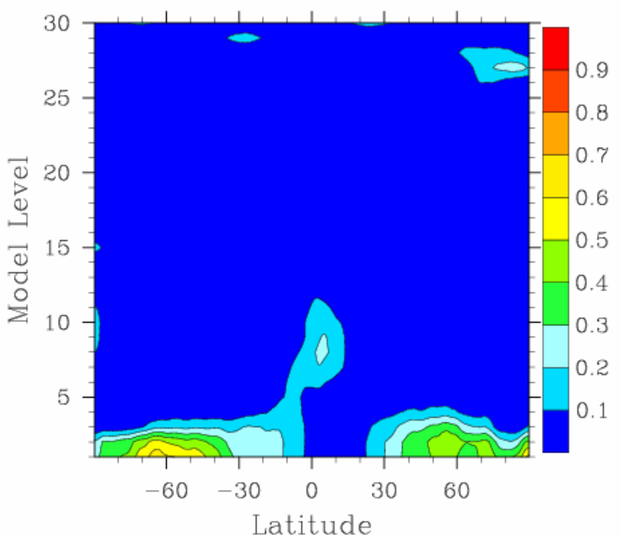
WRF-Var Statistical Balance Constraints

- Define statistical balance after Wu et al (2002):

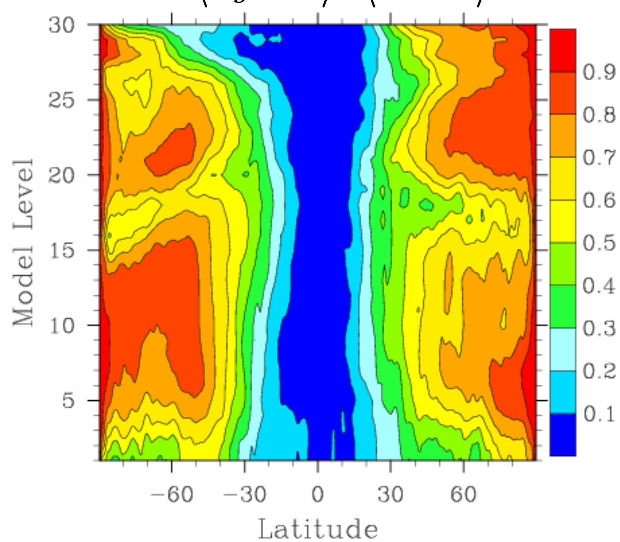
$$\chi_b^{'} = c\,\psi^{'} \quad T_b^{'}(k) = \sum_{k1} G(k,k1)\,\psi^{'}(k1) \quad p_{sb}^{'} = \sum_{k} W(k)\,\psi^{'}(k)$$

- How good are these balance constraints? Test on KMA global model data. Plot correlation between “Full” and balanced components of field:

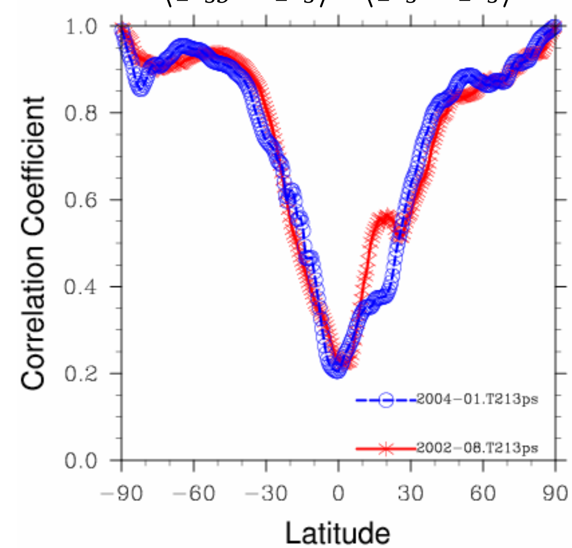
$\langle \chi_b \bullet \chi \rangle / \langle \chi \bullet \chi \rangle$



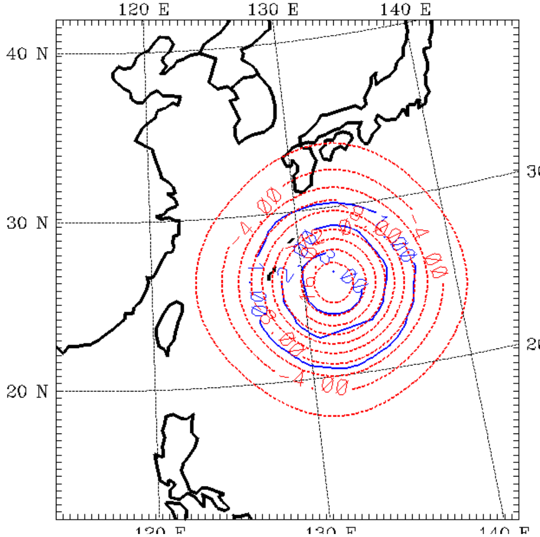
$\langle T_b \bullet T \rangle / \langle T \bullet T \rangle$



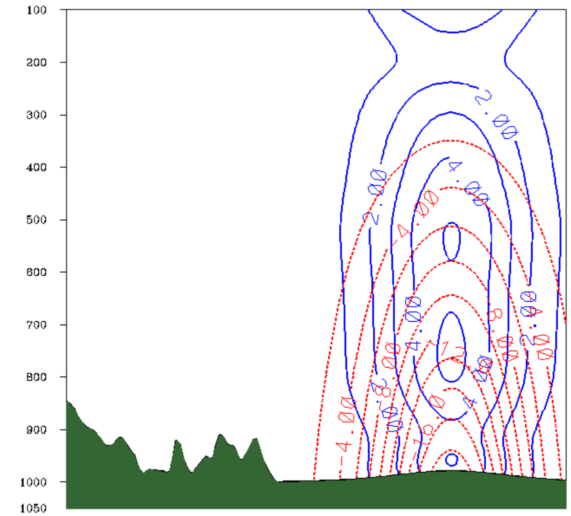
$\langle p_{sb} \bullet p_s \rangle / \langle p_s \bullet p_s \rangle$

