Regional 4D-Var

Hans Huang

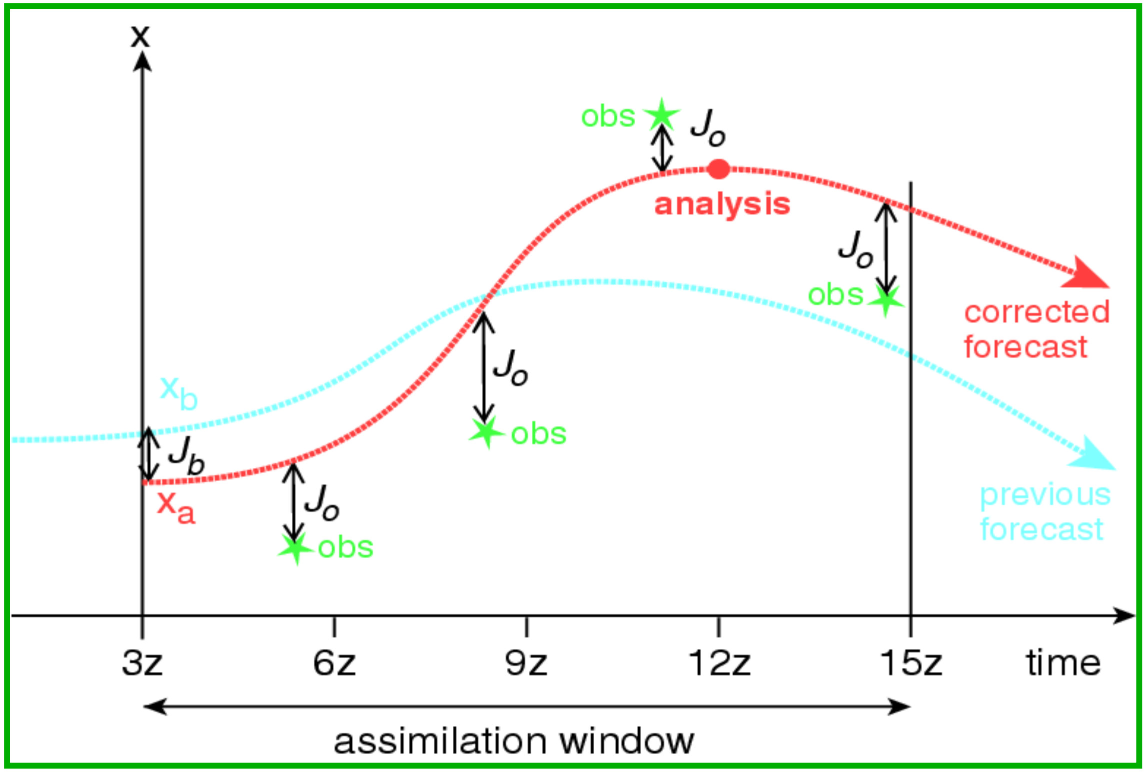
NCAR

Acknowledge: **NCAR/MMM/DAG**
NCAR/RAL/JNT/DATC
USWRP, NSF-OPP, NCAR
AFWA, KMA, CWB, CAA, EUMETSAT, AirDat
Dale Barker (UKMO), Nils Gustafsson (SMHI)

□Outline

- 4D-Var formulation and why regional
- The WRFDA approach
- Regional 4D-Var issues
- Summary

4D-Var



$$J = J_b + J_o$$

$$J_b = \frac{1}{2}\left(\mathbf{x}_0 - \mathbf{x}^b\right)^T \mathbf{B}^{-1}\left(\mathbf{x}_0 - \mathbf{x}^b\right)$$

$$J_o = \frac{1}{2}\sum_k \left[\mathbf{y}_k - H\left(M_k\left(\mathbf{x}_0\right)\right)\right]^T \mathbf{R}^{-1}\left[\mathbf{y}_k - H\left(M_k\left(\mathbf{x}_0\right)\right)\right]$$

Why 4D-Var?

- Use observations over a time interval, which suits most asynoptic data and uses tendency information from observations.
- Use a forecast model as a constraint, which enhances the dynamic balance of the analysis.
- Implicitly use flow-dependent background errors, which ensures the analysis quality for fast developing weather systems.
- NOT easy to build and maintain!

Why regional?

- High resolution model
- Regional observation network
- Radar data and other high resolution observing systems
- Cloud- and precipitation-affected satellite observations
- Model **B** with only regional balance considerations, e.g. tropical **B**
- …

Regional 4D-Var systems

Zupanski M (1993); the Eta model

Zou, et al. (1995); the MM5 model

Sun and Crook (1998); a cloud model

Huang, et al. (2002); the HIRLAM model

Zupanski M, et al. (2005); the RAMS model

Ishikawa, et al. (2005); the JMA mesoscale model

Fillion, et al.; Canadian LAM

Lorenc, et al (2007); UK Met Office UM

…

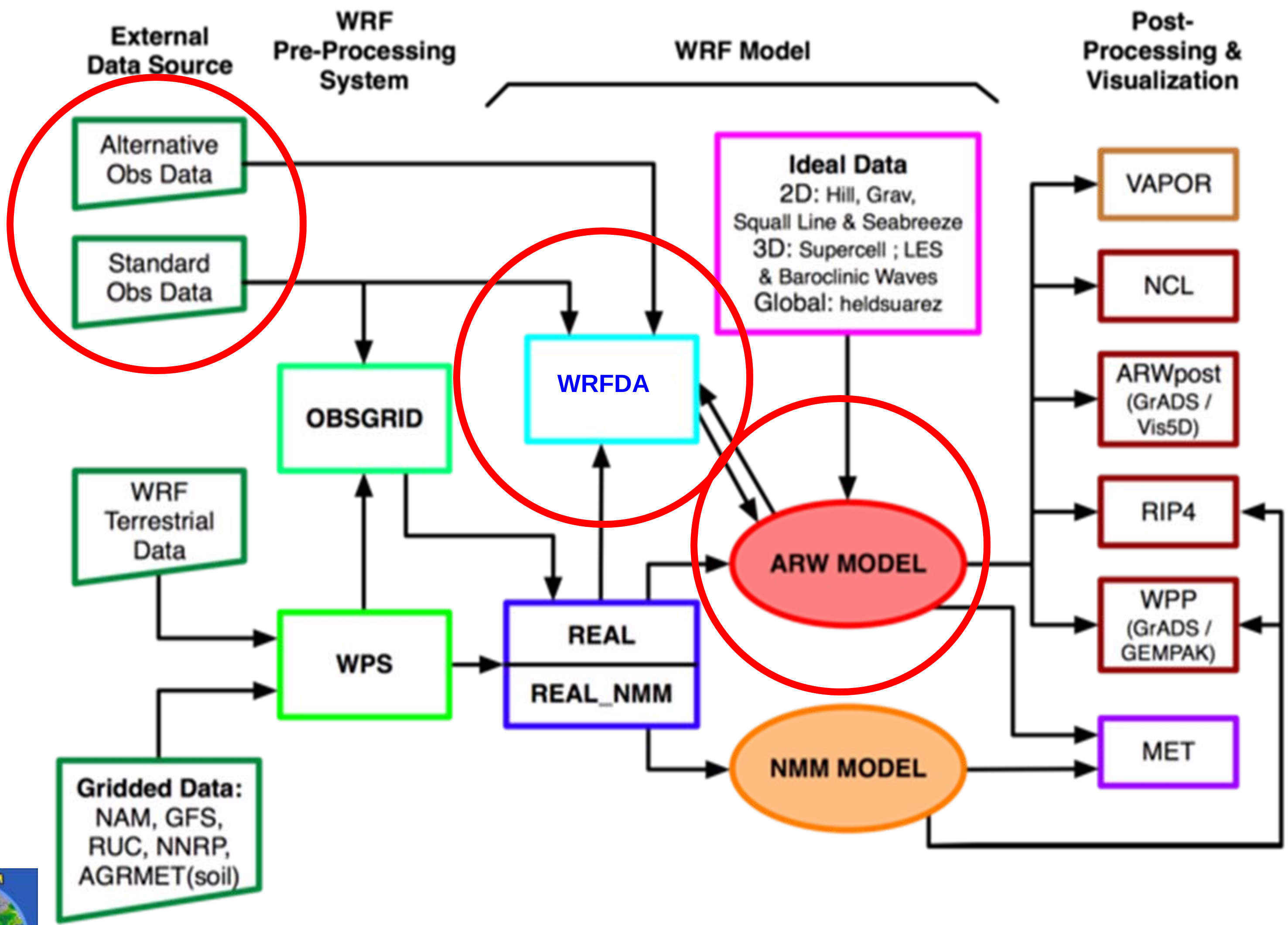
Huang, et al. (2009); the WRF model

□Outline

- 4D-Var formulation and why regional
- The WRFDA approach

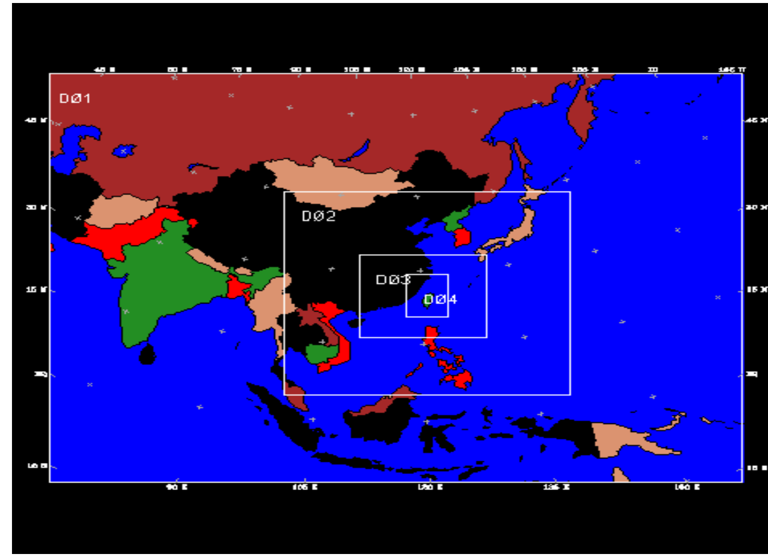
  WRFDA is a **D**ata **A**ssimilation system built within the **WRF** software framework, used for application in both research and operational environments….
- Regional 4D-Var issues
- Summary

WRFDA in WRF Modeling System



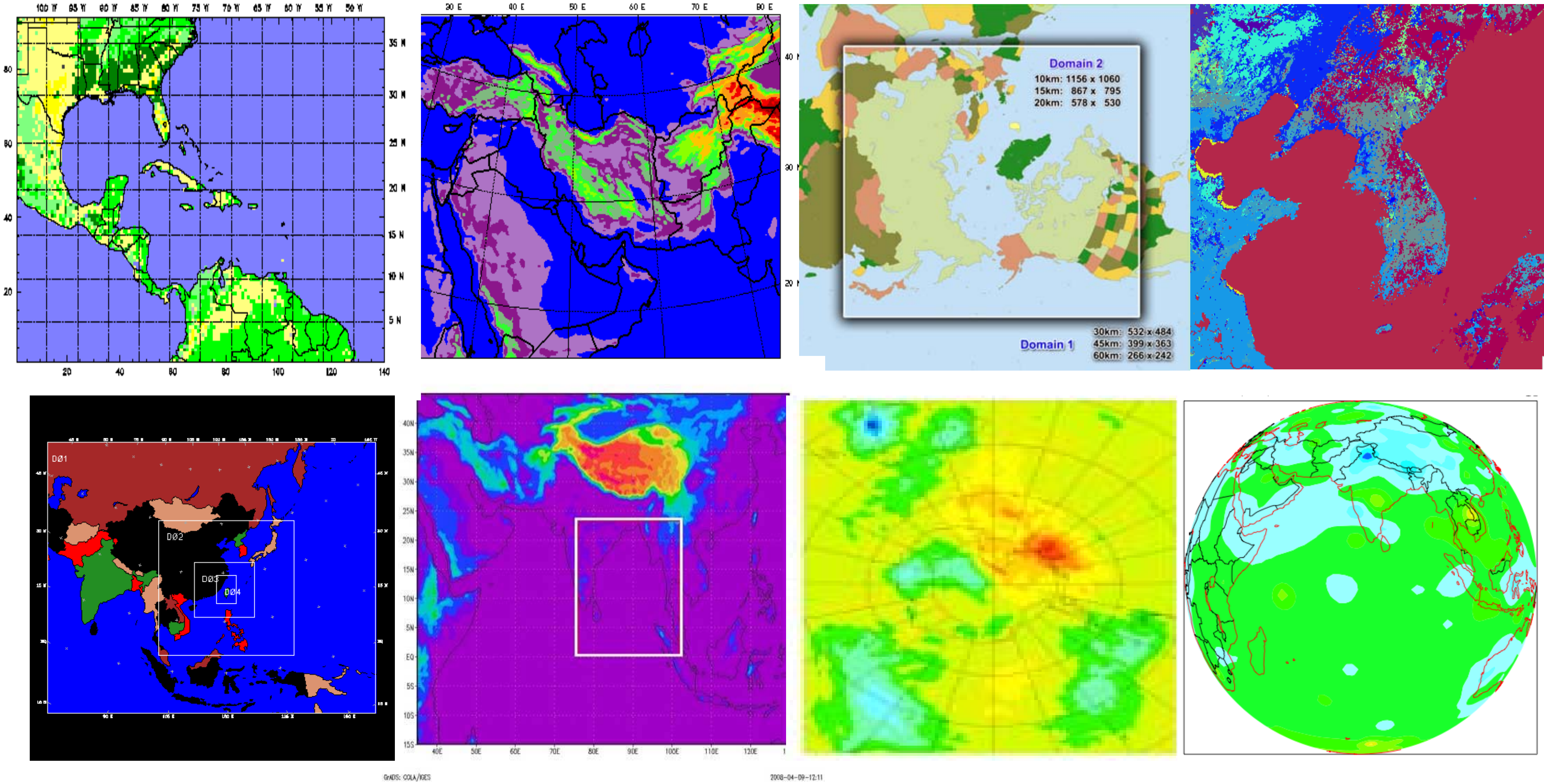
WRFDA

- **Goal:** Community WRF DA system for
  - regional/global,
  - research/operations, and
  - deterministic/probabilistic applications.
- **Techniques:**
  - 3D-Var
  - 4D-Var (regional)
  - Ensemble DA,
  - Hybrid Variational/Ensemble DA.
- **Model:** WRF (ARW, NMM, Global)
- **Support:**
  - NCAR/ESSL/MMM/DAG
  - NCAR/RAL/JNT/DATC
- **Observations:** Conv.+Sat.+Radar (+Bogus)

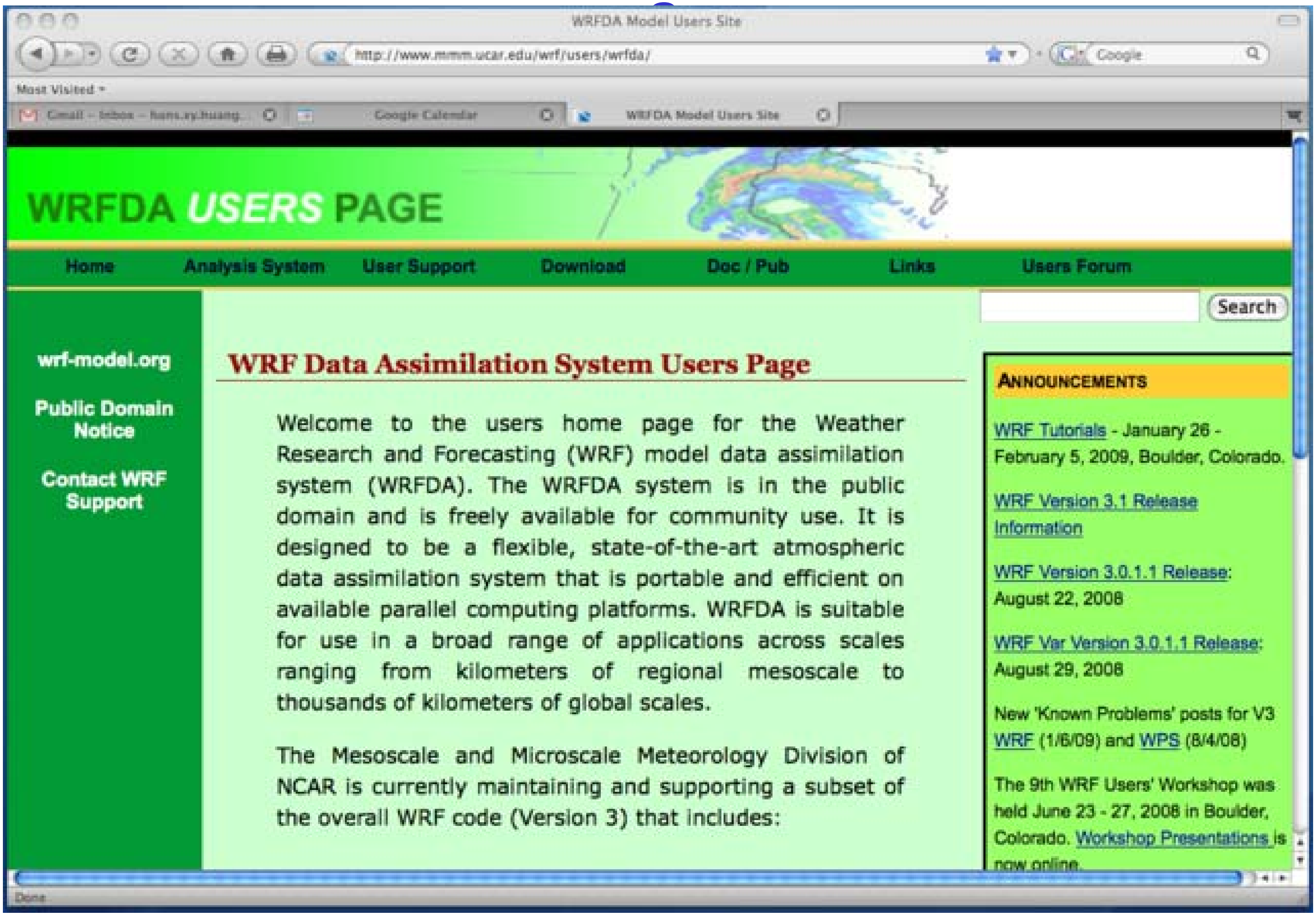


The WRFDA Program

- NCAR staff (DAG,DATC): 20FTE, ~10 projects.
- Non-NCAR collaborators (AFWA, KMA, CWB, BMB, etc): ~10FTE.
- Community users: ~30 (more in 10,000 general WRF downloads?).



www.mmm.ucar.edu/wrf/users/wrfd



WRFDA tutorials

- 21-22 July, 2008. NCAR.
- 10-14 Nov, 2008. CWB, Taiwan.
- 2-4 Feb, 2009. NCAR.
- 17-24 Feb, 2009. Kunming, Yunnan, China.
- 18 April, 2009. South Korea.
- 20-22 July, 2009, NCAR

WRFDA tutorial agenda and presentations

http://www.mmm.ucar.edu/wrf/users/wrfda/tutorial.html

WRFDA online tutorial and user guide

http://www.mmm.ucar.edu/wrf/users/wrfda/Docs/user_guide_V3.1/users_guide_chap6.htm

WRFDA tutorial

Monday - July 20, 2009		
08:30	Welcome and Participants' Introduction	Hans Huang
09:00	WRFDA Overview	Hans Huang
10:30	Observation Pre-Processing	Yong-Run Guo
12:30	WRF-Var System	Michael Duda
01:30	WRF-Var Setup, Run and Diagnostics	Hui Shao
03:00	Practice Session (obsproc, 3D-Var, single-ob tests)	
Tuesday - July 21, 2009		
08:30	WRF-Var Background Error Estimations	Rizvi Syed
09:30	Radar Data Assimilation	Qingnong Xiao
10:45	Radiance Data Assimilation	Tom Auligne/Zhiquan Liu
12:30	WRF 4D-Var	Xin Zhang
03:10	Hybrid Data Assimilation System	Meral Demirtas
02:30	Practice Session (gen_be, radar, radiance, 4D-Var, hybrid)	
Wednesday - July 22, 2009		
08:30	WRF-Var Tools and Verification Package	Rizvi Syed
09:00	Ensemble	Chris Snyder
09:20	GSI	Ming Hu
10:50	Optional Practice Session (advanced practice)	

The analysis problem for a given time

Consider a scalar x.