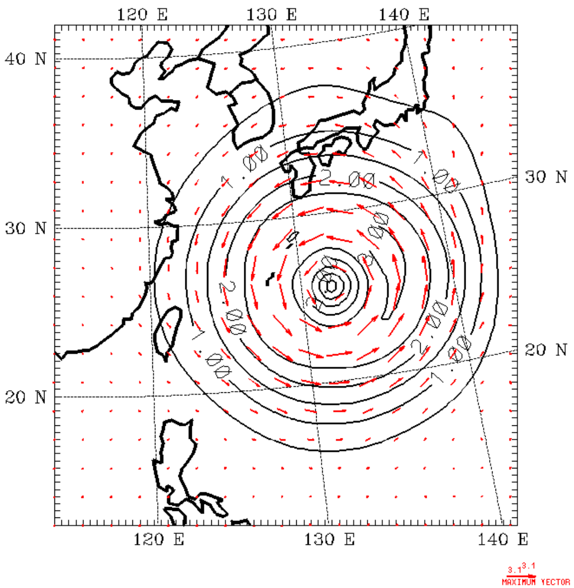


3D-Var response to a single P_s Observation

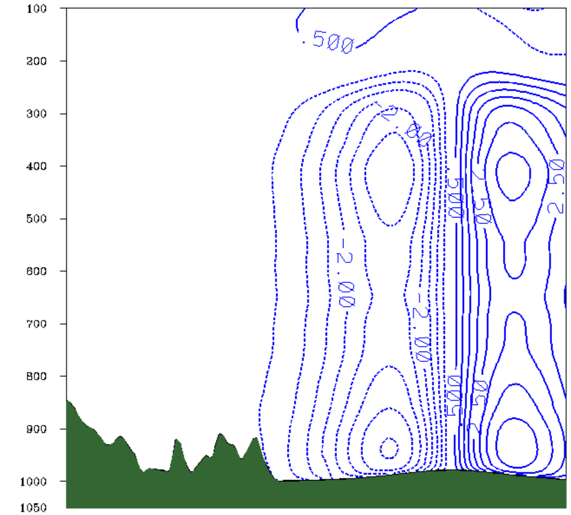


Pressure,
Temperature

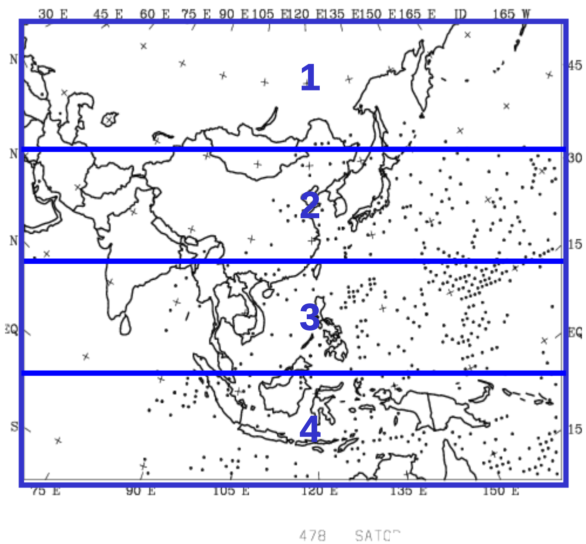




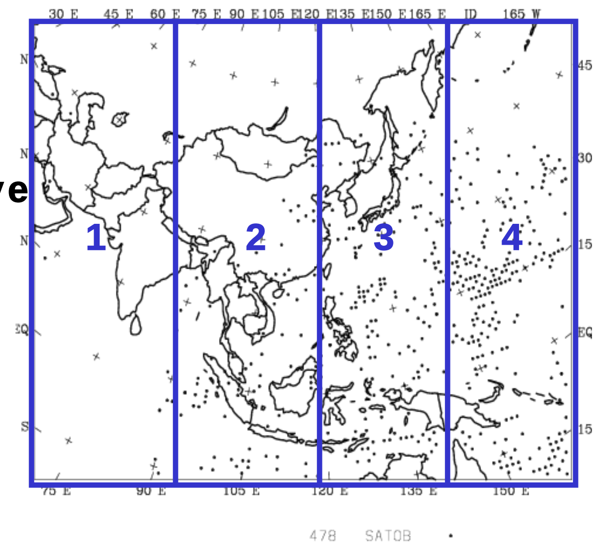
Wind Speed,
Vector,
v-wind component.



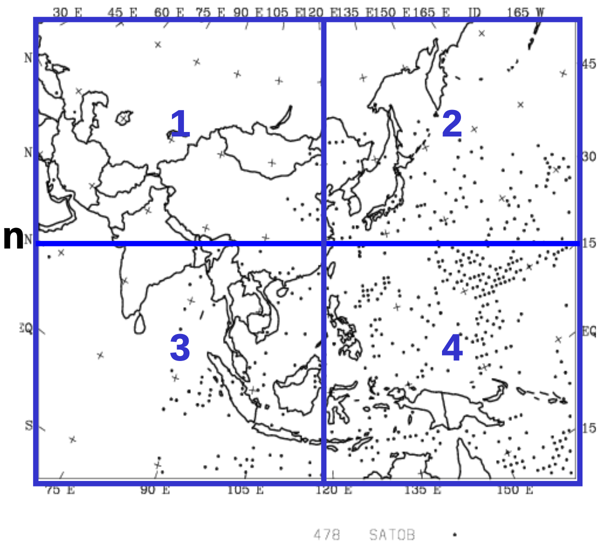
WRF-Var Parallelism (e.g. 4 processors)