The background (normally a short-range forecast):

$$x^b = x^t + b.$$

The observation:

$$x^r = x^t + r.$$

The error statistics are assumed to be known:

$< b > = 0$, mean error (unbiased),

$< r > = 0$, mean error (unbiased),

$< b^2 > = B$, background error variance,

$< r^2 > = R$, observation error variance,

$< br > = 0$, no correlation between b and r,

where $<\cdot>$ is ensemble average.

The analysis: $x^a = x^b + \frac{B}{B+R}\left(x^r - x^b\right)$

$0 < \frac{B}{B+R} < 1$	The analysis value should be between b ackground and observation.
$\lim_{B\to 0} x^a = x^b$	If B is too small, observations are less u seful.
$\lim_{R\to 0} x^a = x^r$	If R can be tuned, analysis can fit observations as close as one wants!

The analysis error variance: $A^{-1} = B^{-1} + R^{-1}$

$A < B$	Statistically, analyses are better than background.
$A < R$	Statistically, analyses are better than observations!

Sequential data assimilation

EKF

$$\mathbf{K}_i = \mathbf{P}_i^f \mathbf{H}_i^T \left(\mathbf{H}_i \mathbf{P}_i^f \mathbf{H}_i^T + \mathbf{R}\right)^{-1}$$

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}_i \left[\mathbf{y} - H\left(\mathbf{x}_i^f\right)\right]$$

$$\mathbf{P}_i^a = \left(\mathrm{I} - \mathbf{K}_i \mathbf{H}_i\right) \mathbf{P}_i^f$$

$$\mathbf{x}_{i+1}^f = M\left(\mathbf{x}_i^a\right)$$

$$\mathbf{P}_{i+1}^f = \mathbf{M}_i \mathbf{P}_i^a \mathbf{M}_i^T + \mathbf{Q}_i$$

$\left(\mathbf{B} = \mathbf{P}_{i+1}^f\right)$ EKF $\rightarrow$ OI or VAR

$$\left(\mathbf{K} = \mathbf{B}\mathbf{H}^T \left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right)^{-1}\right)$$

(Ensemble KF use ensembles to calculate P^a and P^f)

OI or VAR

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}\left[\mathbf{y} - H\left(\mathbf{x}_i^f\right)\right]$$

$$\mathbf{x}_{i+1}^f = M\left(\mathbf{x}_i^a\right)$$

A short list of 4D-Var issues

$$J = \frac{1}{2}\left(\mathbf{x}_0 - \mathbf{x}^b\right)^T \mathbf{B}^{-1}\left(\mathbf{x}_0 - \mathbf{x}^b\right) + \frac{1}{2}\sum_k \left[\mathbf{y}_k - H\left(M_k\left(\mathbf{x}_0\right)\right)\right]^T \mathbf{R}^{-1}\left[\mathbf{y}_k - H\left(M_k\left(\mathbf{x}_0\right)\right)\right]$$

□$\mathbf{y}$ observations, preprocessing, quality control

□H new observation types, improving operators

□$\mathbf{R}$ observational errors, tuning, correlated observational errors

□$\mathbf{x}^b$ improve models (an important part of DA!!)

□$\mathbf{B}$ Statistical models, flow-dependent features

□$\mathbf{M}$ and $\mathbf{M}^T$:Źdevelopment and validity (and ensemble approaches)

□Minimization algorithm (Quasi−Newton; Conjugate Gradient; …)

(WRFDA) Observations, y

- **In-Situ:**
  - Surface (SYNOP, METAR, SHIP, BUOY).
  - Upper air (TEMP, PIBAL, AIREP, ACARS, TAMDAR).
- **Remotely sensed retrievals:**
  - Atmospheric Motion Vectors (geo/polar).
  - SATEM thickness.
  - Ground-based GPS Total Precipitable Water/Zenith Total Delay.
  - SSM/I oceanic surface wind speed and TPW.
  - Scatterometer oceanic surface winds.
  - Wind Profiler.
  - Radar radial velocities and reflectivities.
  - Satellite temperature/humidity/thickness profiles.
  - GPS refractivity (e.g. COSMIC).
- **Radiative Transfer (RTTOV or CRTM):**
  - HIRS from NOAA-16, NOAA-17, NOAA-18, METOP-2
  - AMSU-A from NOAA-15, NOAA-16, NOAA-18, EOS-Aqua, METOP-2
  - AMSU-B from NOAA-15, NOAA-16, NOAA-17
  - MHS from NOAA-18, METOP-2
  - AIRS from EOS-Aqua
  - SSMIS from DMSP-16

- **Bogus:**
  - TC bogus.
  - Global bogus.

H - Observation operator

H maps variables from "model space" to "observation space"

$$\mathbf{x} \Longrightarrow \mathbf{y}$$

- Interpolations from model grids to observation locations
- Extrapolations using PBL schemes
- Time integration using full NWP models (4D-Var in generalized form)
- Transformations of model variables (u, v, T, q, p_s, etc.) to "indirect" observations (e.g. satellite radiance, radar radial winds, etc.)
  - Simple relations like PW, radial wind, refractivity, …
  - Radar reflectivity $Z = Z(T, IWC, LWC, RWC, SWC)$
  - Radiative transfer models $L(\nu) \approx \int_0^\infty B(\nu, T(z)) \left[\frac{dTR(\nu)}{dz} \right] dz$
  - Precipitation using simple or complex models
  - …

!!! Need H, $\mathbf{H}$ and $\mathbf{H}^T$, not $\mathbf{H}^{-1}$!!!

A short list of 4D-Var issues

$$J = \frac{1}{2}\left(\mathbf{x}_0 - \mathbf{x}^b\right)^T \mathbf{B}^{-1}\left(\mathbf{x}_0 - \mathbf{x}^b\right) + \frac{1}{2}\sum_k \left[\mathbf{y}_k - H\left(M_k\left(\mathbf{x}_0\right)\right)\right]^T \mathbf{R}^{-1}\left[\mathbf{y}_k - H\left(M_k\left(\mathbf{x}_0\right)\right)\right]$$

- $\mathbf{y}$ observations, preprocessing, quality control
- H new observation types, improving operators
- $\mathbf{R}$ observational errors, tuning, correlated observational errors
- $\mathbf{x}^b$ improve models (an important part of DA!!)
- $\mathbf{B}$ Statistical models, flow-dependent features
- $\mathbf{M}$ and $\mathbf{M}^T$:Źdevelopment and validity (and ensemble approaches)
- Minimization algorithm (Quasi–Newton; Conjugate Gradient; …)

Model-Based Estimation of Climatological Background Errors

- Assume background error covariance estimated by model perturbations *x'* :

$$\mathbf{B} = \overline{(\boldsymbol{x}^b - \boldsymbol{x}^t)(\boldsymbol{x}^b - \boldsymbol{x}^t)^T} \approx \overline{\boldsymbol{x}'\boldsymbol{x}'^T}$$

Two ways of defining ***x'***:

- The NMC-method (Parrish and Derber 1992):

$$\mathbf{B} = \overline{\boldsymbol{x}'\boldsymbol{x}'^T} \approx A\overline{(\boldsymbol{x}^{t2} - \boldsymbol{x}^{t1})(\boldsymbol{x}^{t2} - \boldsymbol{x}^{t1})^T}$$

  where e.g. t2=24hr, t1=12hr forecasts…

- …or ensemble perturbations (Fisher 2003):

$$\mathbf{B} = \overline{\boldsymbol{x}'\boldsymbol{x}'^T} \approx C\overline{(\boldsymbol{x}^k - \langle \boldsymbol{x} \rangle)(\boldsymbol{x}^k - \langle \boldsymbol{x} \rangle)^T}$$

- Tuning via innovation vector statistics and/or variational methods.

Single observation experiment - one way to view the structure of **B**

The solution of 3D-Var should be

$$\boldsymbol{x}^a = \boldsymbol{x}^b + \mathbf{BH}^T\left[\mathbf{HBH}^T + \mathbf{R}\right]^{-1}\left[\boldsymbol{y} - H\boldsymbol{x}^b\right]$$

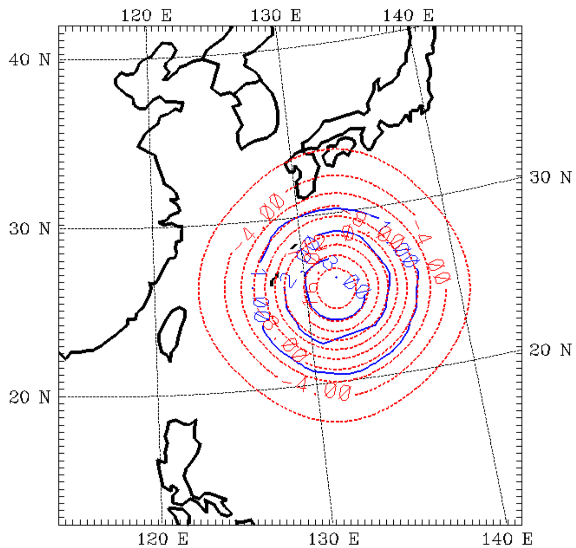
Single observation

$$\underline{\boldsymbol{x}^a - \boldsymbol{x}^b = \mathbf{B}}_i\left[\sigma_b^2 + \sigma_o^2\right]^{-1}\left[\boldsymbol{y}_i - x_i\right]$$

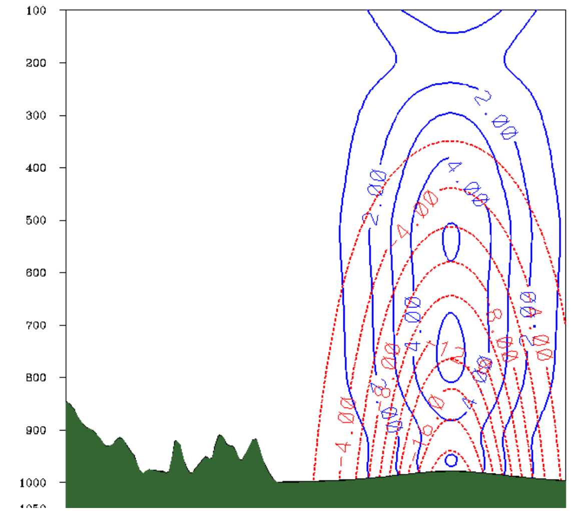
WRF DA

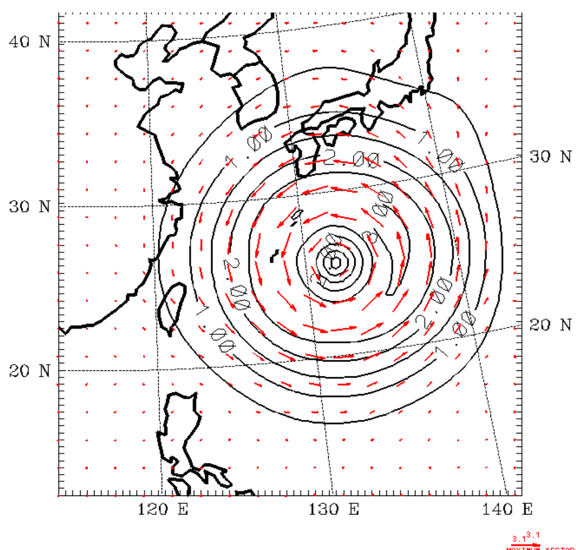
Example of B

3D-Var response to a single p_s observation

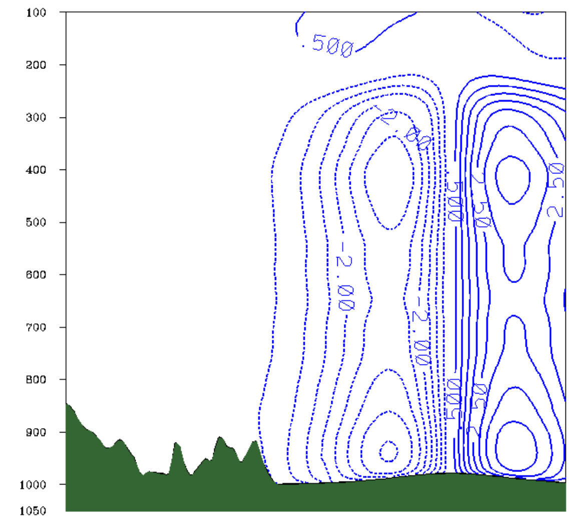


Pressure,
Temperature





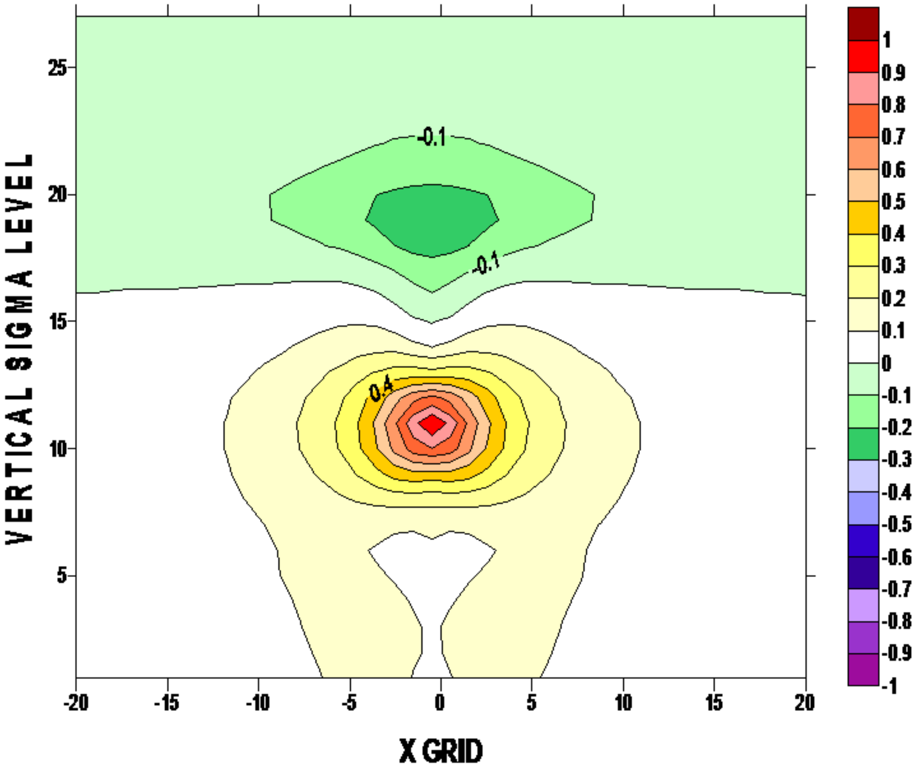
Wind Speed,
Vector,
v-wind component.



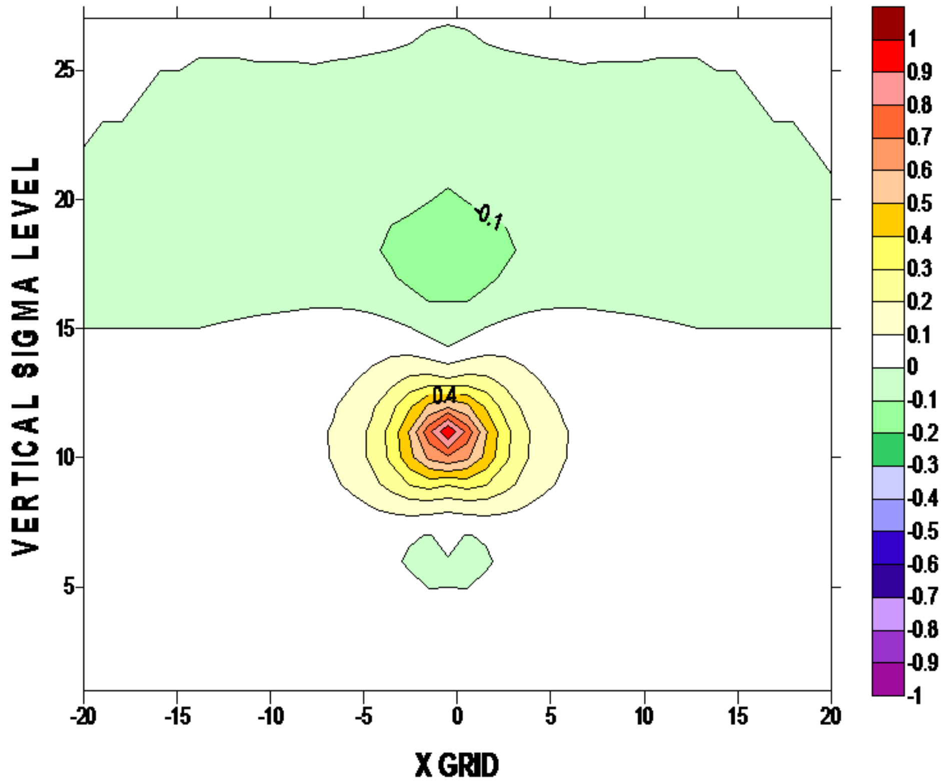
Examples of B

T increments : T Observation (1 Deg , 0.001 error around 850 hPa)

B from NMC method



B from ensemble method



WRF DA

Incremental WRFDA 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.

WRFDA Background Error Modeling

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

U_p: Change of variable, impose balance.

U_v: Vertical correlations EOF Decomposition

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

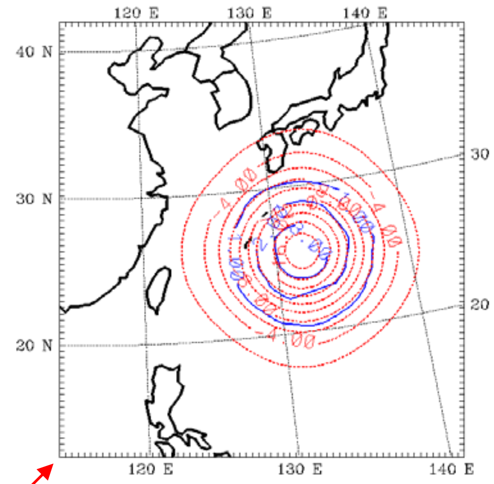
WRFDA Background Error Modeling

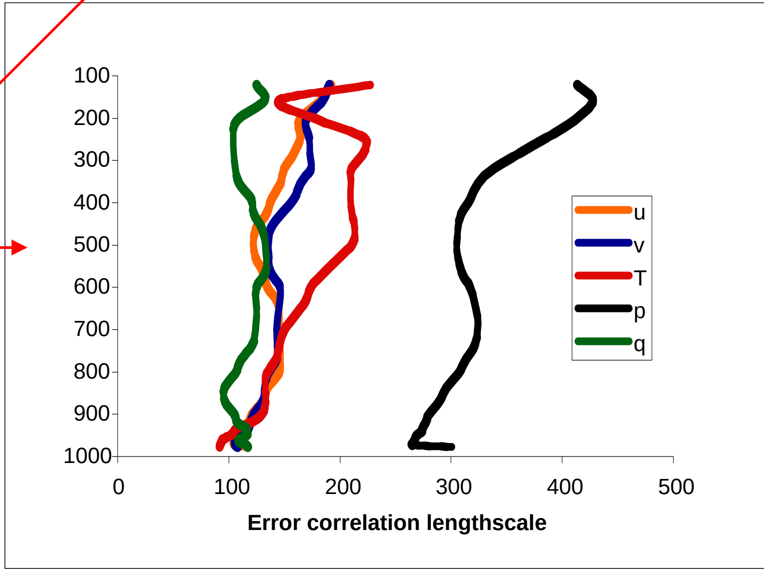
$$\delta\mathbf{x}(t_0) = \mathbf{U}\mathbf{v} = \mathbf{U}_p\mathbf{U}_v\mathbf{U}_h\mathbf{v}$$

$\mathbf{U}_p$: Change of variable, impose balance.

$\mathbf{U}_v$: Vertical correlations EOF Decomposition

$\mathbf{U}_h$: RF = Recursive Filter, e.g. Purser et al 2003



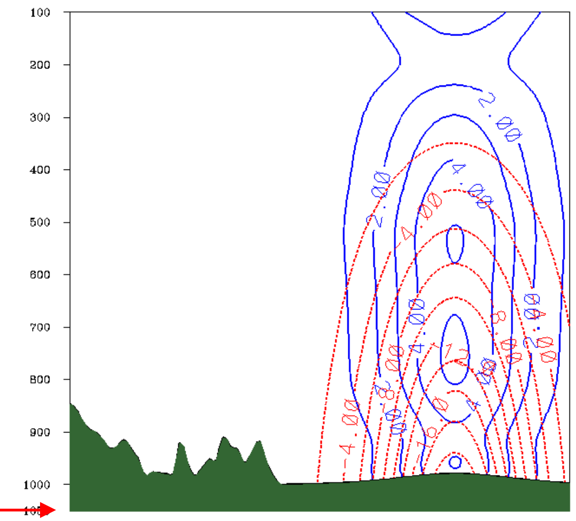


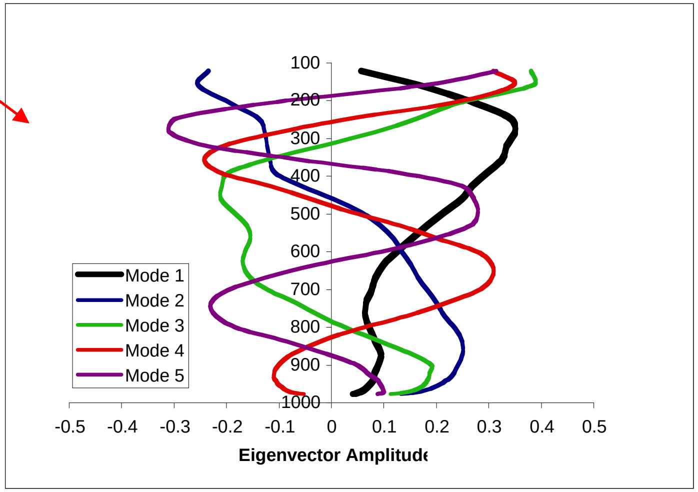
WRFDA Background Error Modeling

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

U_p: Change of variable, impose balance.

U_v: Vertical correlations EOF Decomposition





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

WRFDA Background Error Modeling

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

U_p: Change of variable, impose balance.

U_v: Vertical correlations EOF Decomposition

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

Define control variables:

$$\psi'$$

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

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

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

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

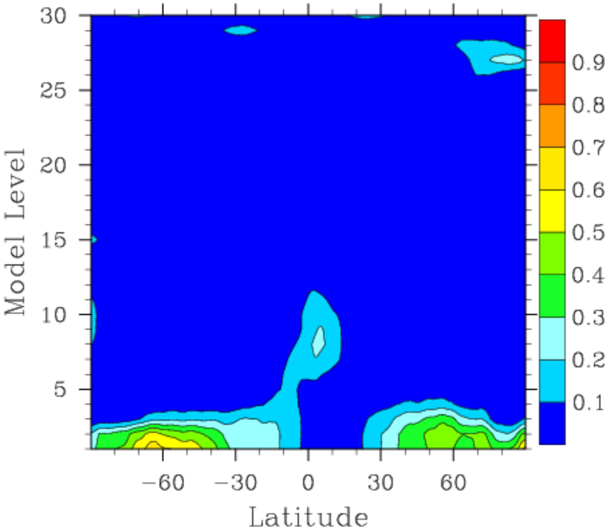
WRFDA Statistical Balance Constraints

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

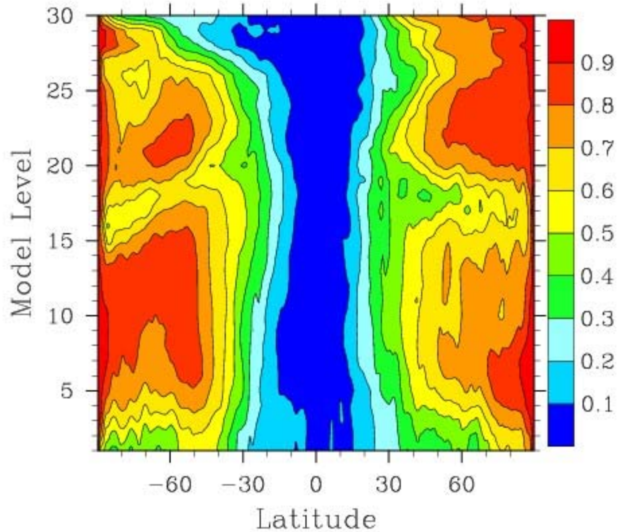
$$\chi_b^{'} = c\,\psi^{'} \quad T_b^{'}(k) = \sum_{k1} G(k,k1)\,\psi^{'}(k1) \quad p_{sb}^{'} = \sum_{k} W(k)\,\psi^{'}(k)$$

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

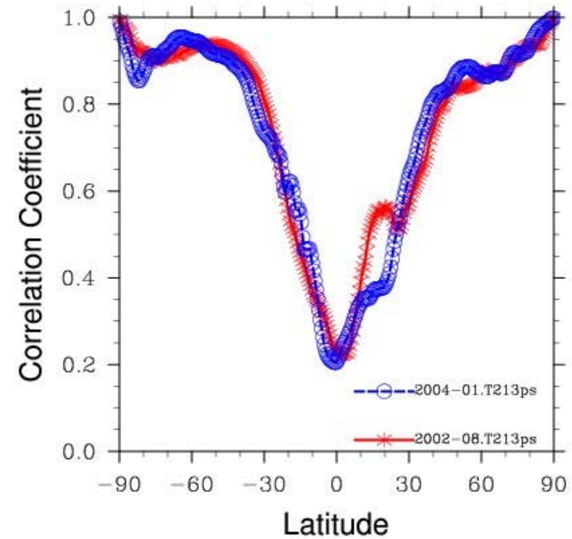
$\langle \chi_b \bullet \chi \rangle / \langle \chi \bullet \chi \rangle$



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



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



A short list of 4D-Var issues

$$J = \frac{1}{2}\left(\mathbf{x}_0 - \mathbf{x}^b\right)^T \mathbf{B}^{-1}\left(\mathbf{x}_0 - \mathbf{x}^b\right) + \frac{1}{2}\sum_k \left[\mathbf{y}_k - H\left(M_k\left(\mathbf{x}_0\right)\right)\right]^T \mathbf{R}^{-1}\left[\mathbf{y}_k - H\left(M_k\left(\mathbf{x}_0\right)\right)\right]$$

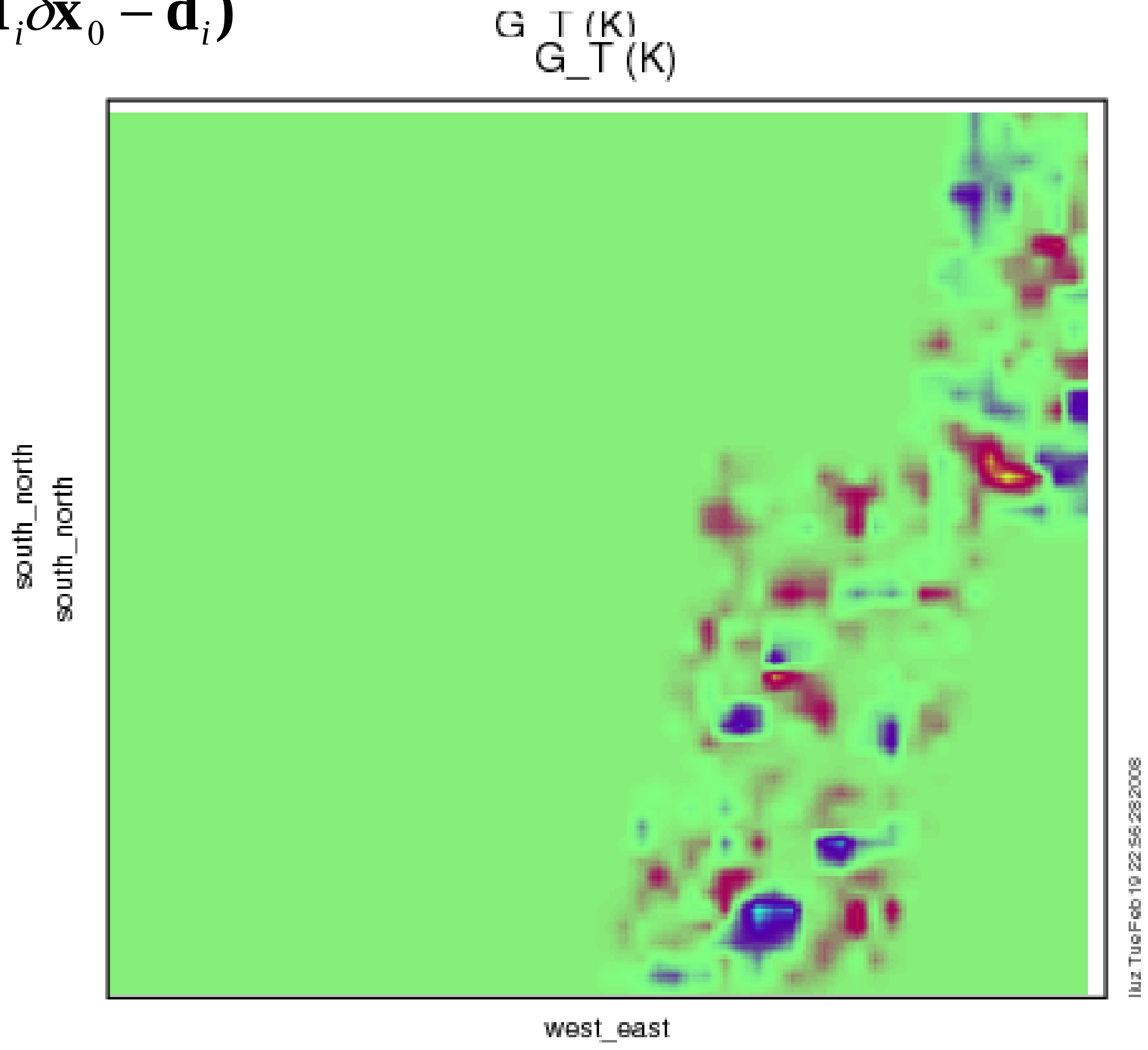
- $\mathbf{y}$ observations, preprocessing, quality control
- H new observation types, improving operators
- $\mathbf{R}$ observational errors, tuning, correlated observational errors
- $\mathbf{x}^b$ improve models (an important part of DA!!)
- $\mathbf{B}$ Statistical models, flow-dependent features
- $\mathbf{M}$ and $\mathbf{M}^T$:Źdevelopment and validity (and ensemble approaches)
- Minimization algorithm $\left(\text{Quasi}-\text{Newton; Conjugate Gradient; …}\right)$

4D-Var

- Use observations over a time interval, which suits most asynoptic data and use tendency information from observations.
- Use a forecast model as a constraint, which enhances the dynamic balance of the analysis.
- Implicitly use flow-dependent background errors, which ensures the analysis quality for fast developing weather systems.
- NOT easy to build and maintain!

Radiance Observation Forcing at 7 data slots

$$\mathbf{H}_i^{\mathrm{T}}\mathbf{R}_i^{-1}(\mathbf{H}_i\mathbf{M}_i\delta\mathbf{x}_0 - \mathbf{d}_i)$$



Why 4D-Var?

- Use observations over a time interval, which suits most asynoptic data and use tendency information from observations.
- Use a forecast model as a constraint, which enhances the dynamic balance of the analysis.
- Implicitly use flow-dependent background errors, which ensures the analysis quality for fast developing weather systems.
- NOT easy to build and maintain!

WRF DA

Single observation experiment

The idea behind single ob tests:

The solution of 3D-Var should be

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{B}\mathbf{H}^T\left[\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right]^{-1}\left[\mathbf{y} - H\mathbf{x}^b\right]$$

Single observation

$$\underline{\mathbf{x}^a - \mathbf{x}^b} = \mathbf{B}_i\left[\sigma_b^2 + \sigma_o^2\right]^{-1}\left[\mathbf{y}_i - x_i\right]$$

3D-Var → 4D-Var: $H \rightarrow HM$; $\mathbf{H} \rightarrow \mathbf{HM}$; $\mathbf{H}^T \rightarrow \mathbf{M}^T\mathbf{H}^T$

The solution of 4D-Var should be

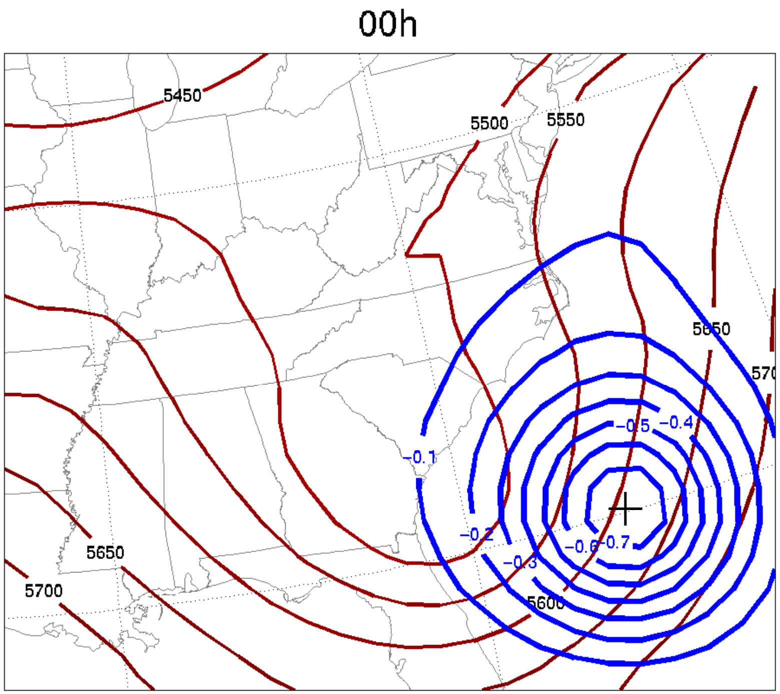
$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{B}\mathbf{M}^T\mathbf{H}^T\left[\mathbf{H}\left(\mathbf{M}\mathbf{B}\mathbf{M}^T\right)\mathbf{H}^T + \mathbf{R}\right]^{-1}\left[\mathbf{y} - HM\mathbf{x}^b\right]$$