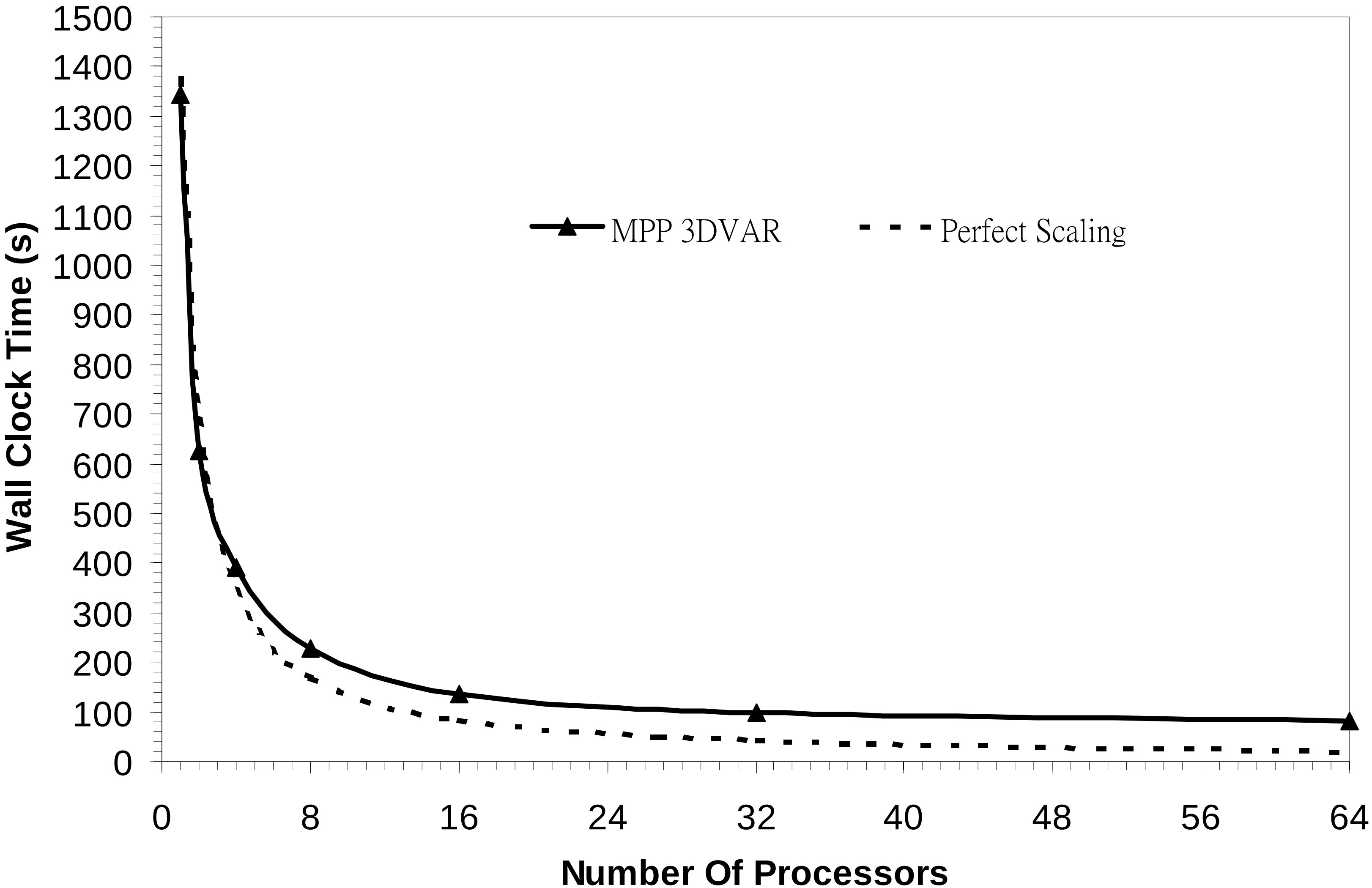
Recursive Filter and FFTs



Minimization and Forecast Model



MPI Scalability – NCAR IBM



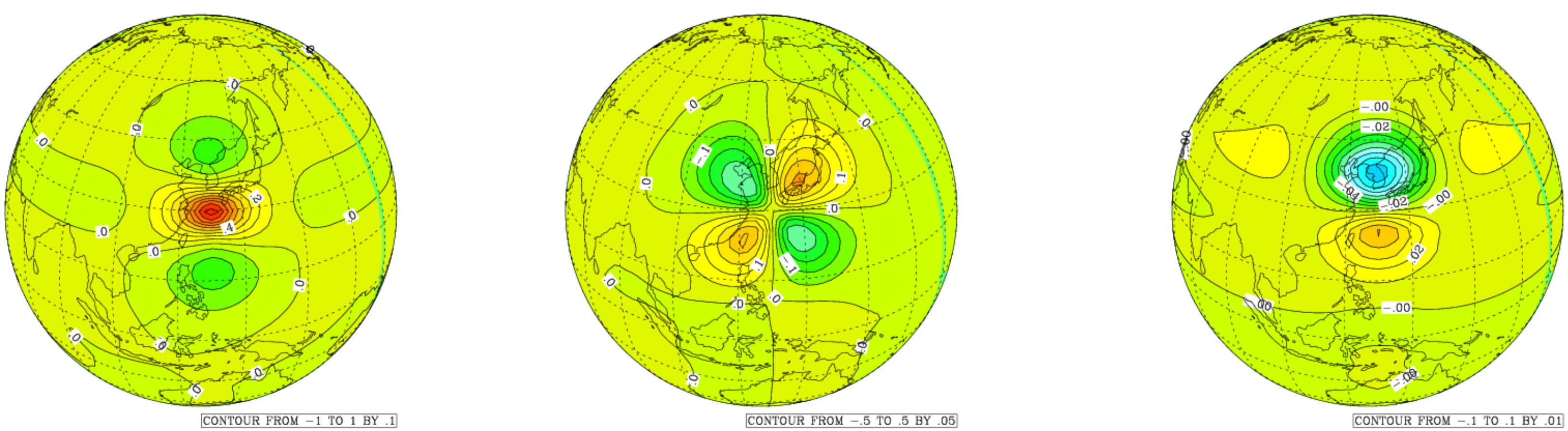
- **Test Case: 140x150x41 AFWA 45km "T4 theater" – 25th Jan 2002.**
- **Background error tuning – Old Its = 98, New = 49 (64PE = 58s).**

Global Applications of WRF-Var

Major technical changes to regional system are

1. Periodic boundary conditions.
2. New global WRF Registry created.
3. Minor changes to treat pole as a special point.
4. Spectral-Grid transformation for horizontal error correlations (FFTPACK).

U-wind Observation (O-B = 1m/s, s_o = 1m/s) at 120E, 45N, level 15:



U V T

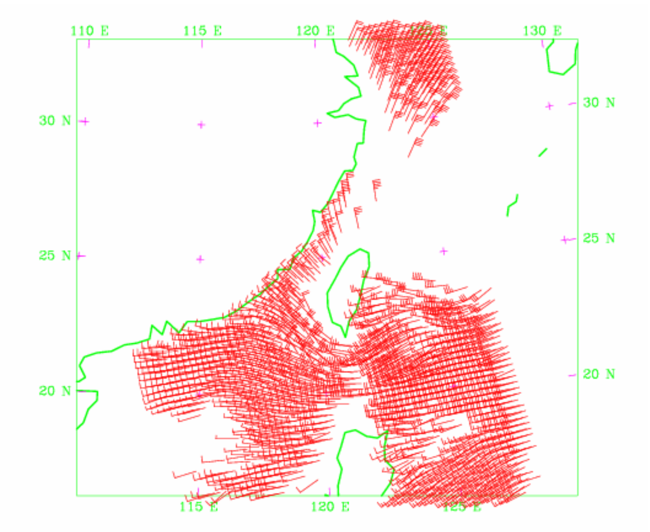
4. Observational Issues

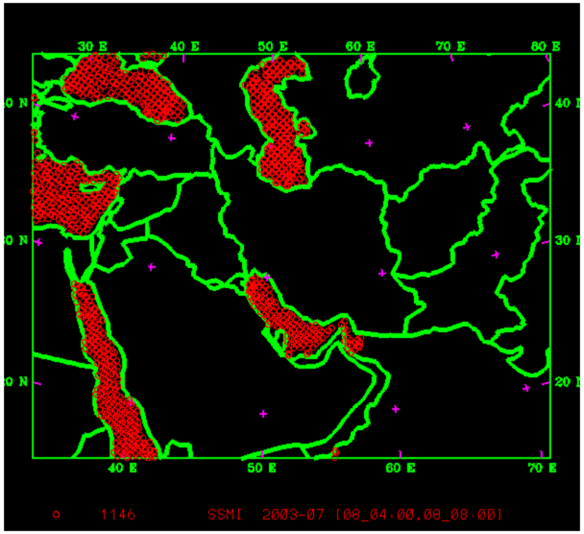
WRF-Var Observations (August 2005, V2.1 Release)

- Conventional:
  - Surface (SYNOP, METAR, SHIP, BUOY).
  - Upper air (TEMP, PIBAL, AIREP, ACARS).

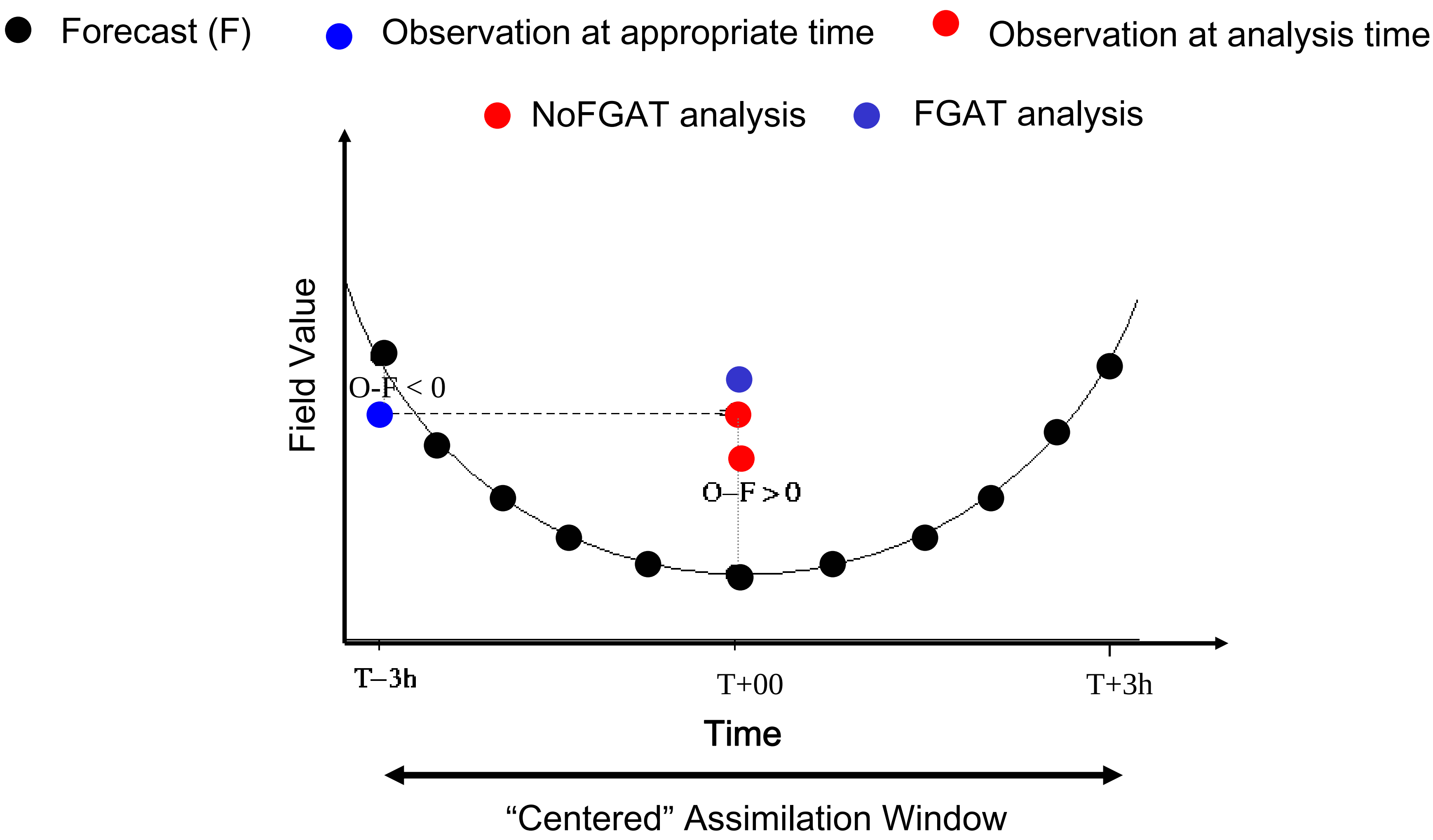
- Remotely sensed retrievals:
  - Atmospheric Motion Vectors (SATOBS, MODIS
  - Ground-based GPS Total Precipitable Water.
  - SSM/I oceanic surface wind speed and TPW.
  - Scatterometer (Quikscat) oceanic surface winds.
  - Wind Profiler.
  - **Radar radial velocity and reflectivity.**
  - ATOVS/AIRS/MODIS temperature/humidities.
  - GPS “local” refractivity.