Single observation, solution at observation time

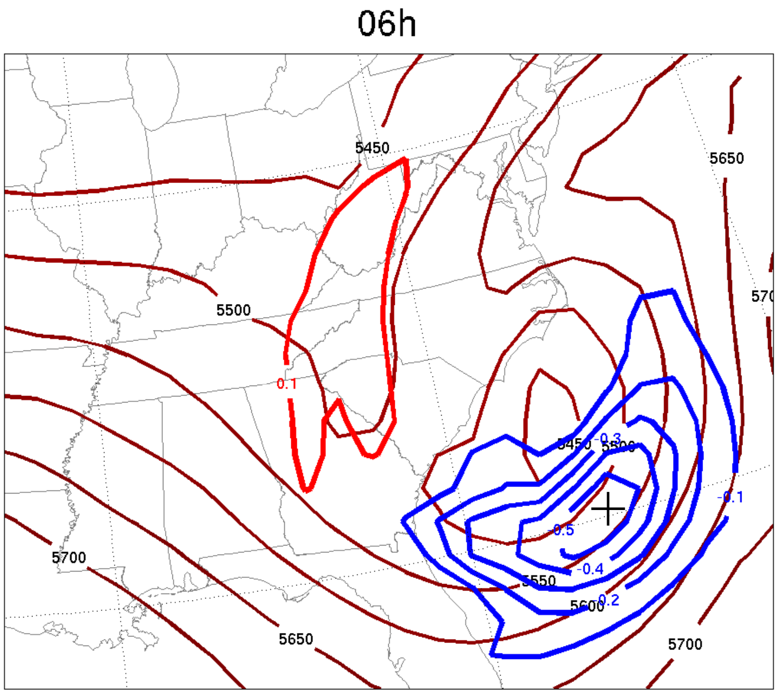
$$\underline{\mathbf{M}\left(\mathbf{x}^a - \mathbf{x}^b\right) = \left(\mathbf{M}\mathbf{B}\mathbf{M}^T\right)_i}\left[\sigma_b^2 + \sigma_o^2\right]^{-1}\left[\mathbf{y}_i - x_i\right]$$

Analysis increments of 500mb θ

from 3D-Var at 00h and from 4D-Var at 06h
due to a 500mb T observation at 06h

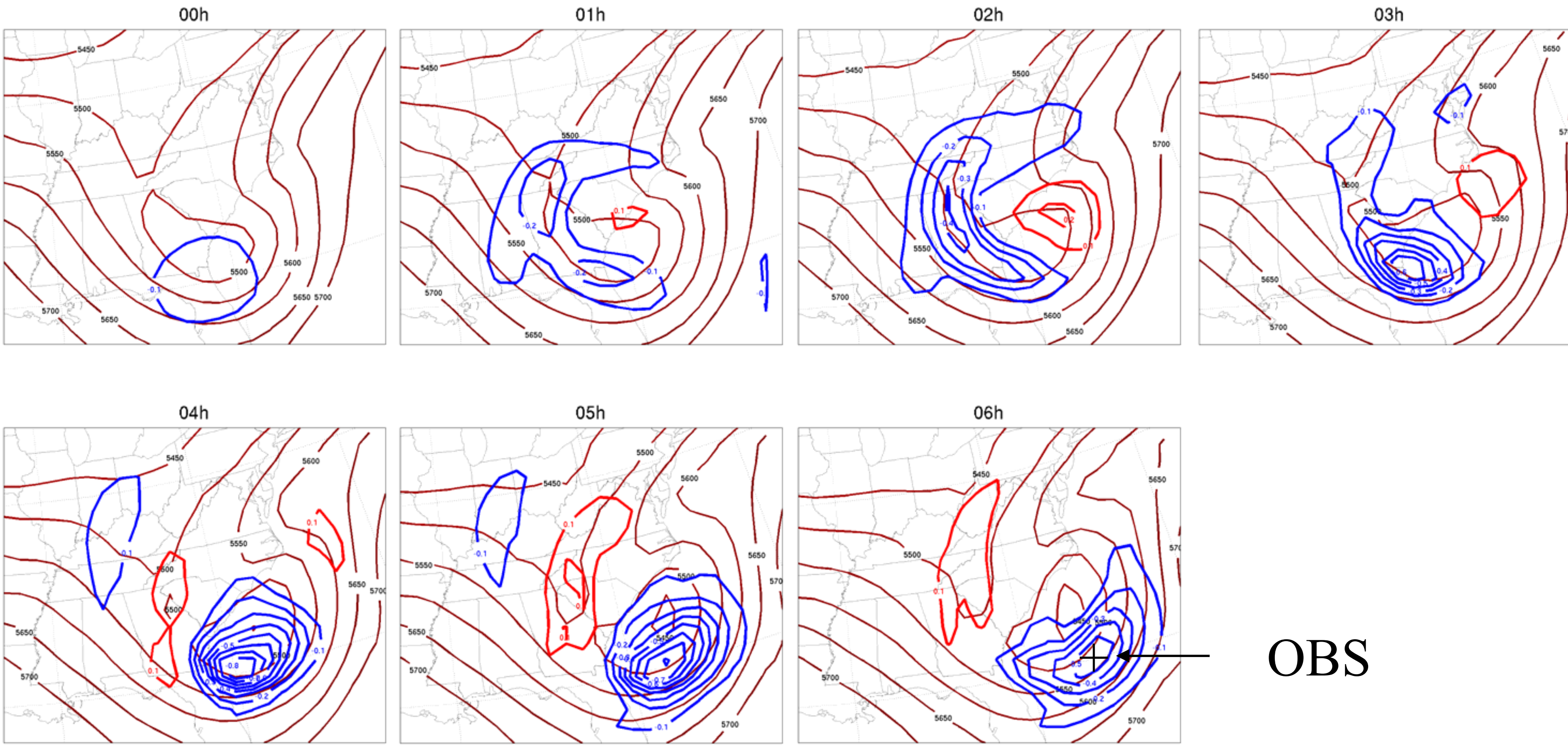


3D-Var

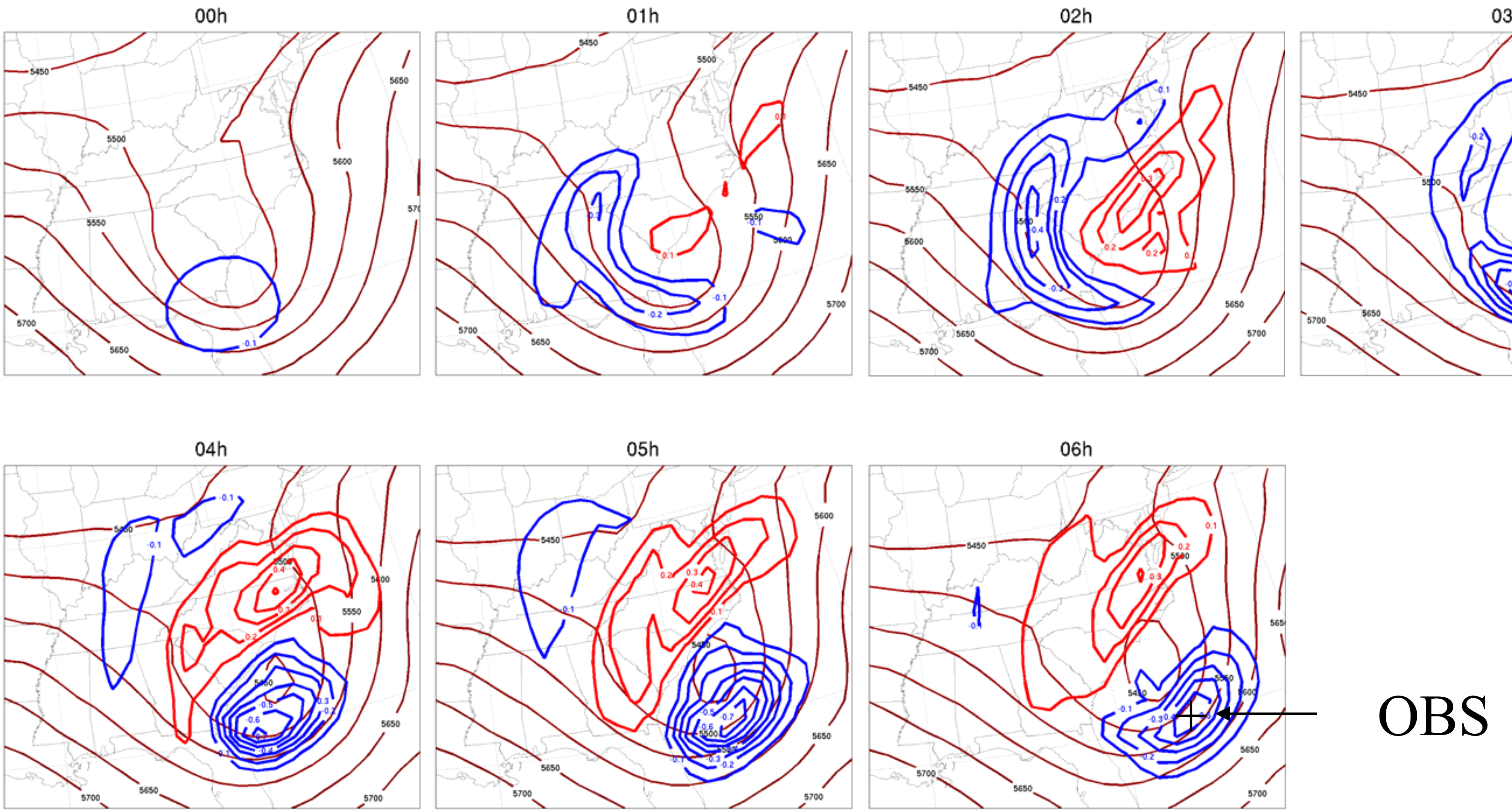


4D-Var

500mb θ increments at 00,01,02,03,04,05,06h to a 500mb T ob at 06h

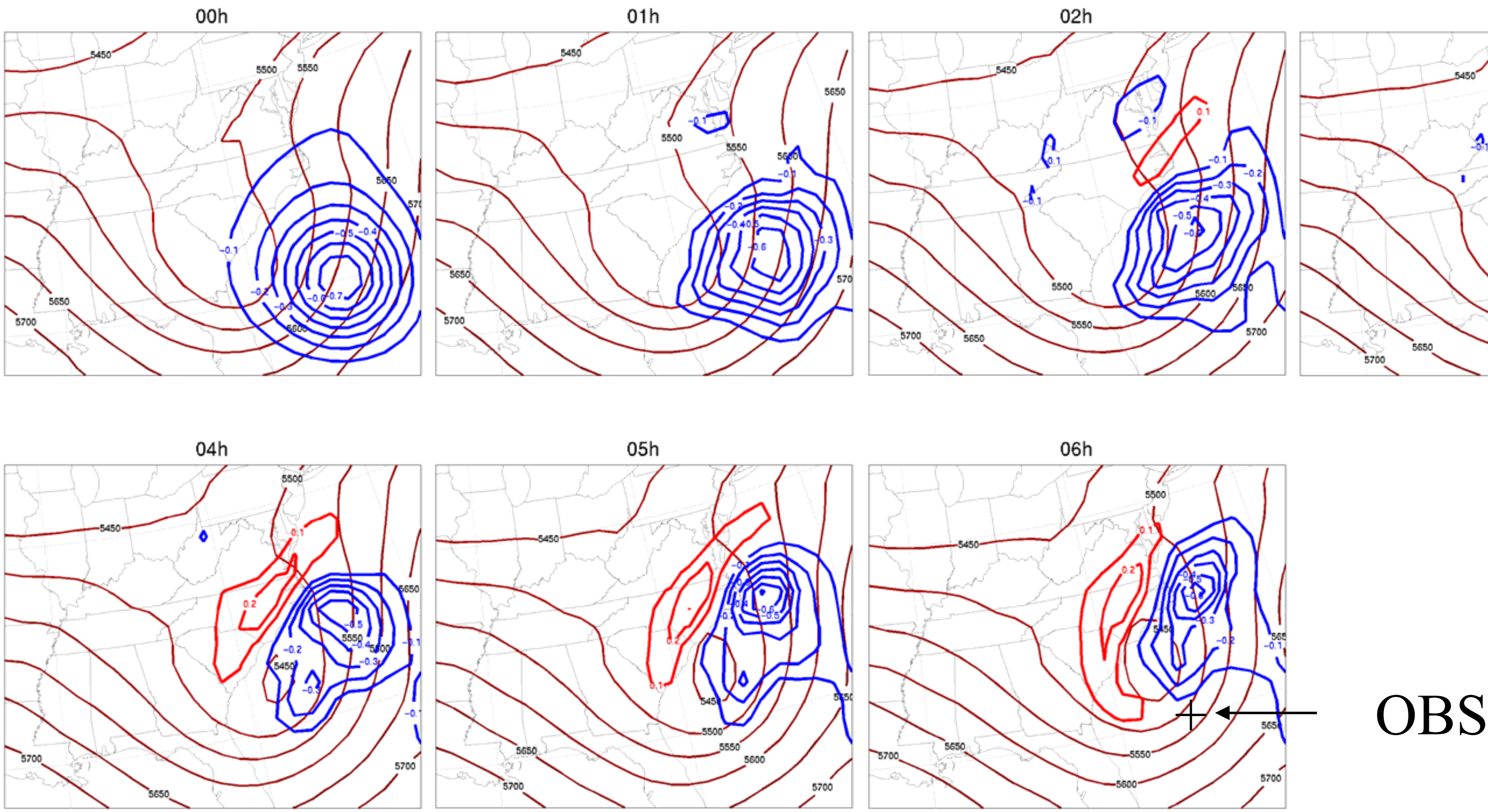


500mb θ difference at 00,01,02,03,04,05,06h from two nonlinear runs (one from background; one from 4D-Var)



WRF DA

500mb θ difference at 00,01,02,03,04,05,06h from two nonlinear runs (one from background; one from FGAT)



Control of noise: J_c

$$J_c(\mathbf{x}_0)=\frac{\gamma_{df}}{2}\left[\left(\delta\mathbf{x}_{N/2}-\delta\mathbf{x}_{N/2}^{df}\right)^T\mathbf{C}^{-1}\left(\delta\mathbf{x}_{N/2}-\delta\mathbf{x}_{N/2}^{df}\right)\right]$$

$$=\frac{\gamma_{df}}{2}\left[\left(\delta\mathbf{x}_{N/2}-\sum_{i=0}^{N}f_i\delta\mathbf{x}_i\right)^T\mathbf{C}^{-1}\left(\delta\mathbf{x}_{N/2}-\sum_{i=0}^{N}f_i\delta\mathbf{x}_i\right)\right]$$

$$=\frac{\gamma_{df}}{2}\left[\left(\sum_{i=0}^{N}h_i\delta\mathbf{x}_i\right)^T\mathbf{C}^{-1}\left(\sum_{i=0}^{N}h_i\delta\mathbf{x}_i\right)\right]$$

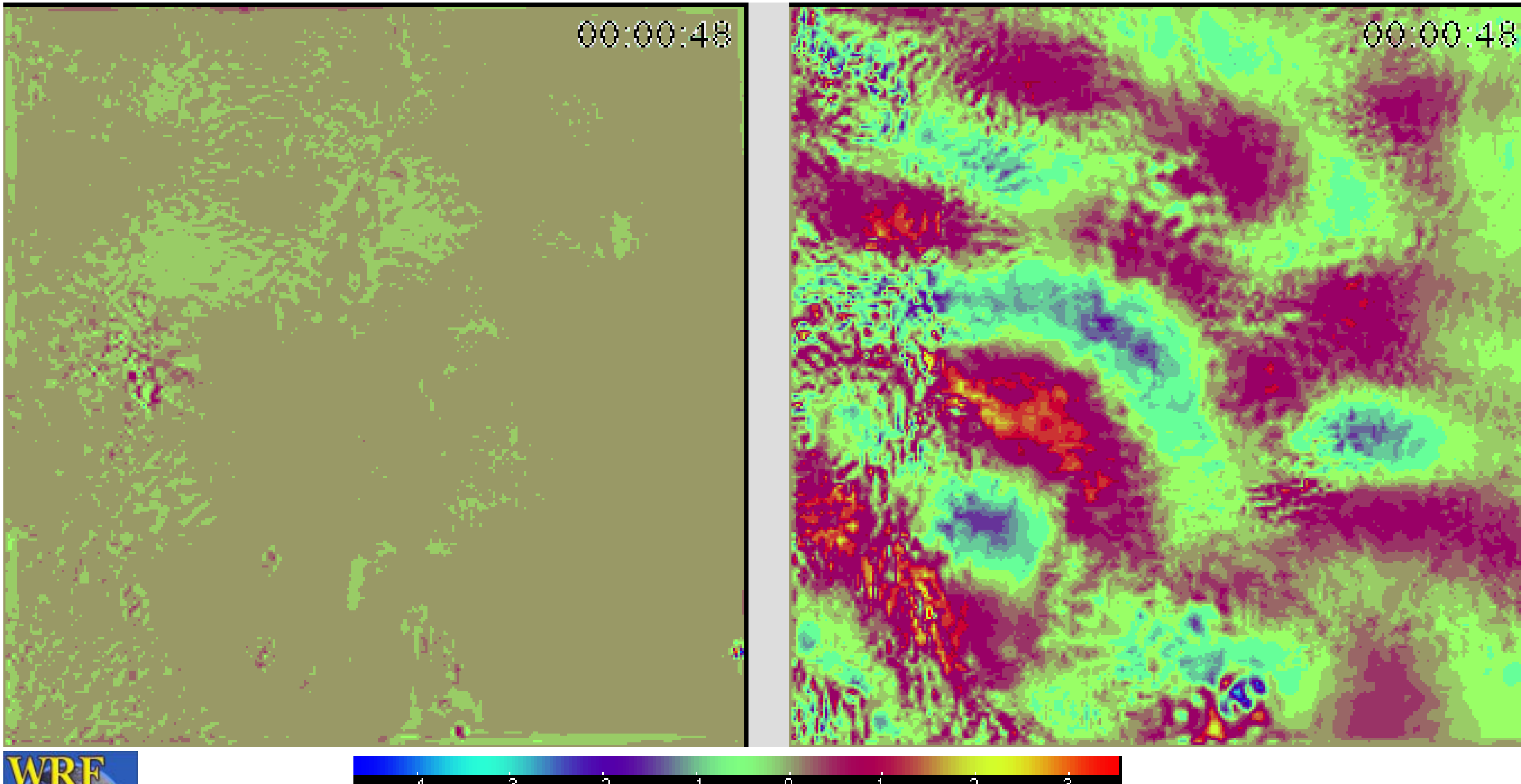
where:

$$h_i=\begin{cases}-f_i, & if\ i\neq N/2\\ 1-f_i, & if\ i=N/2\end{cases}$$

(Dry) Surface Pressure Tendency

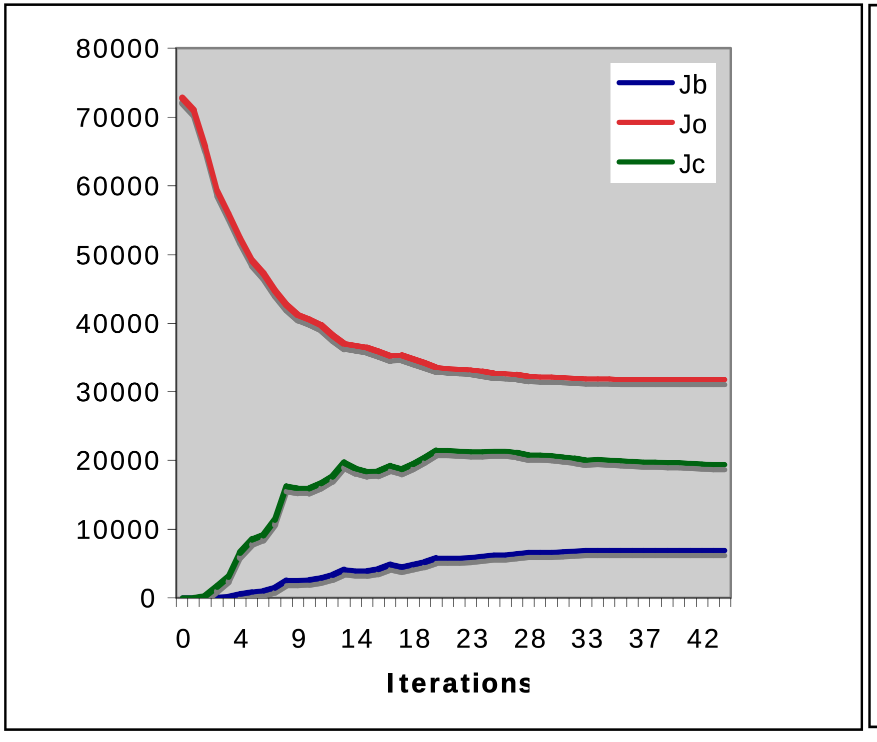
DFI

No DFI



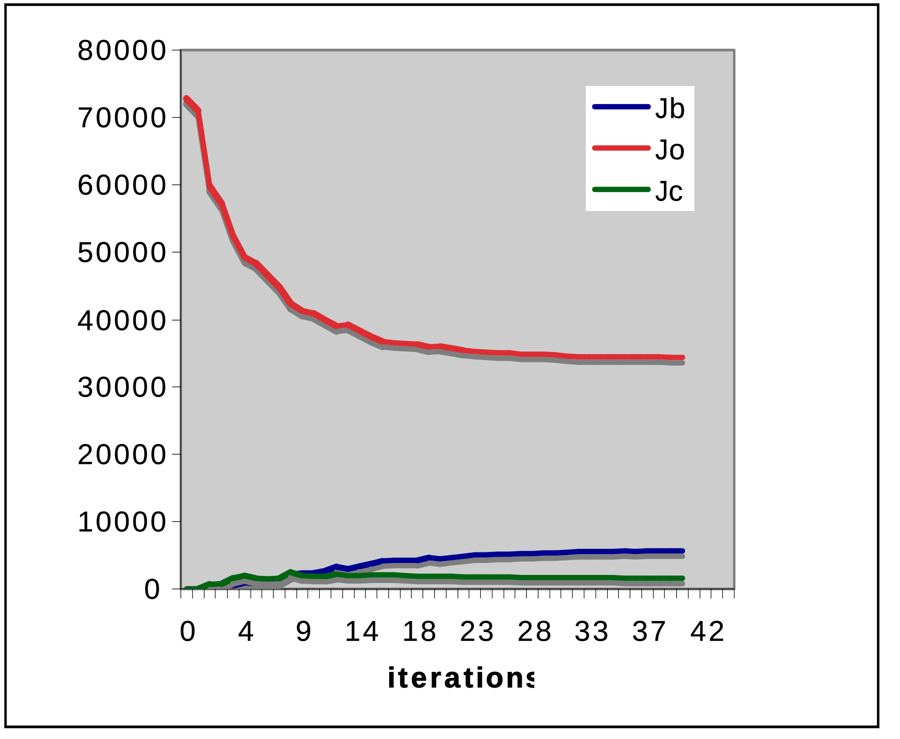
Cost functions when minimize

$J = J_b + J_o$



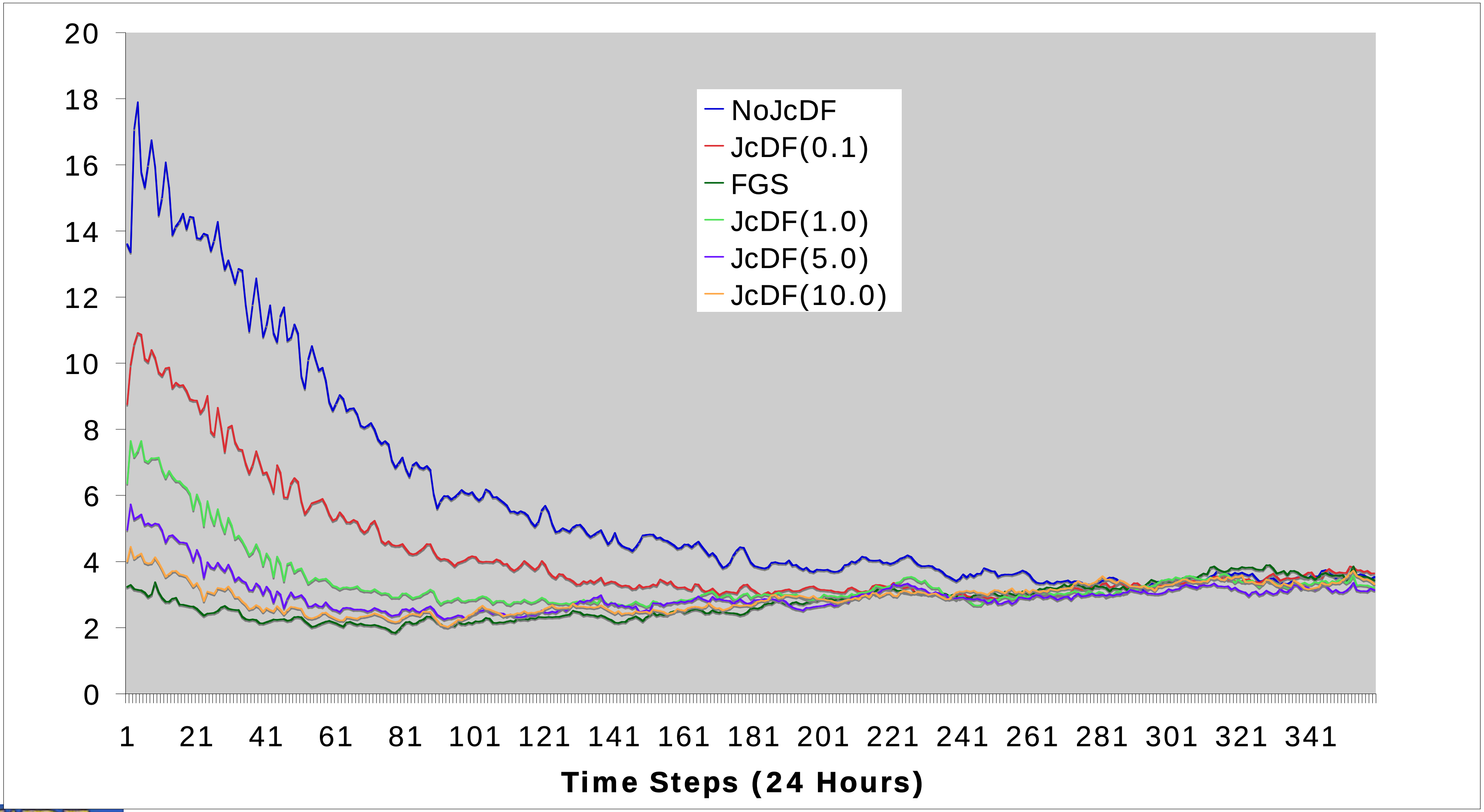
Cost functions when minimize

$J = J_b + J_o + J_c$



$\gamma_{df} = 0.1$

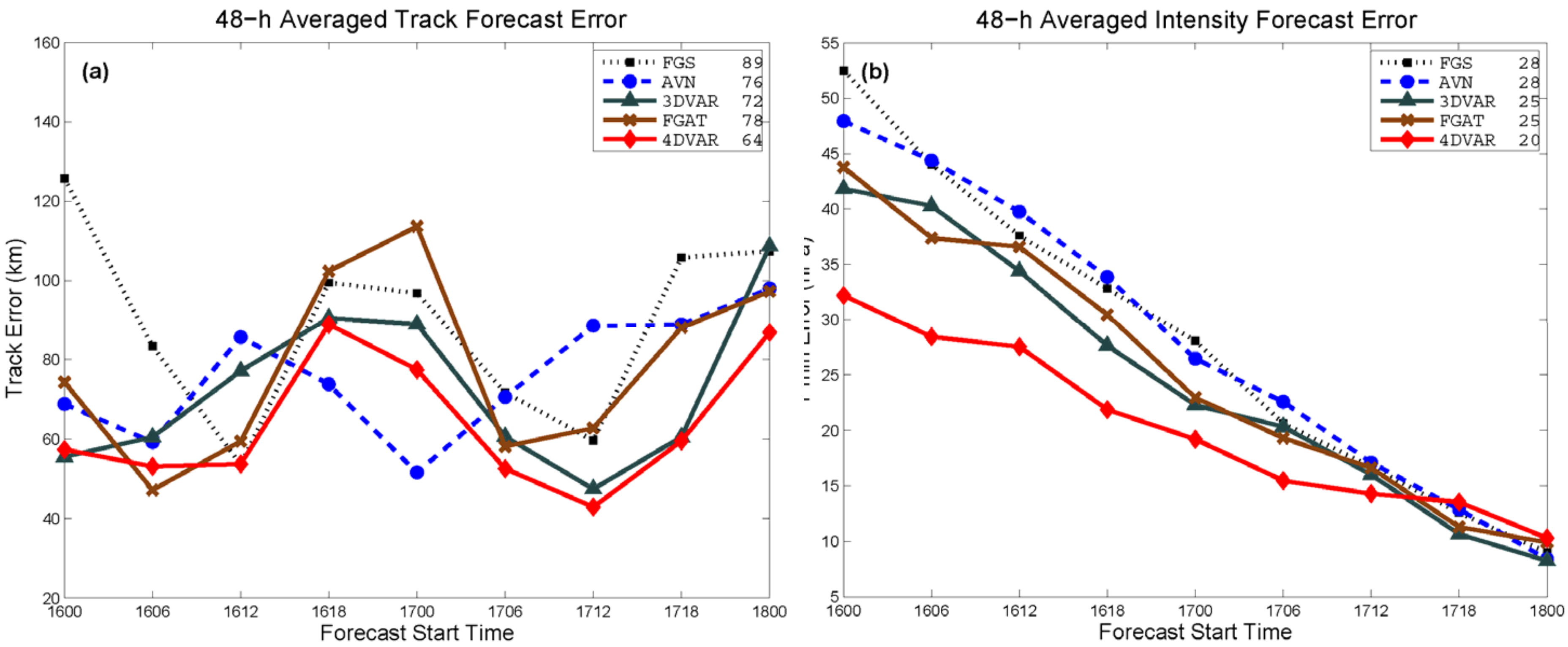
3-hour Surface Pressure Tendency



Real Case: Typhoon Haitang
Experimental Design (Cold-Start)

- Domain configuration: 91x73x17, 45km
- Period: 00 UTC 16 July - 00 UTC 18 July, 2005
- Observations from Taiwan CWB operational database.
- 5 experiments are conducted:
  - FGS – forecast from the background [The background fields are 6-h WRF forecasts from National Center for Environment Prediction (NCEP) GFS analysis.]
  - AVN- forecast from the NCEP AVN analysis
  - 3DVAR – forecast from WRF-Var3d using FGS as background
  - FGAT - forecast from WRF-Var3dFGAT using FGS as background
  - 4DVAR – forecast from WRF-Var4d using FGS as background

Typhoon Haitang 2005



A KMA Heavy Rain Case

Period: 12 UTC 4 May - 00 UTC 7 May, 2006

Assimilation window: 6 hours

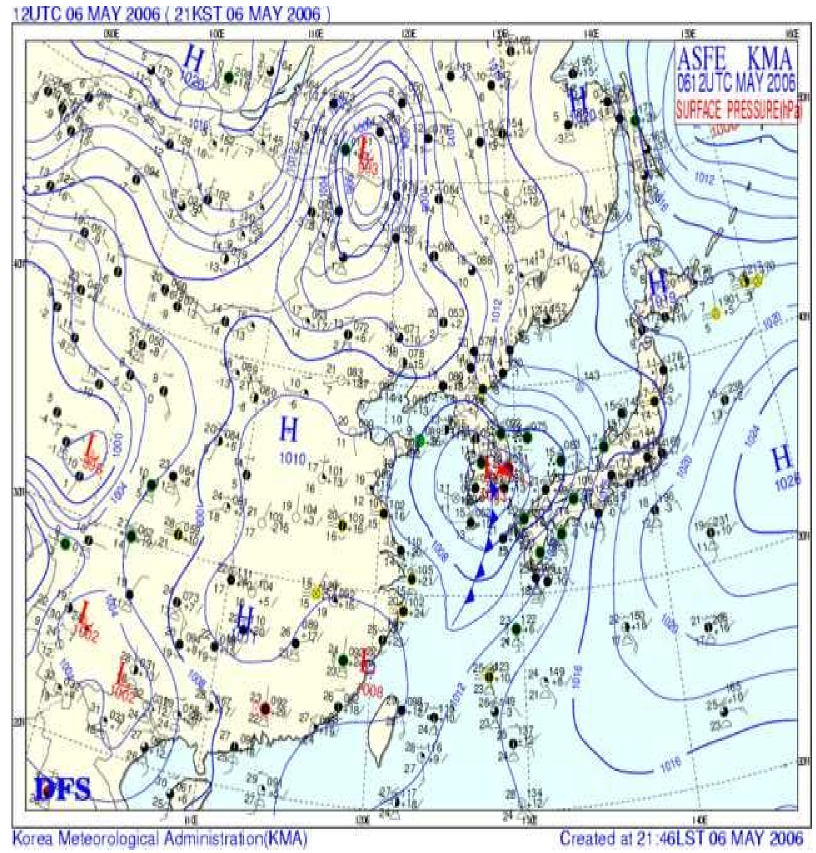
Cycling (6h forecast from previous cycle as background for analysis)

All KMA operational data

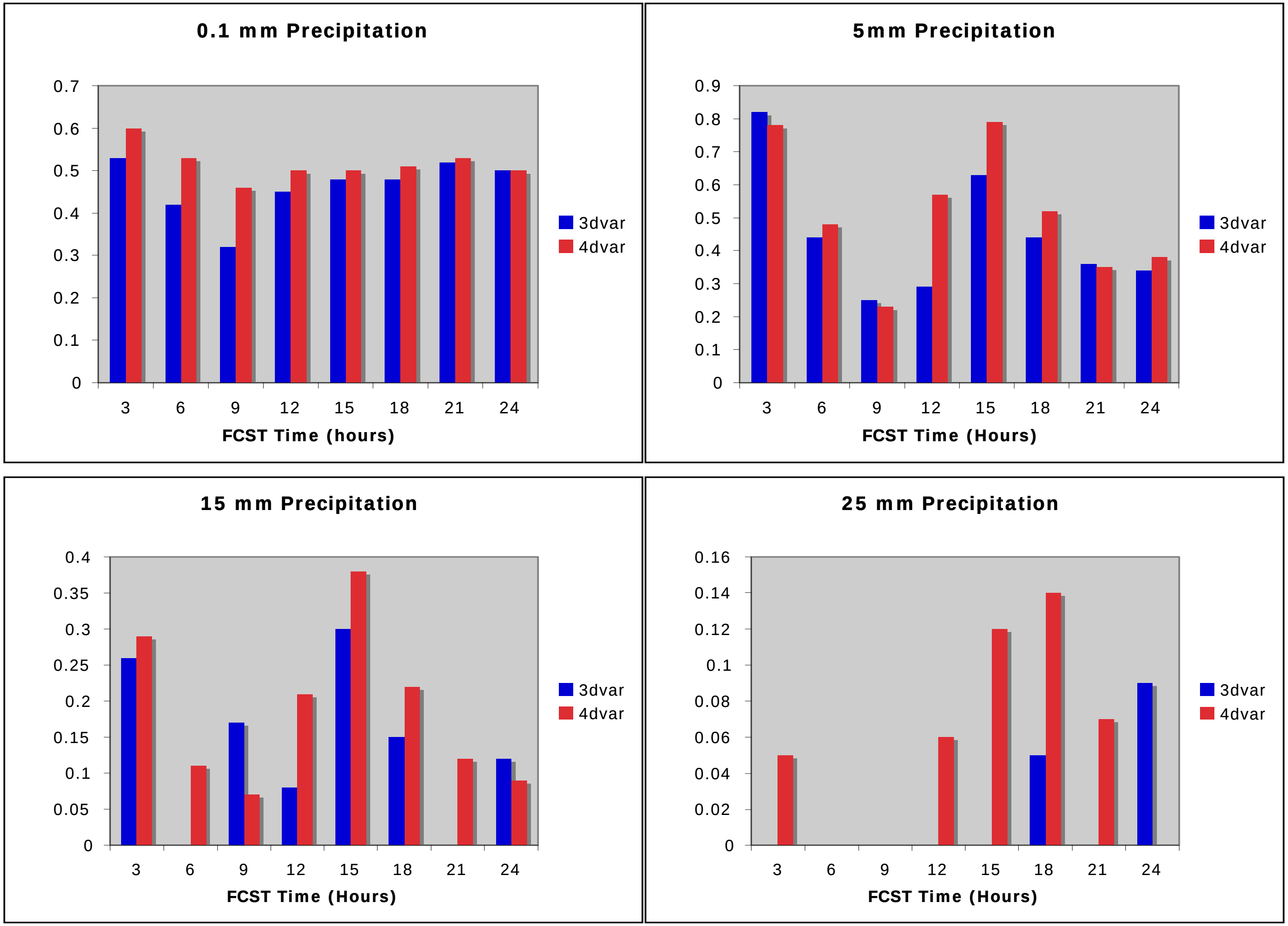
Grid : 60x54x31

Resolution : 30km

Domain size: the same as the KMA operational 10km domain.