- Radiances:
  - SSM/I brightness temperatures (Shu-Hua Chen).



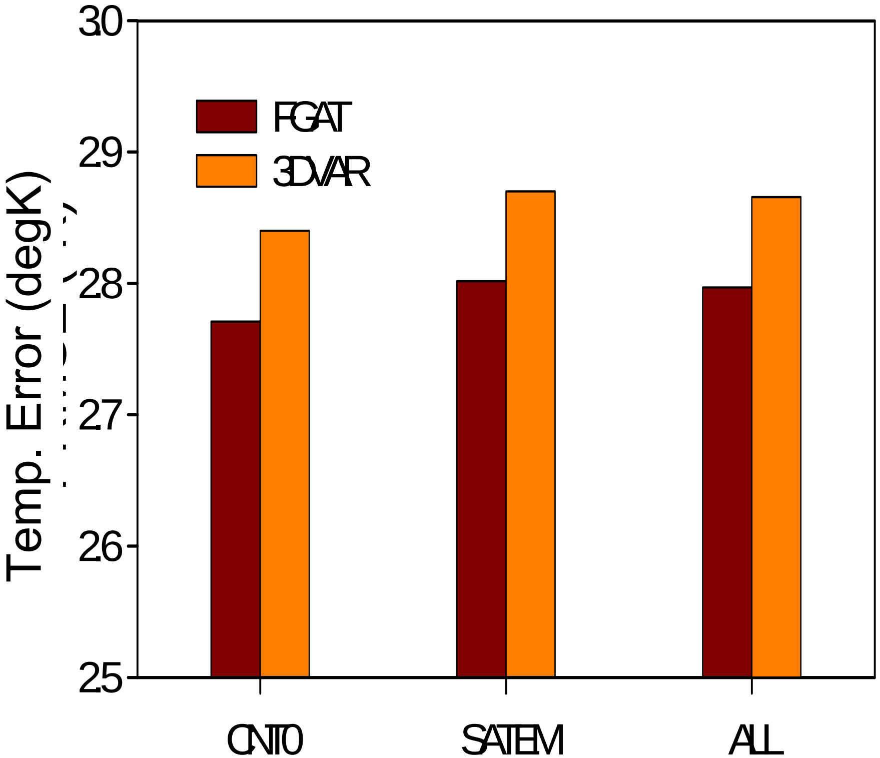


3D-Var FGAT: First Guess at Appropriate Time

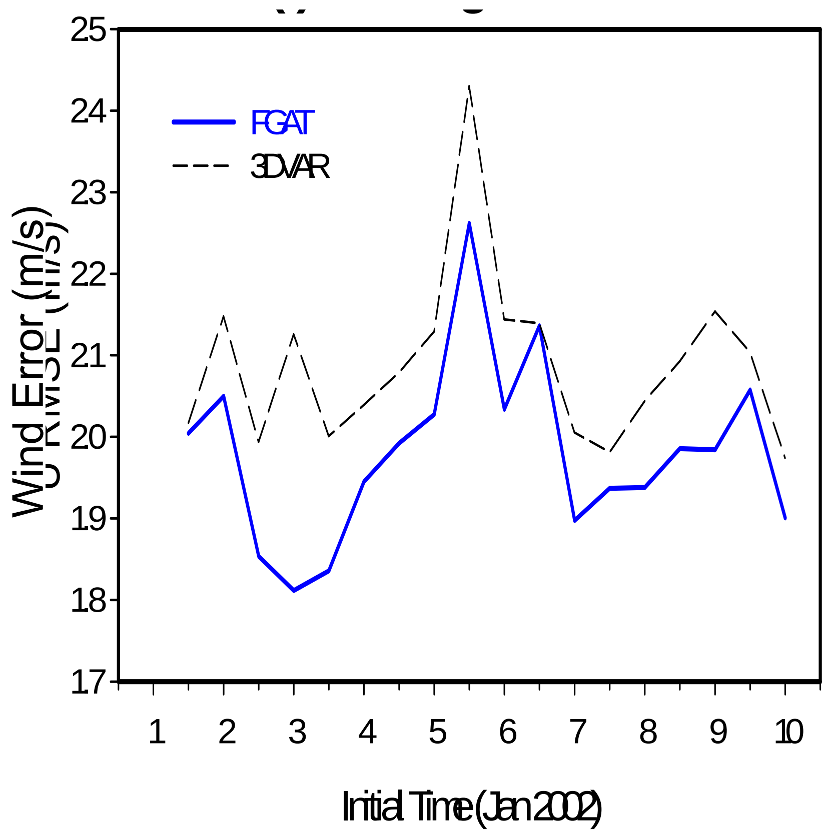


3D-Var-FGAT Forecast Impact

6hr Forecast T Error



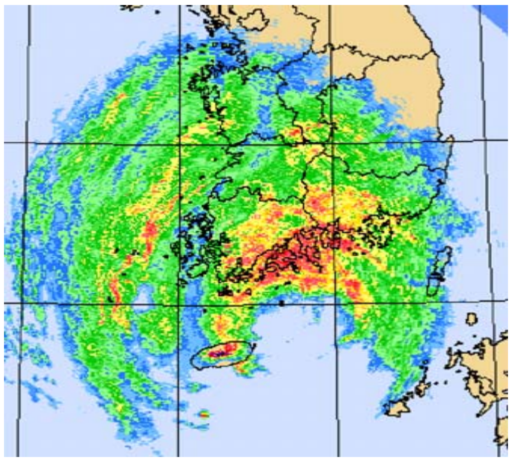
12hr Forecast U Error



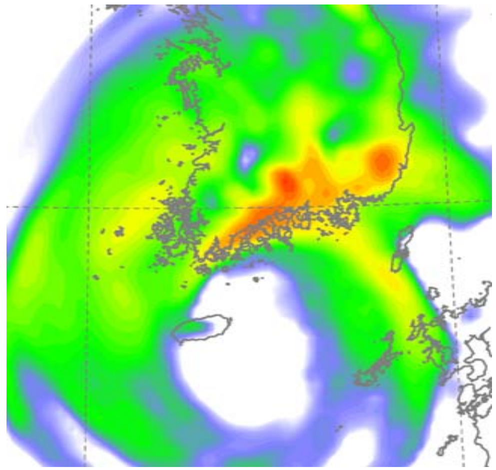
Korean Radar Data Assimilation in WRF-Var

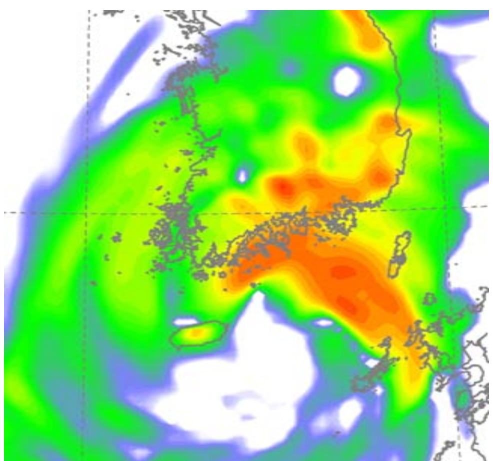
Typhoon Rusa Test Case 3hr Precip:

Obs (03Z, 31/08)

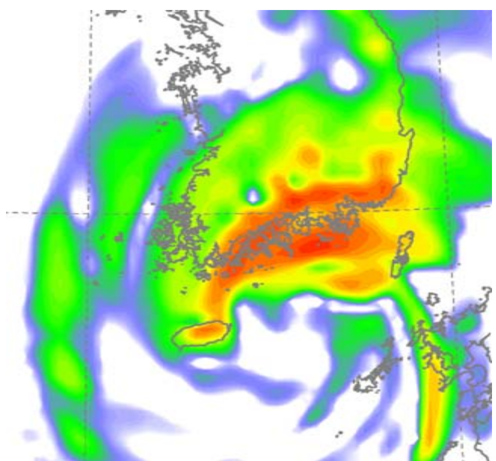


No Radar



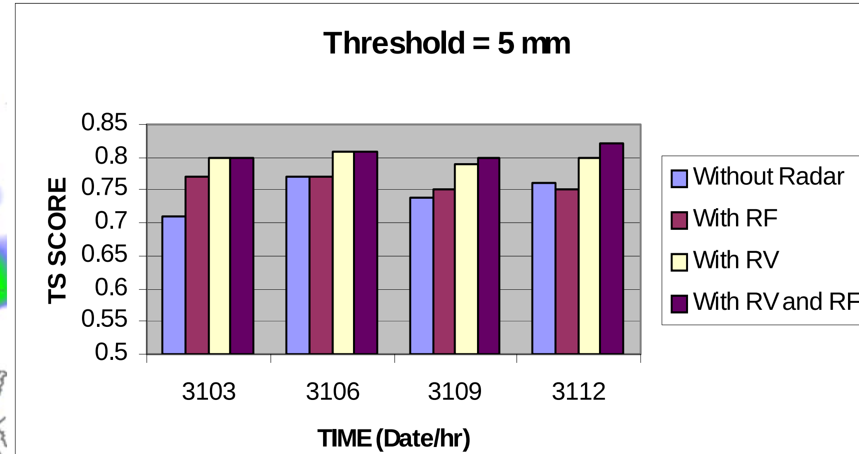


Radar RV

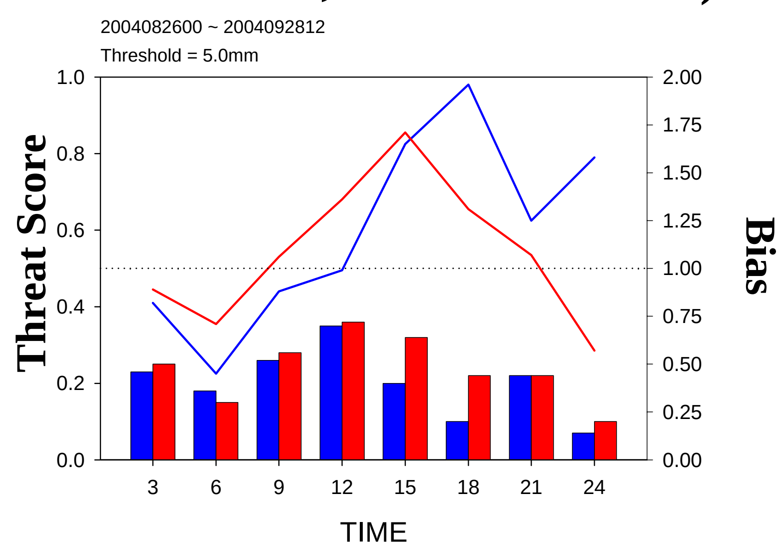


Radar RV+RF

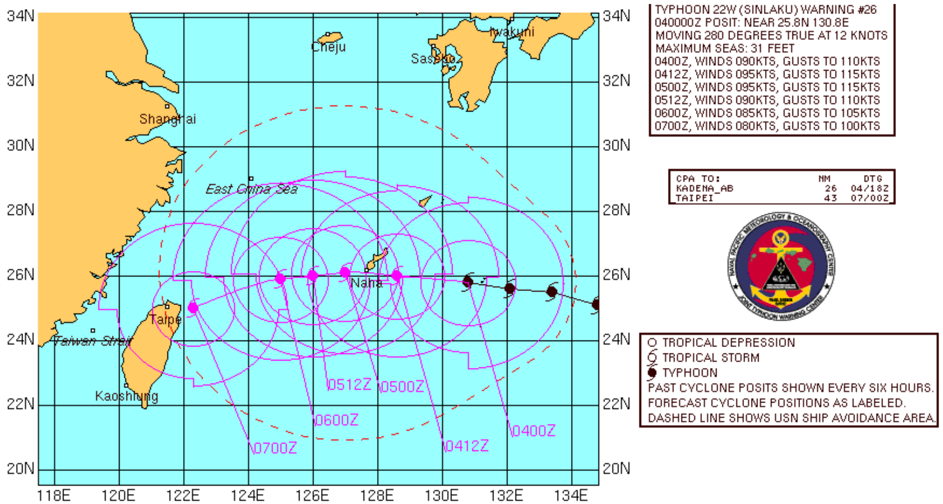
Typhoon Rusa 3hr Precip. Verification:

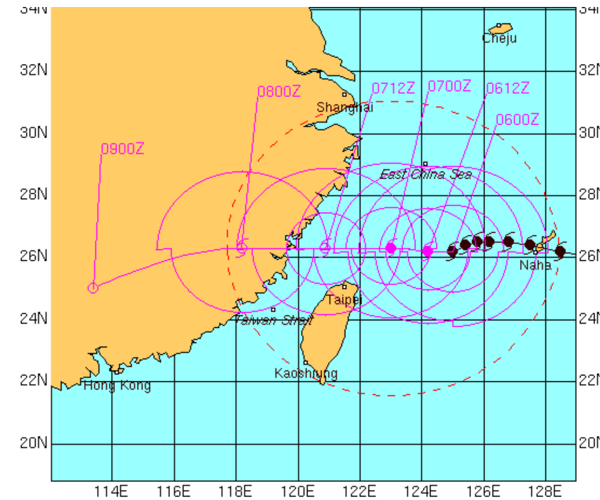


KMA Pre-operational Verification:
(no radar: blue, with radar: red)