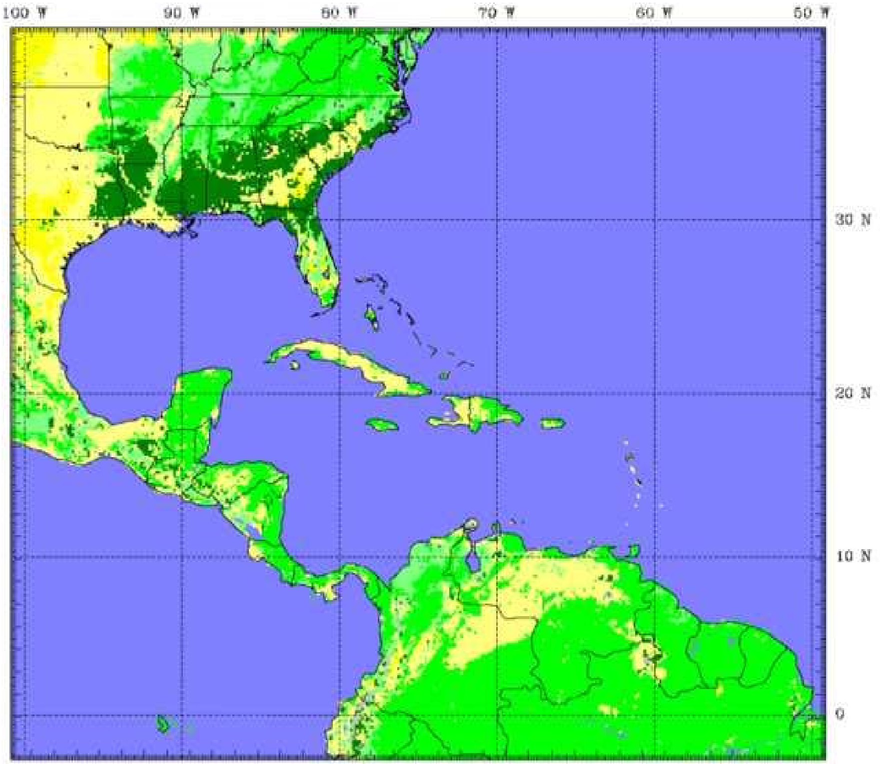
Precipitation Verification



Test domain: AFWA T8 45-km

Domain configurations:

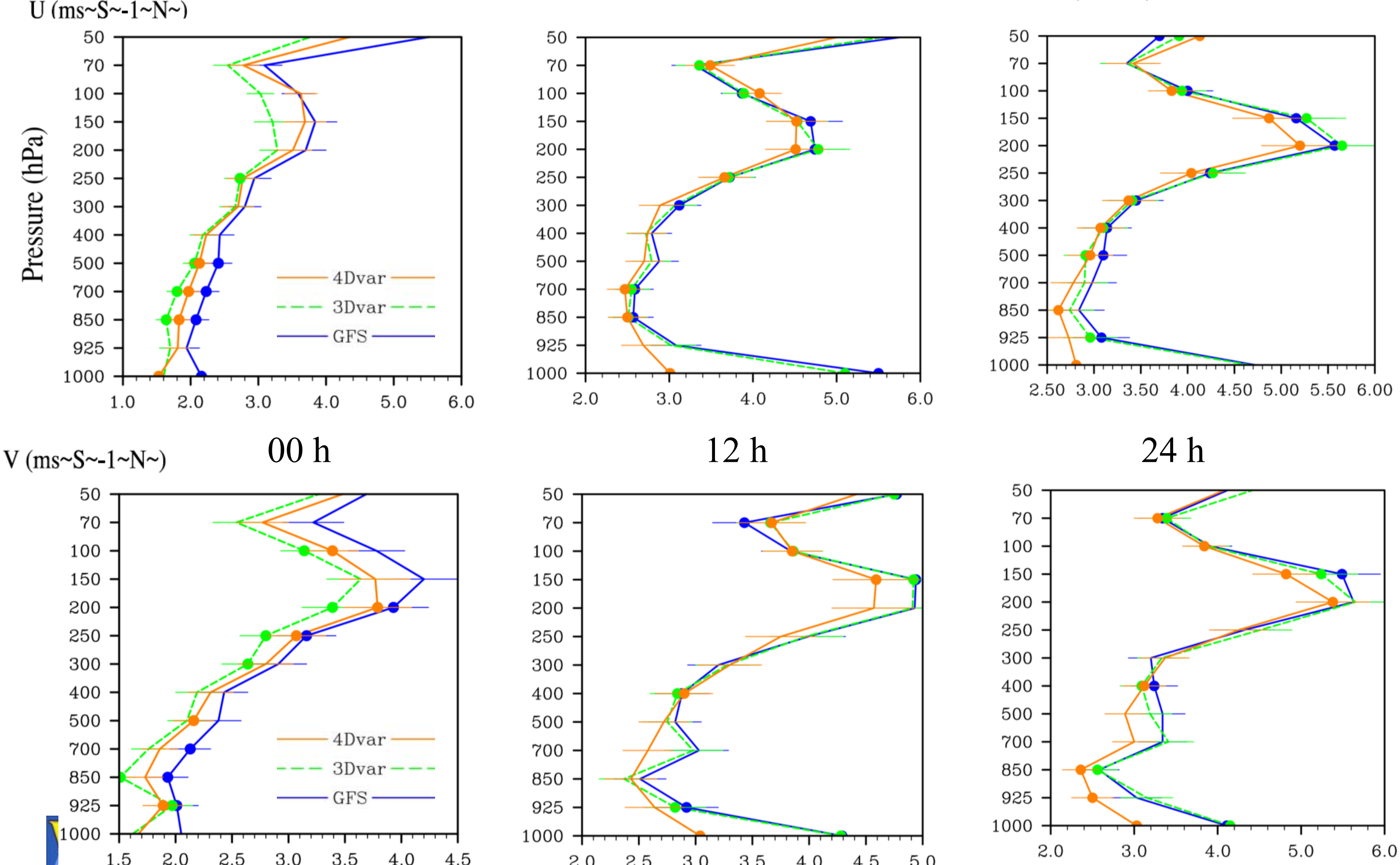
- 2007-08-15-00 to 2007-08-21-00
- Horizontal grid: 123 x 111 @ 45 km grid spacing, 27 vertical levels
- 3/6 hours GFS forecast as FG
- Every 6 hours



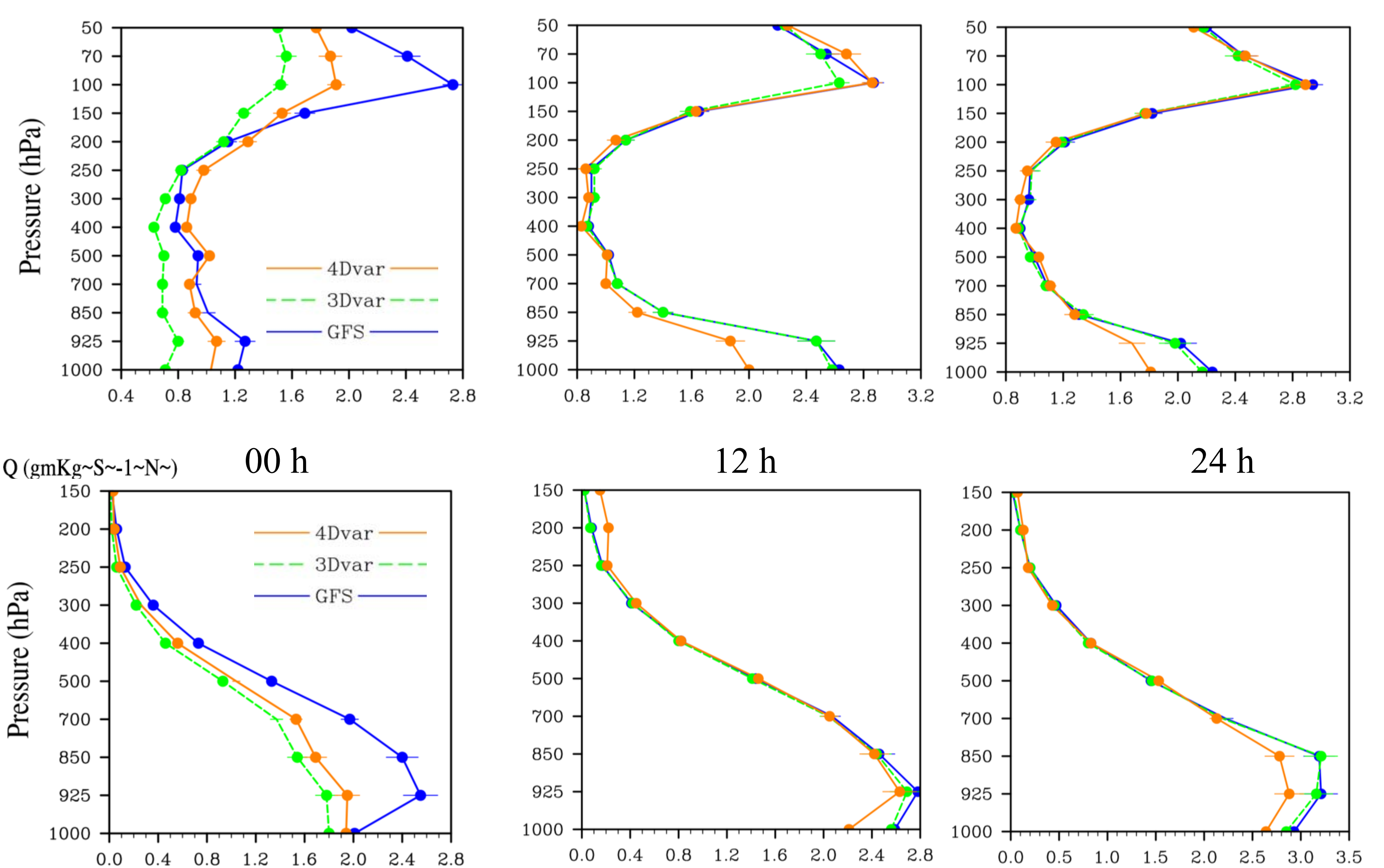
Observations used:

sound sonde_sfc
synop geoamv
airep pilot
ships qscat
gpspw gpsref

RMSE Profiles at Verification Times (u,v)



RMSE Profiles at Verification Times (T,q)



The first radar data assimilation experiment using WRF 4D-Var (OSSE)

TRUTH ----- Initial condition from TRUTH (13-h forecast initialized at 2002061212Z from AWIPS 3-h analysis) run cutted by ndown**,** boundary condition from NCEP GFS data.

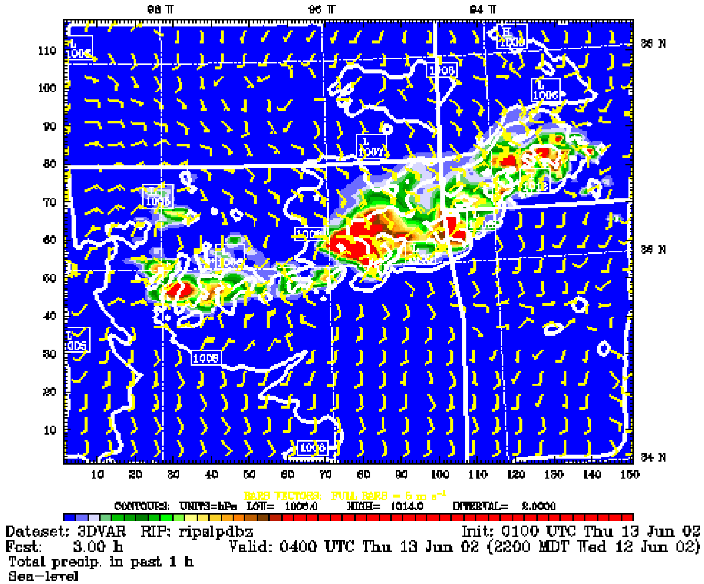
NODA ----- Both initial condition and boundary condition from NCEP GFS data.

3DVAR ----- 3DVAR analysis at 2002061301Z used as the initial condition, and boundary condition from NCEP GFS. Only Radar radial velocity at 2002061301Z assimilated (total # of data points = 65,195).

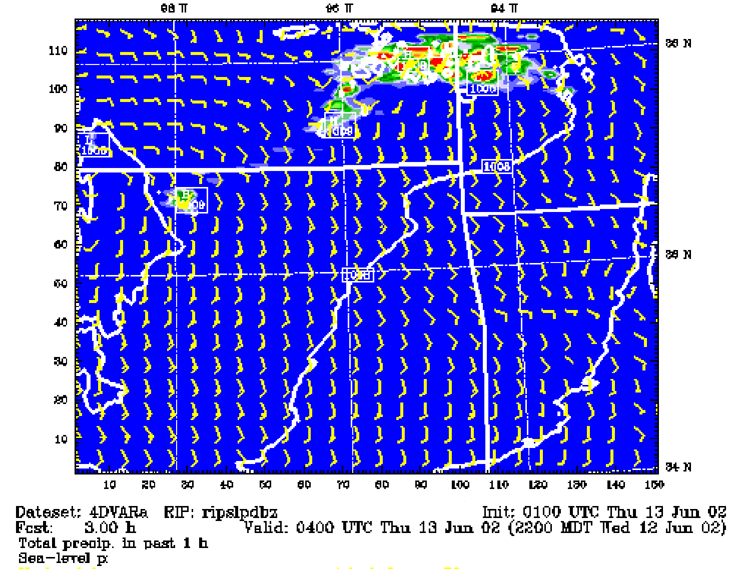
4DVAR ----- 4DVAR analysis at 2002061301Z used as initial condition, and boundary condition from NCEP GFS. The radar radial velocity at 4 times: 200206130100, 05, 10, and 15, are assimilated (total # of data points = 262,445).