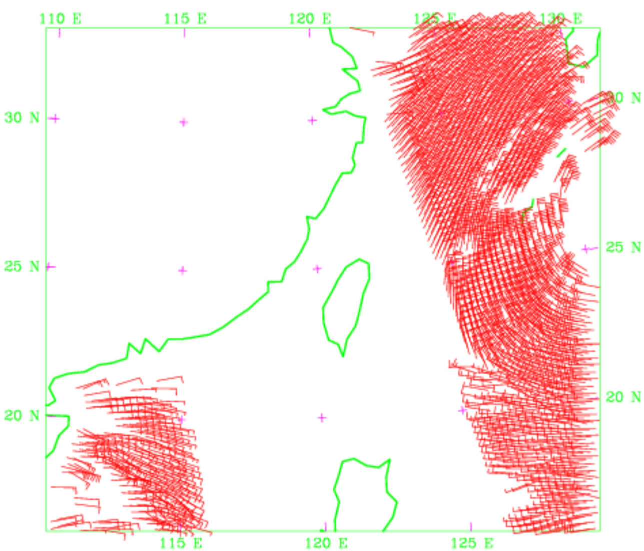


Typhoon Sinlaku: Quikscat Data



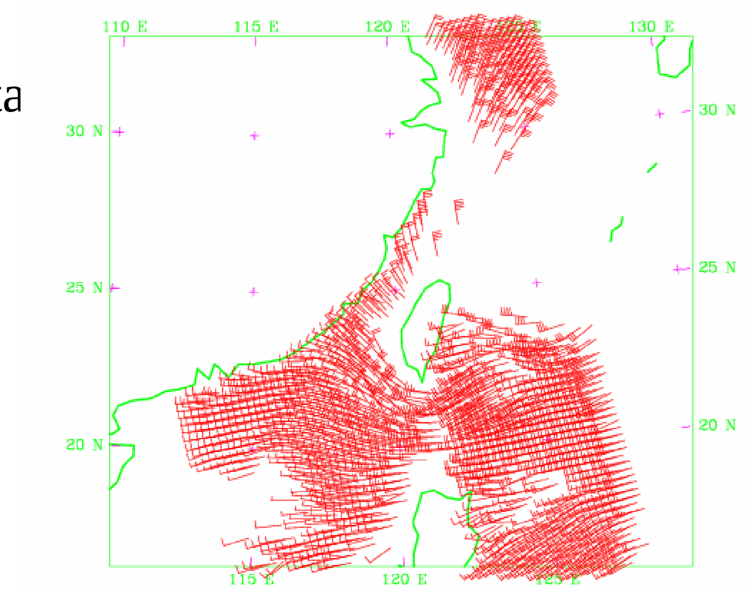


00Z September 4th 2002 **00Z September 6th 2002**

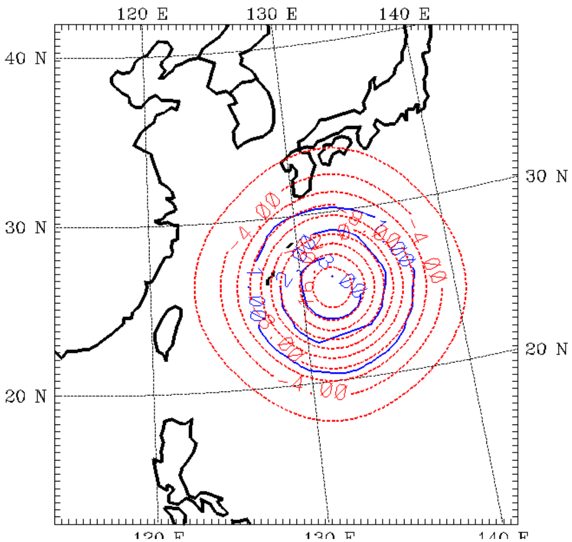


Quikscat Data

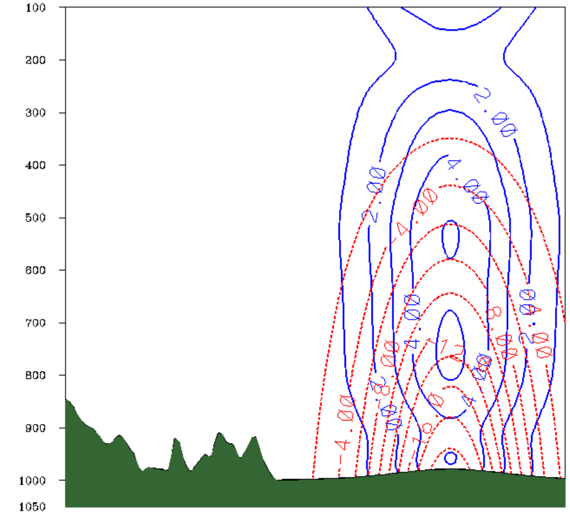
Barker et al (2004)