Hourly precipitation ending at 03-h forecast

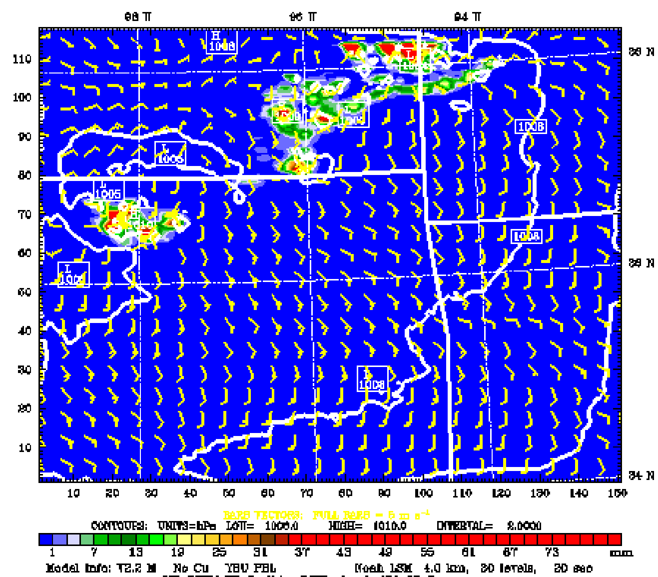
TRUTH



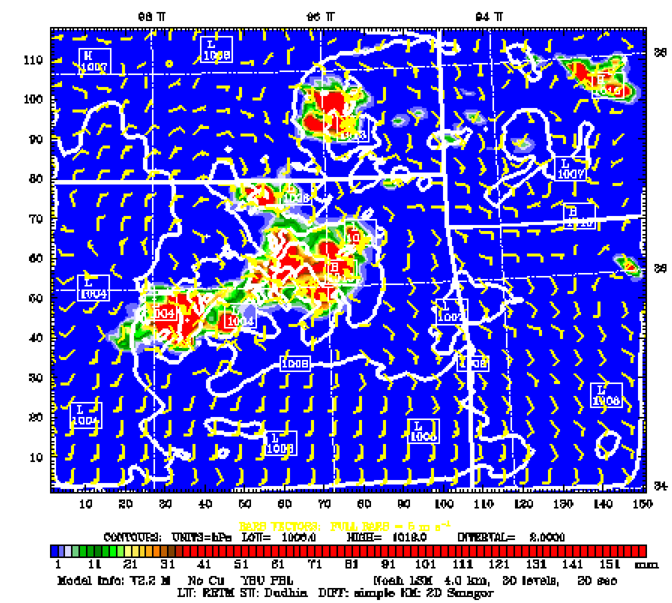
NODA



3DVAR

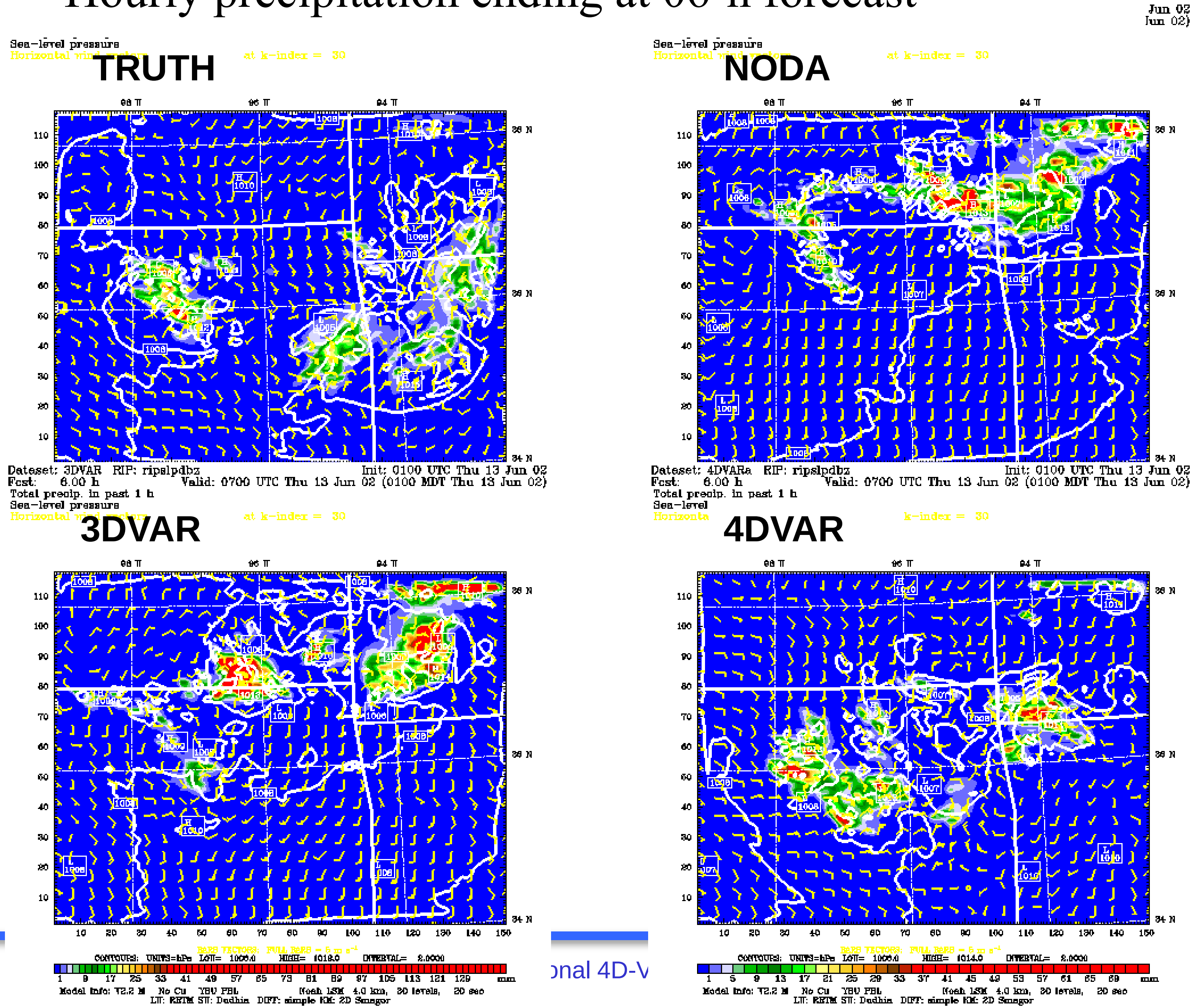


4DVAR

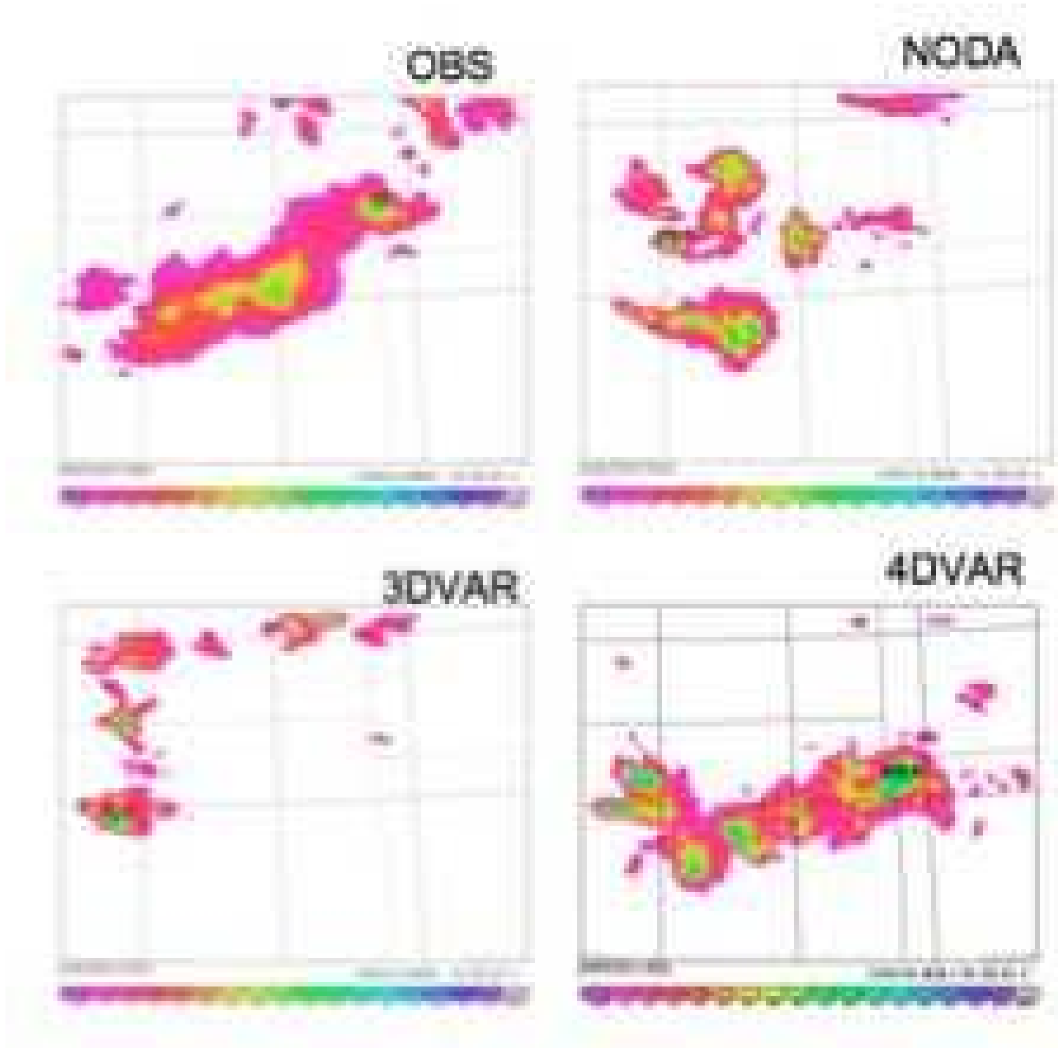


onal 4D

Hourly precipitation ending at 06-h forecast



Real data experiments



□Outline

- 4D-Var formulation and why regional
- The WRFDA approach
- Regional 4D-Var issues
- Summary

Specific Regional 4D-Var Issues

- Background error covariance estimations.
- Control of lateral boundaries.
- Control of the large scales or coupling to a large scale model?
- Mesoscale balances.

Model-Based Estimation of Climatological Background Errors

- Assume background error covariance estimated by model perturbations *x'* :

$$\mathbf{B} = \overline{(\boldsymbol{x}^b - \boldsymbol{x}^t)(\boldsymbol{x}^b - \boldsymbol{x}^t)^T} \approx \overline{\boldsymbol{x}'\boldsymbol{x}'^T}$$

Two ways of defining ***x'***:

- The NMC-method (Parrish and Derber 1992):

$$\mathbf{B} = \overline{\boldsymbol{x}'\boldsymbol{x}'^T} \approx A\overline{(\boldsymbol{x}^{t2} - \boldsymbol{x}^{t1})(\boldsymbol{x}^{t2} - \boldsymbol{x}^{t1})^T}$$

  where e.g. t2=24hr, t1=12hr forecasts…

- …or ensemble perturbations (Fisher 2003):

$$\mathbf{B} = \overline{\boldsymbol{x}'\boldsymbol{x}'^T} \approx C\overline{(\boldsymbol{x}^k - \langle \boldsymbol{x} \rangle)(\boldsymbol{x}^k - \langle \boldsymbol{x} \rangle)^T}$$

- Tuning via innovation vector statistics and/or variational methods.

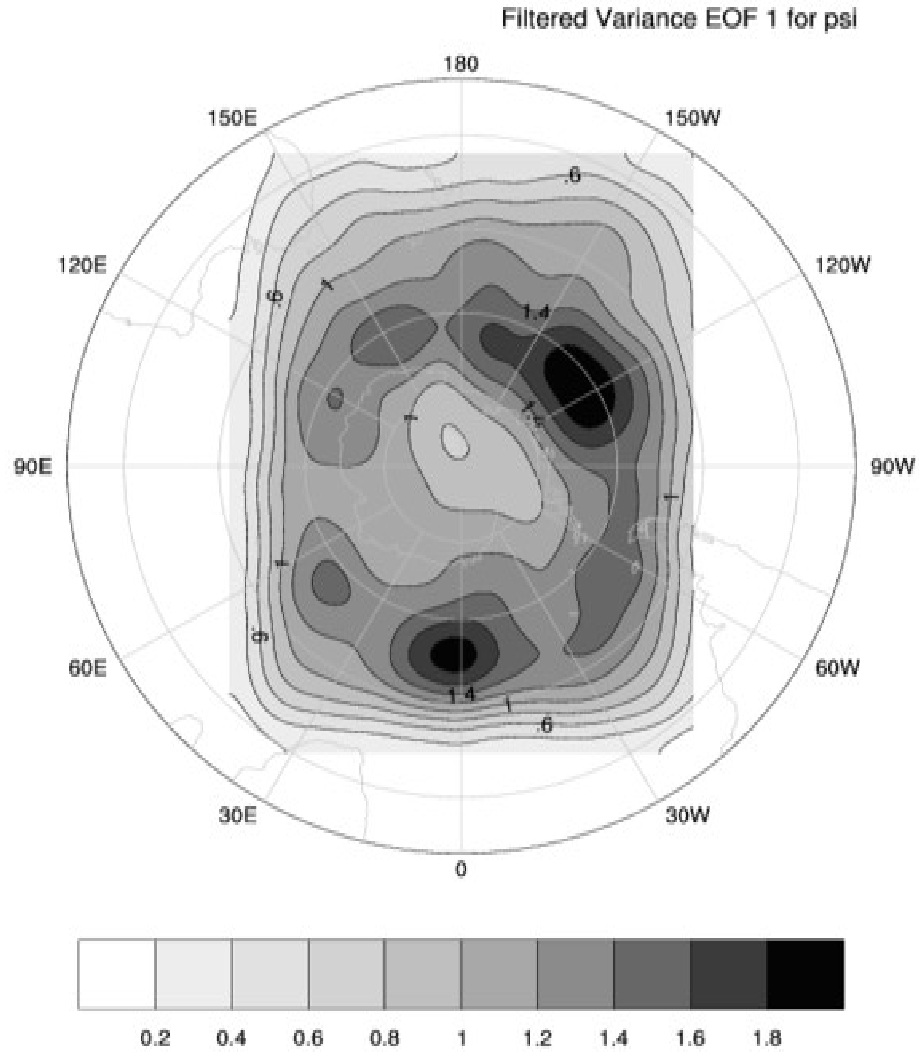
Background Error

Variances are computed for unbalanced fields projected on vertical Empirical Orthogonal Function.

This allows to give different weights to the observations depending on their geographical positions.

This example shows that mid-tropospheric wind observations are likely to be given more weight in the circumpolar vortex, where uncertainty is high.

Variance is also lower near the boundaries, because of the large scale coupling.



Yann MICHEL

Sensitivity of forecast error with respect to initial and lateral boundary conditions

(From Gustafsson et al 1997)

- Consider a case with a significant forecast error at, for example +12h.
- Calculate a forecast error norm and a gradient of this error norm with respect to the model state variables by comparing with, for example, an analysis
- Project the gradient of the forecast error norm back to the initial state by the aid of the adjoint model. We will have the gradient of the forecast error norm with respect to the initial and lateral boundary conditions. “Sensitivity patterns”
- Add a negative fraction of these gradients to the initial and lateral boundary data and re-run the forecast. “Sensitivity forecast”

Forecast error

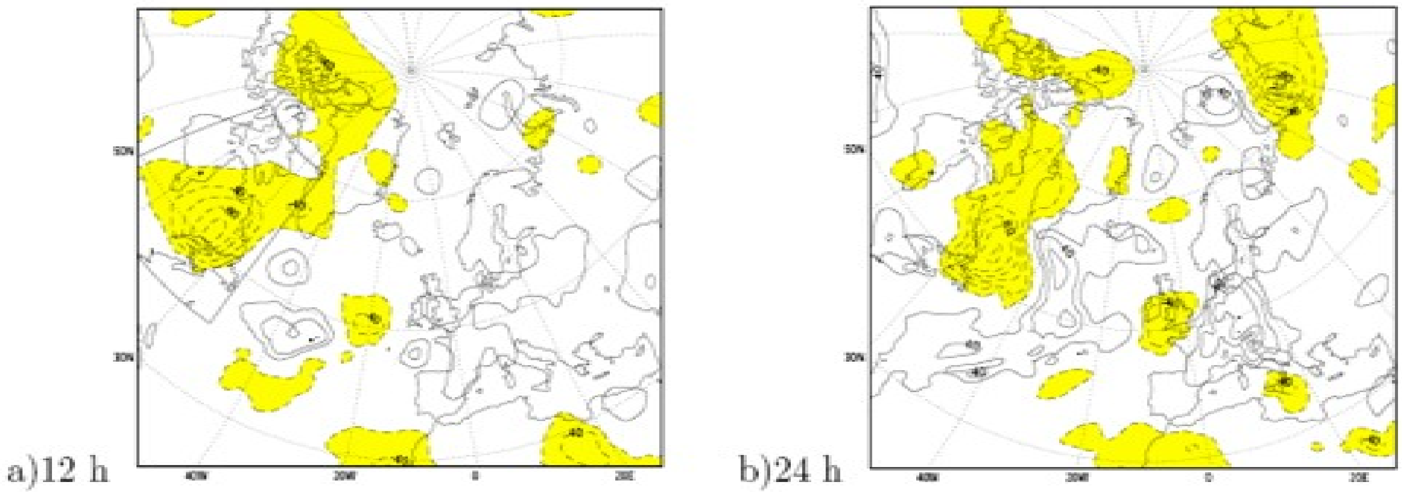


Figure 2: 300 hPa height error fields (verification analysis - forecasts from operational SMHI HIRLAM) for forecast lengths of 12 h and 24 h from 0000 UTC on 16 February 1995. Positive (negative) values are represented by solid (dashed) contours. Contours every 20 m. The shaded areas correspond to errors $\leq$ -20 m. The sector-formed area in a) is used for calculation of the area-concentrated 12 h forecast error norm.

Sensitivity with respect to initial conditions

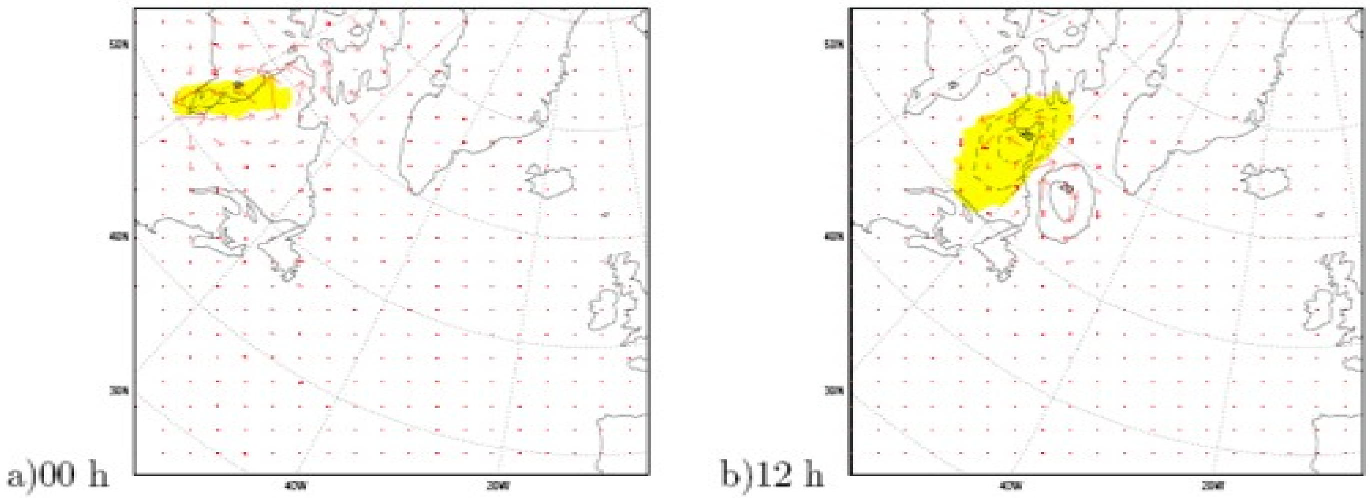


Figure 3: 300 hPa height and wind differences between the sensitivity and reference 0 h and 12 h forecasts from 0000 UTC on 16 February 1995 using initial perturbations only with $\alpha_i = 0.1$ and the area-concentrated 12 h forecast error norm. Positive (negative) values are represented by solid (dashed) contours. Contours every 10 m. The shaded areas correspond to retrieved forecast errors $\leq$ -5 m. The western boundary of the model integration area coincides with the left side of the maps.

Sensitivity with respect to lateral boundary conditions

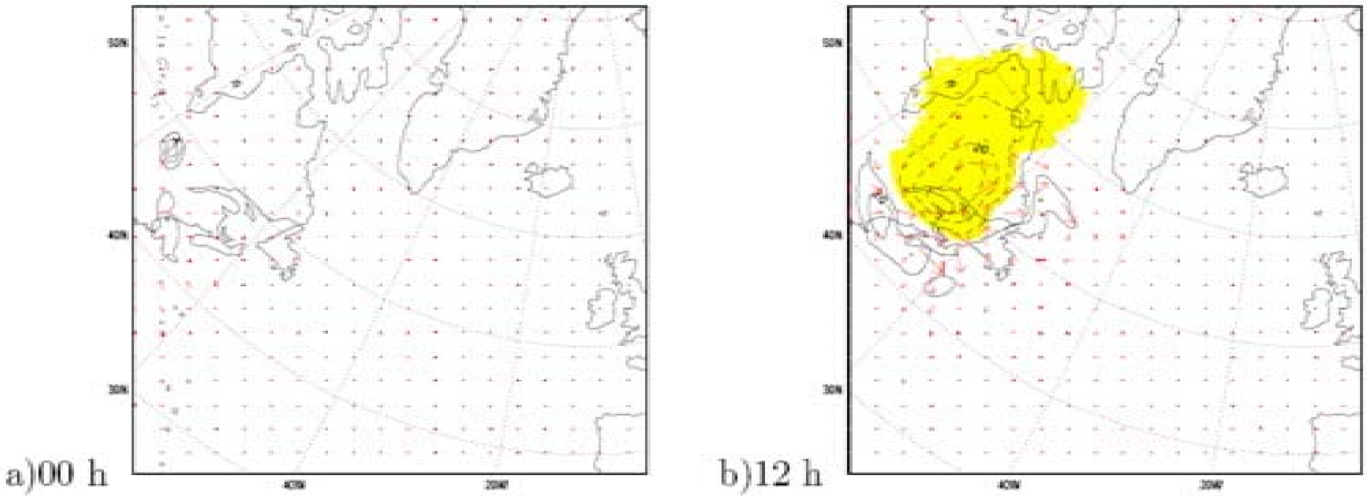


Figure 4: As in Fig. 3, but using only the lateral boundary perturbations for the sensitivity forecast run with $\alpha_b = 5.0$ and with a time-averaging period of ± 1 h: a) 0 h (boundary perturbation), b) 12 h. (a) Contour interval of 2 m, (b) contour interval of 10 m.