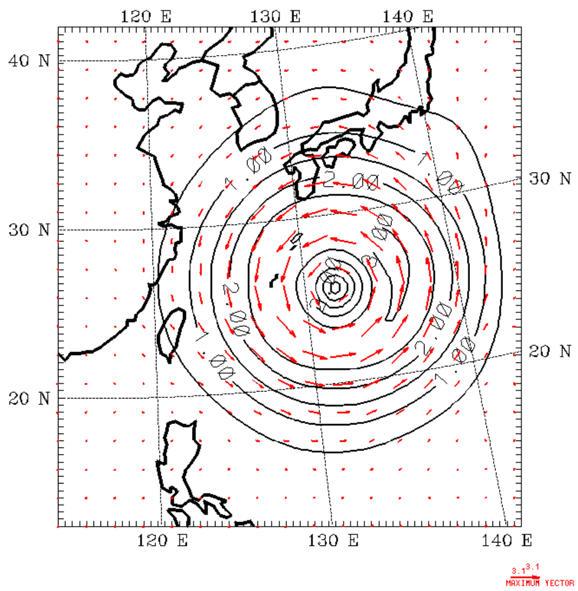


WRF-Var Sinlaku Bogus:Analysis Increments

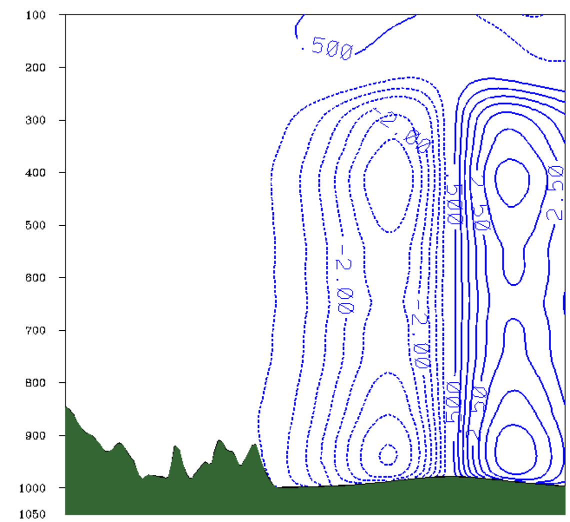


Pressure,
Temperature

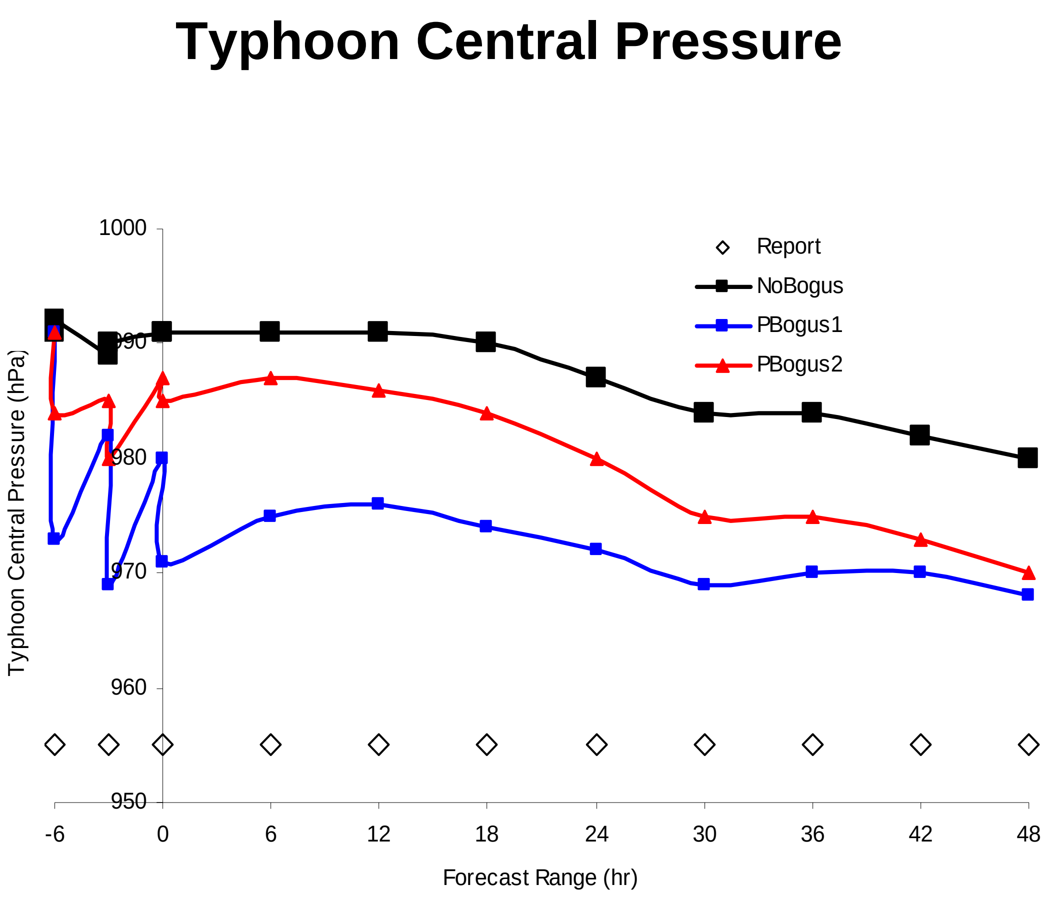




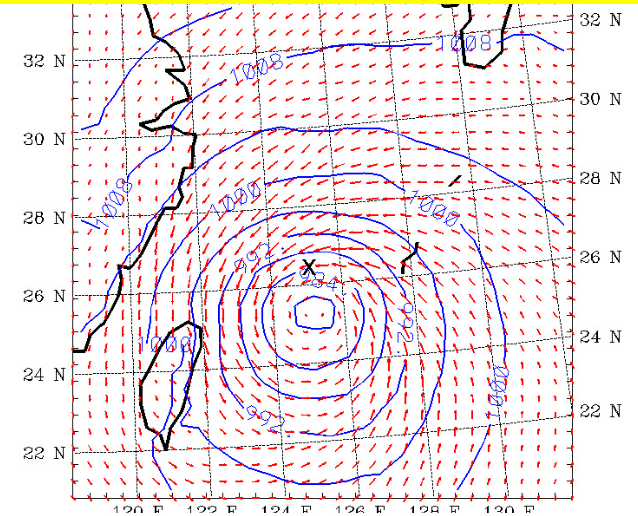
Wind Speed,
Vector,
v-wind component.



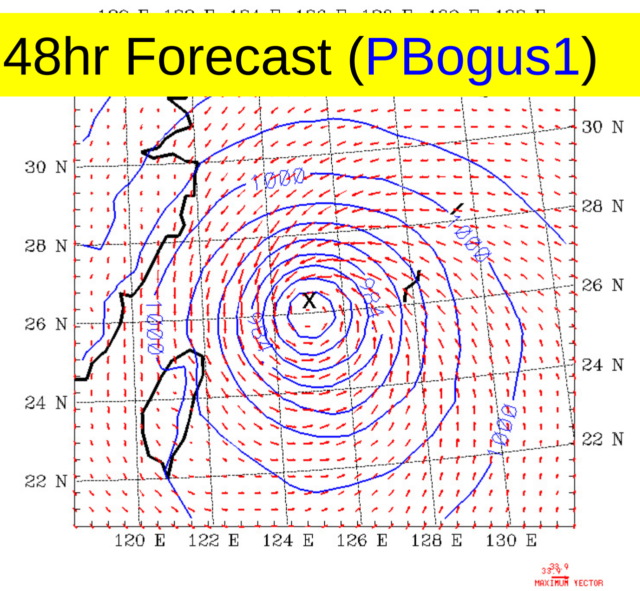
Single Surface Pressure Bogus: Forecast Impact



48hr Forecast (NoBogus)



48hr Forecast (PBogus1)



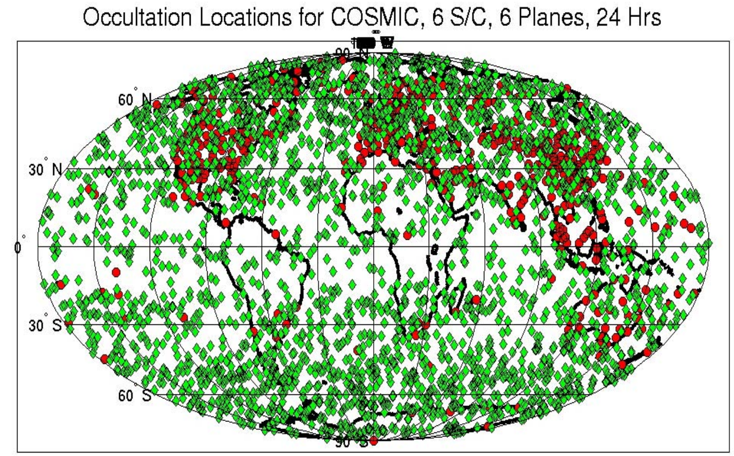
5. Current Status and Future Plans

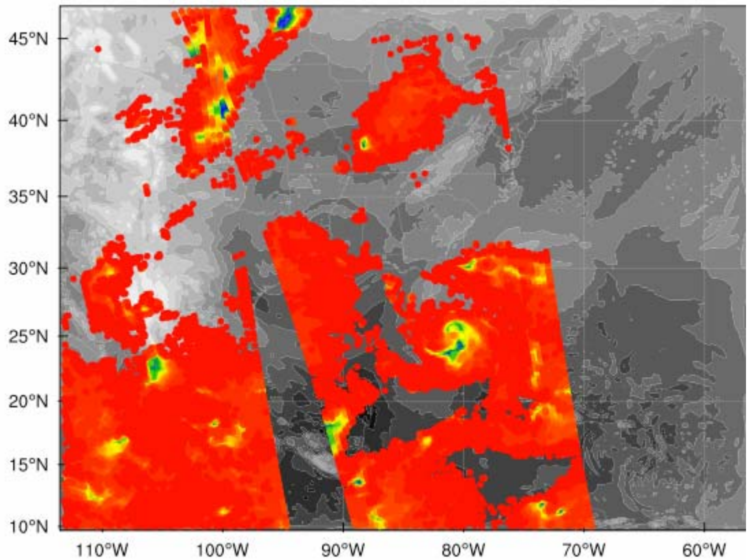
WRF-Var Observations (December 2006)

- Conventional:
  - Surface (SYNOP, METAR, SHIP, BUOY).
  - Upper air (TEMP, PIBAL, AIREP, ACARS).

- Remotely sensed retrievals:
  - Atmospheric Motion Vectors (geo/polar).
  - Ground-based GPS Total Precipitable Water.
  - SSM/I oceanic surface wind speed and TPW.
  - Scatterometer oceanic surface winds.
  - Wind Profiler.
  - Radar radial velocity and reflectivity.
  - Satellite temperature/humidities.
  - GPS refractivity (e.g. COSMIC).

- Radiances:
  - SSM/I brightness temperatures.
  - Direct radiance assimilation (SSM/I, TMI, AMSU, AIRS).



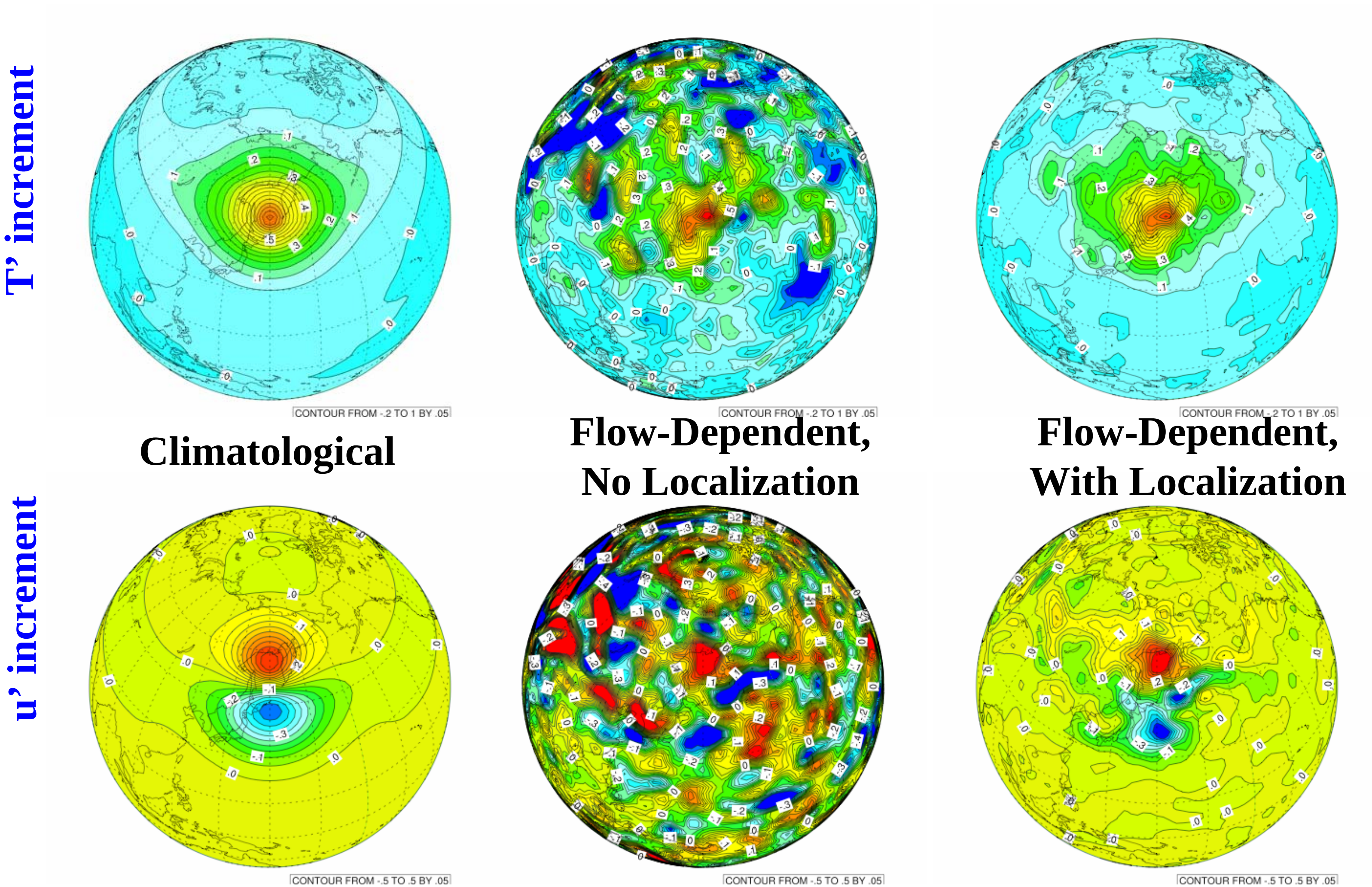


4D-Var Summary (X.-Y. Huang)

1. WRF-(4D)Var project: 2004-2006.
2. WRF-(4D)Var formulation: Multi-incremental, based on the existing WRF 3D-Var and WRF model.
3. Current status of WRF-(4D)Var: The prototype has been put together and can run. An installation has been made at AFWA. On going work:
   - Case studies and data assimilation experiments.
   - Code merge.
   - MPP.
   - JcDFI
   - Simple physics
4. Near future plan: The basic system; Multi-incremental
5. FY07 plan: Lateral boundary control (J_bdy); more physics, parallel runs, efficiency.

WRF-Var Example Flow-Dependence: Global

- Specify single T observation (O-B, σ_o=1K) at 50N, 150E, 500hPa.
- Flow-Dependence given by KMA’s Ensemble Prediction System.

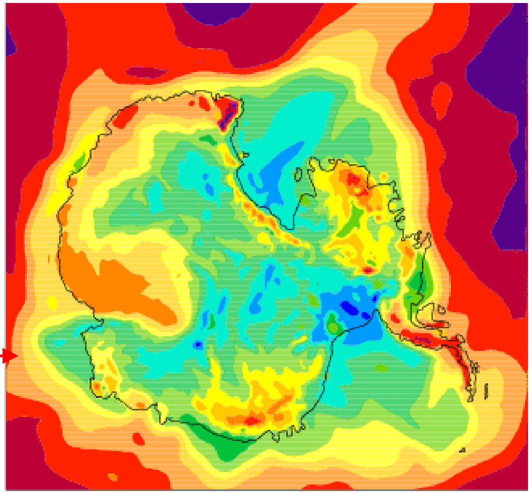


WRF-Var Plans

WRF-Var:

- Further development of radiance, radar, GPS assimilation.
- Port JCSDA's CRTM RT model to WRF-Var,
- Global 3D-Var operational at KMA (2007).
- 4D-Var operational at AFWA (March 2008).
- Applications (US, Korea, Taiwan, India, Antarctica).
- Flow-dependent covariances (4D-Var, Hybrid VAR/ENS).

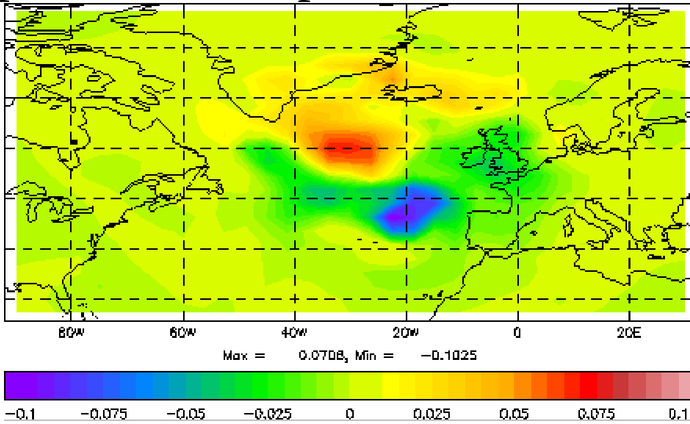
WRF 20km Antarctica:



Unified WRF Data Assimilation System (WADAS?):

- Further comparison of EnKF and 3/4D-Var.
- Hybrid variational/ensemble-based system.
- Leverages satellite radiance expertise of JCSDA, EUMETSAT, universities, etc.
- Suitable for MMM research, operations, and academic community use.

Alpha CV: T response to wind ob:



Observation Cost Function (Jo) Implementation

$$J_0\left[\delta\mathbf{x}(t_0)\right] = \frac{1}{2}\sum_{i=0}^{n}\left[\mathbf{H}_i\delta\mathbf{x}(t_i) - \mathbf{d}_i\right]^T \mathbf{R}_i^{-1}\left[\mathbf{H}_i\delta\mathbf{x}(t_i) - \mathbf{d}_i\right]$$

- Observation error correlations increase computational cost of J_o, so…

- Assume observation errors uncorrelated (specified by std. dev. σ_o^2):

$$J_0\left[\delta\mathbf{x}(t_0)\right] = \frac{1}{2}\sum_{i=0}^{n}\left[\mathbf{H}_i\delta\mathbf{x}(t_i) - \mathbf{d}_i\right]^2 / \sigma_{oi}^2$$

- This assumption is not good for satellite retrievals (e.g. temperature, humidity) as retrieval process creates correlated errors.

- Approximation is better for the "raw" radiance observations.

- Use of "super observations" improves assumptions and reduces costs.

Effects of simplifications of cost function

- We started from the full cost-function:

$$J\left[\mathbf{x}(t_0)\right] = \frac{1}{2}\left[\mathbf{x}(t_0) - \mathbf{x}^b(t_0)\right]^T \mathbf{B}_o^{-1}\left[\mathbf{x}(t_0) - \mathbf{x}^b(t_0)\right] + \frac{1}{2}\sum_{i=0}^{n}\left[\mathbf{y}_i - \mathbf{y}_i^o\right]^T \mathbf{R}_i^{-1}\left[\mathbf{y}_i - \mathbf{y}_i^o\right]$$

- We have made various practical assumptions to get the following:

$$J = J_b + J_o = \frac{1}{2}\sum_{n} v_n^2 + \frac{1}{2}\sum_{i=0}^{n}\left[\mathbf{H}_i \delta\mathbf{x}(t_i) - \mathbf{d}_i\right]^2 / \sigma_{oi}^2$$

- This is much cheaper (sums rather than matrix operations).

- Very important to choose accurate background error model $\mathbf{B_o} = \mathbf{U}\mathbf{U}^T$.