Sensitivity with respect to initial and lateral boundary conditions

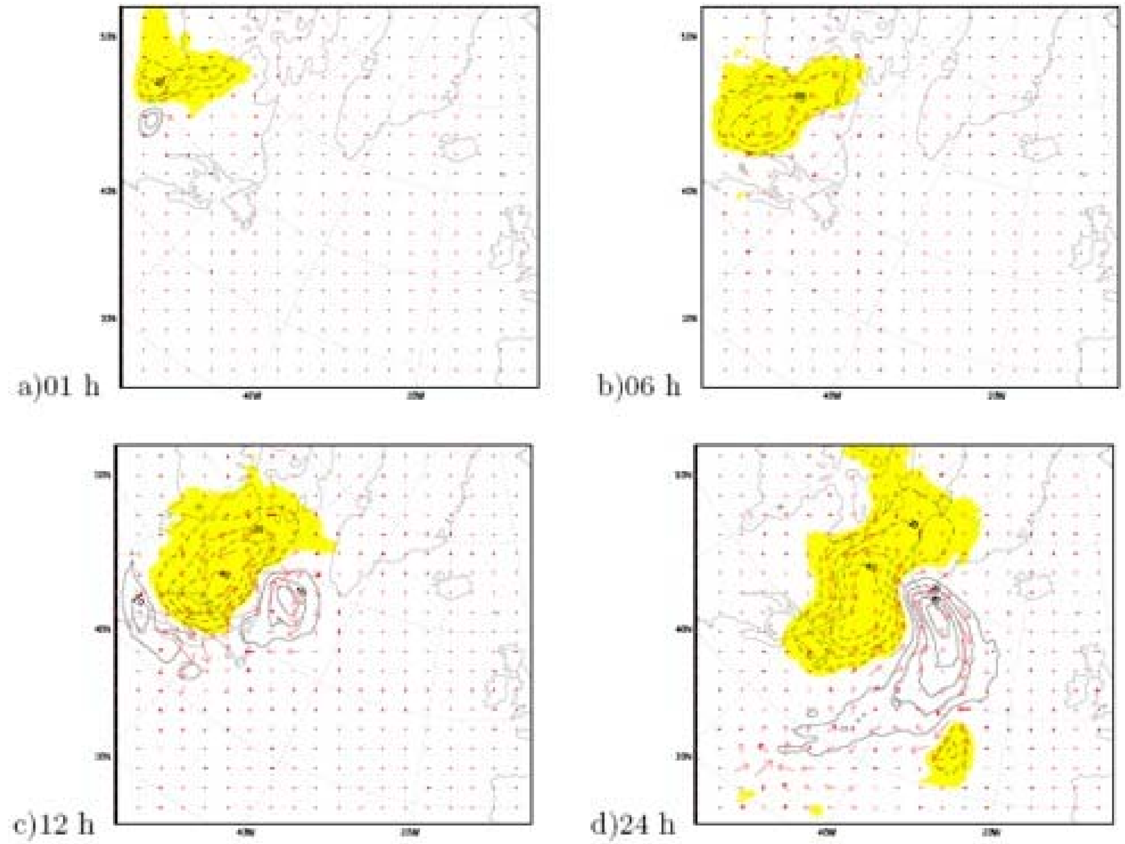


Figure 5: As in Fig. 3, but using initial as well as lateral boundary perturbations for the sensitivity forecast run with $\alpha_i = 0.1$ and $\alpha_b = 5.0$ and with a time-averaging period of ± 1 h: a) 1 h, b) 6 h, c) 12 h, d) 24 h.

Options for LBC in 4D-Var

(Nils Gustafsson)

1. Trust the large scale model providing the LBC.

Apply relaxation toward zero on the lateral boundaries during the integration of the TL and AD models in regional 4D-Var.

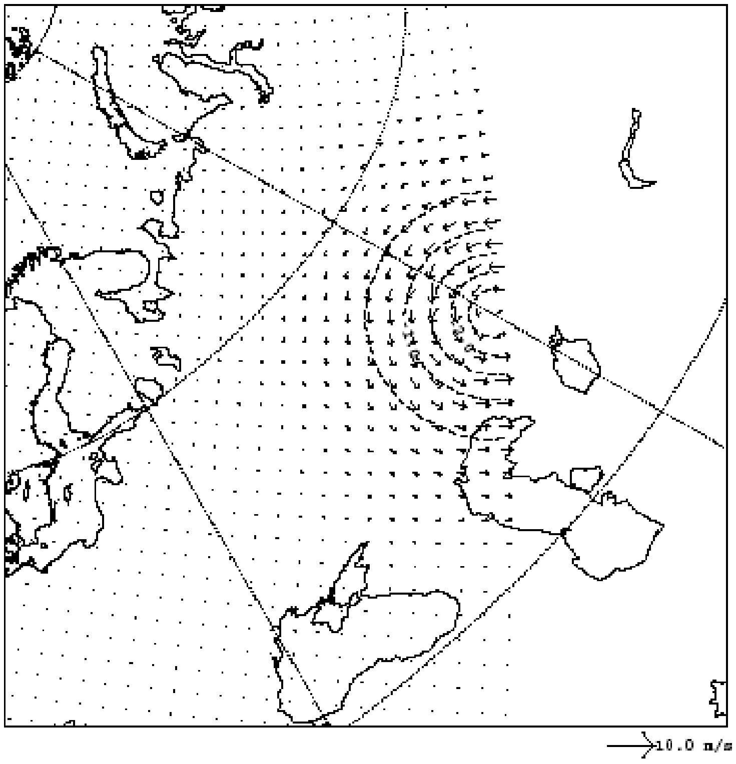
This, however, will have some negative effects:

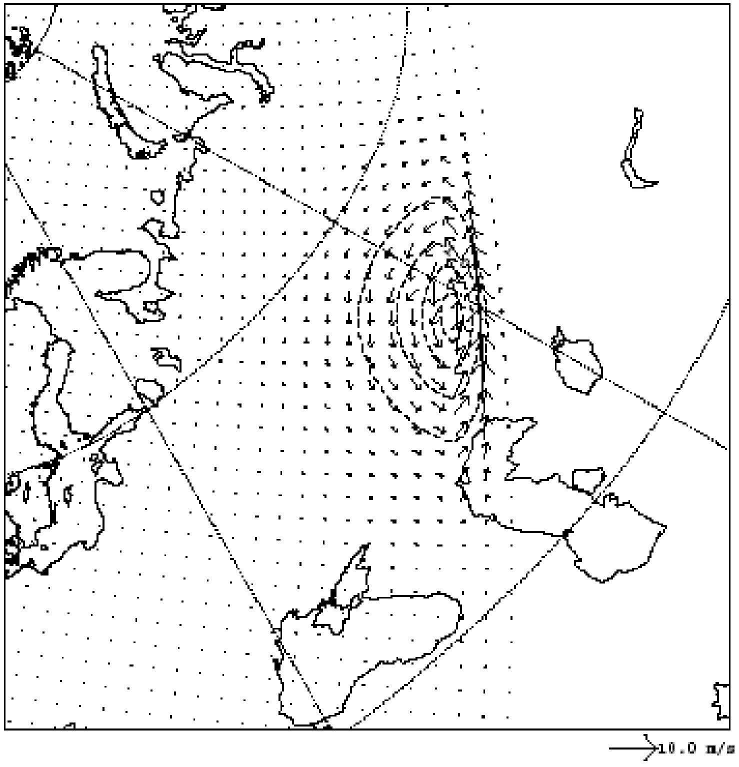
- Larger scale increments as provided via the background error constraint may be filtered during the forward TL model integration.
- Observations close the the lateral boundaries will have reduced effect. In particular, information from observations downstream of the lateral boundaries and from the end of the assimilation window will be filtered due to the adjoint model LBC relaxation.

2. Control the lateral boundary conditions.

Regional data assimilation LBC

3D-Var via the background error constraint formulation:





(Nils Gustafsson)

Control of Lateral Boundary Conditions (JMA and HIRLAM)

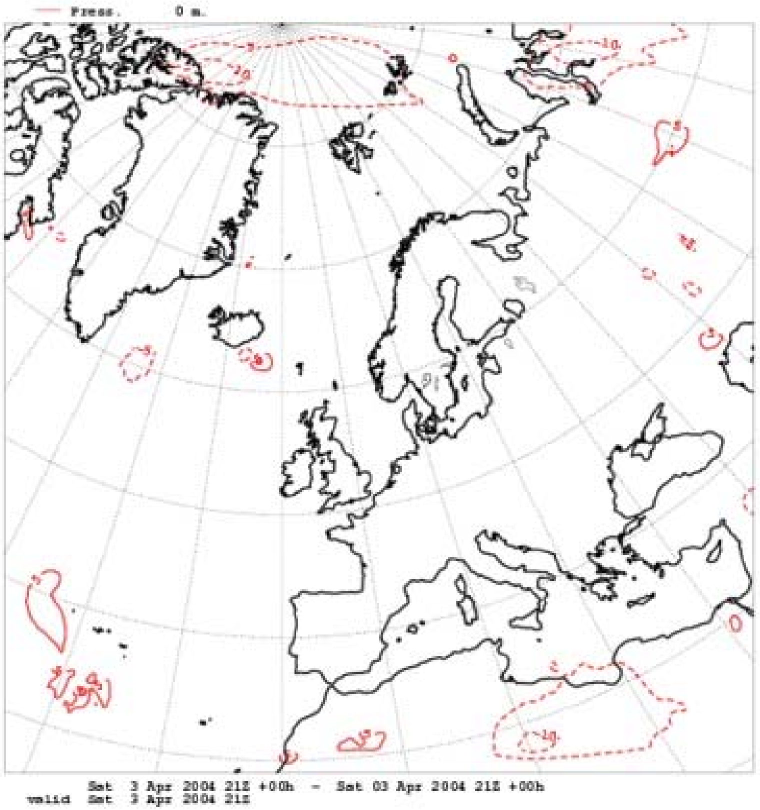
1) Introduce the LBCs at the end of the data assimilation window as assimilation control variables (full model state = double size control vector)
2) Introduce the adjoints of the Davies LBC relaxation scheme and the time interpolation of the LBCs
3) Introduce a "smoothing and balancing" constraint for the LBCs into the cost function to be minimized

$$J = J_b + J_o + J_c + J_{lbc}$$

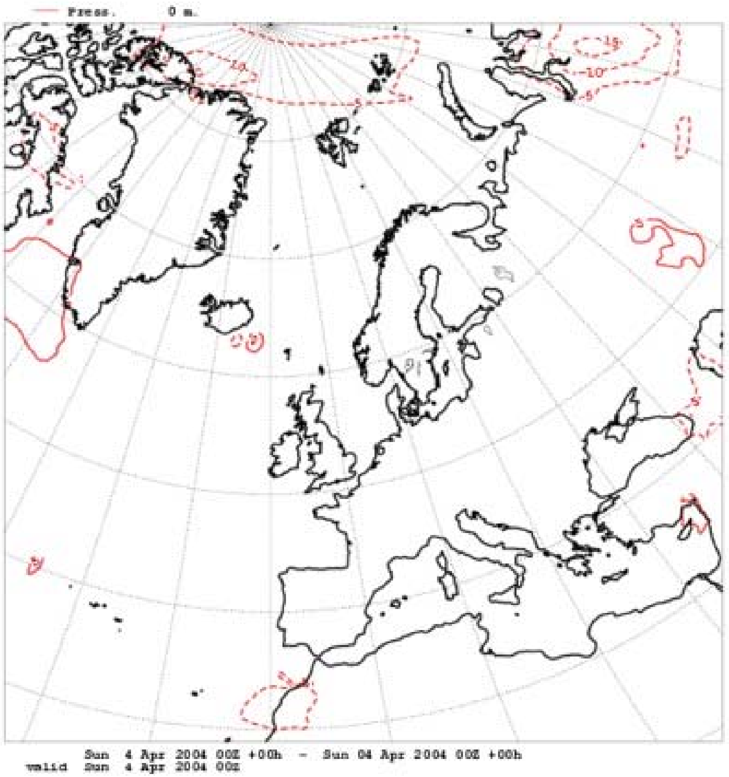
where

$$J_{\mathbf{lbc}} = \frac{1}{2}\left(\mathbf{x}_{lbc} - \mathbf{x}_{lbc}^{b}\right)^{T}\mathbf{B}_{lbc}^{-1}\left(\mathbf{x}_{lbc} - \mathbf{x}_{lbc}^{b}\right)$$

Example, first cycle with lbc control (HIRLAM)

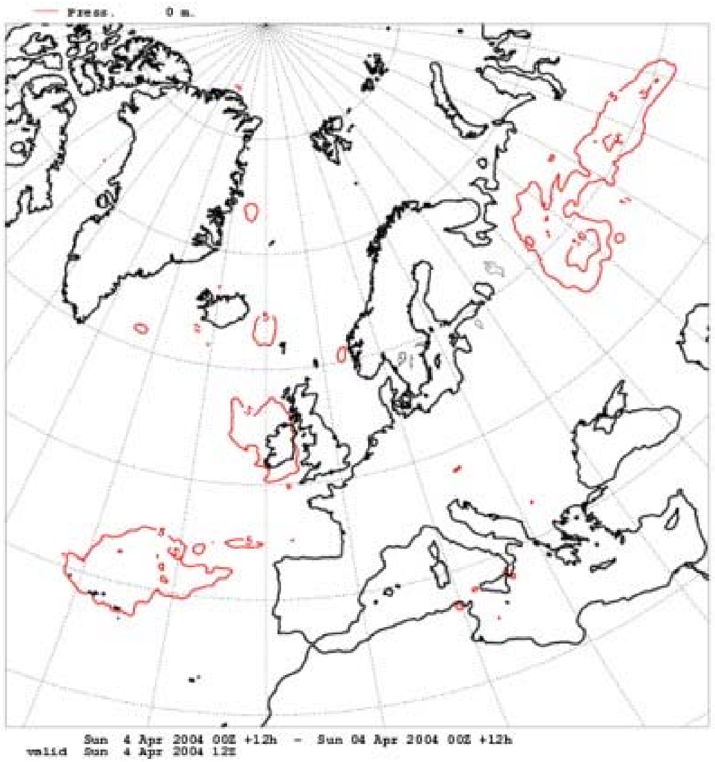


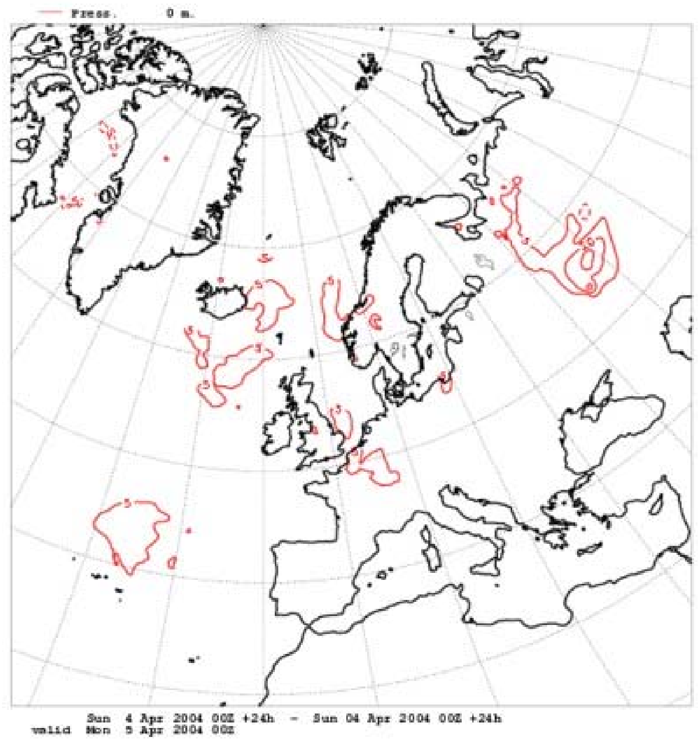
21 UTC, start of window



00 UTC Nominal analysis time

Effects on the forecasts (HIRLAM)





+12 h

+24 h


Press. 0 m.
Sun 4 Apr 2004 00Z +48h - Sun 04 Apr 2004 00Z +48h
valid Tue 6 Apr 2004 00Z

+48 h

Problems with Control of LBC

- Double the control vector length
- Pre-conditioning; LBC_0h and LBC_6h are correlated
- Relative weights for J_b and J_{lbc}??

Control of "larger" horizontal scales in regional 4D-Var

- There are difficulties to handle larger horizontal scales in regional 4D-Var
- Observations inside a small domain may not provide information on larger horizontal scales
- The assimilation basis functions (via the background error constraint) may not represent larger scales properly .
- Possibilities to handle the larger horizontal scales:
  - Via the LBC in the forecast model only; Depending on the cycle for refreshment of LBC, one may miss important advection of information over the LBC
  - Via ad hoc techniques for mixing in information from a large scale data assimilation (applied in HIRLAM).
  - Via a large scale error constraint - J_{ls}

A large scale error constraint - J_{ls}

Add a large scale constraint J_{ls} to the regional 4D-Var cost function:

$$J = J_b + J_o + J_c + J_{lbc} + J_{ls}$$

where

$$J_{ls} = \frac{1}{2}(\mathbf{x} - \mathbf{x}_{ls})^T \mathbf{B}_{ls}^{-1} (\mathbf{x} - \mathbf{x}_{ls})$$

- Has been tried in ALADIN 3D-Var with x_{ls} being a global analysis from the same observation time. In this case one needs to, at least in principle, use different observations for the global and regional assimilations

- Is being tried in HIRLAM 4D-Var with x_{ls} being a global (+3h) forecast valid at the start of the assimilation window.

For both applications one need to check that ($\mathbf{x}$-$\mathbf{x}^b$) and ($\mathbf{x}$-$\mathbf{x}_{ls}$) are not too strongly correlated!

□Summary

- 4D-Var formulation and why 4D-Var
- Why regional (high resolution; regional observation network;…)
- The WRFDA approach: J_b; J_o; J_c
- Regional 4D-Var issues: **B**, J_{lbc}, J_{ls}