WRF Four-dimensional variational data assimilation Tutorial for V3.3

Xin Zhang Xiang-Yu Huang

NCAR Earth System Laboratory

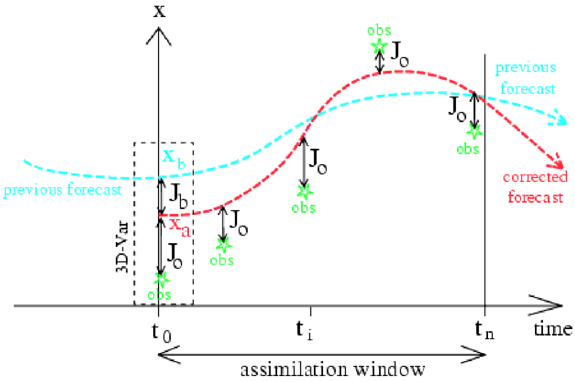
Presented on 24 June 2011, 12th WRF Users' Workshop

NCAR is sponsored by the National Science Foundation

Pre-requirements to run WRF 4D-Var

- Knowledge and experience to run WRF model
- Knowledge and experience to run WRFDA (3D-Var)
- WRF 4D-Var V3.3 is quite different from its previous versions, so no experience for previous WRF 4D-Var versions required.

4D-Var versus 3D-Var (Adopted from ECMWF training Course)



4D-Var versus 3D-Var (Adopted from ECMWF training Course)

- 4D-Var is comparing observations with background model fields at the correct time
- 4D-Var can use observations from frequently reporting stations
- The dynamics and physics of the forecast model in an integral part of 4D-Var, so observations are used in a meteorologically more consistent way
- 4D-Var combines observations at different times during the 4D-Var window in a way that reduces analysis error
- 4D-Var propagates information horizontally and vertically in a meteorologically more consistent way

Incremental 4D-Var formulation

$$J = J_b + J_o$$

Define analysis increment: $\delta \mathbf{x} = \mathbf{x} - \mathbf{x_b}$

$$J_b = \frac{1}{2} \delta \mathbf{x}^T \mathbf{B}^{-1} \delta \mathbf{x}$$

$$J_o = \frac{1}{2} \sum_{k=1}^{K} [(\overbrace{\mathbf{y}_k - H_k M_k \mathbf{x}_b}^{innovation} - \mathbf{H}_k \mathbf{M}_k \delta \mathbf{x})^T \mathbf{R}^{-1}$$

$$(\mathbf{y}_k - H_k M_k \mathbf{x}_b - \mathbf{H}_k \mathbf{M}_k \delta \mathbf{x})]$$

Incremental 4D-Var formulation (Cont'd)

To find the $\delta\mathbf{x}$ which lead the J to minimal:

$$\nabla_{\delta\mathbf{x}} J = 0$$

$$\nabla_{\delta\mathbf{x}} J_b = \mathbf{B}^{-\mathbf{1}}\delta\mathbf{x}$$

$$\nabla_{\delta\mathbf{x}} J_o = \sum_{k=1}^{K}[\underbrace{\mathbf{M}_k^T}_{\text{WRF\_AD}} \mathbf{H}_k^T\mathbf{R}^{-1}(\mathbf{y}_k - H_k \underbrace{M_k}_{\text{WRF\_NL}} \mathbf{x}_b - \mathbf{H}_k \underbrace{\mathbf{M}_k}_{\text{WRF\_TL}} \delta\mathbf{x})]$$

M_k : Model integration from step 0 to step k.
$\mathbf{M}_k$: tangent linear model and ${\mathbf{M}_k}^T$: adjoint model are needed.

Weak constraint with digital filter

$$J = J_b + J_o + J_c$$

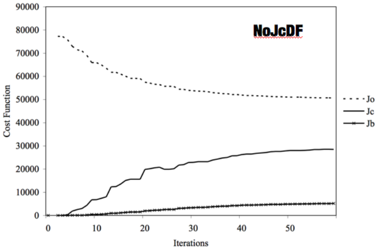
$$J_b(\mathbf{x}_0) = \frac{1}{2}\left[(\mathbf{x}_0 - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b)\right]$$

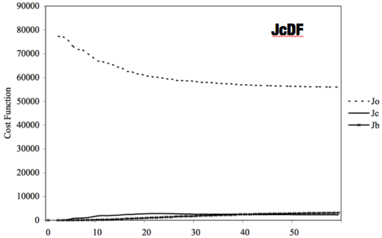
$$J_o(\mathbf{x}_0) = \frac{1}{2}\sum_{k=1}^{K}\left[(\mathbf{H}_k\mathbf{x}_k - \mathbf{y}_k)^T \mathbf{R}^{-1}(\mathbf{H}_k\mathbf{x}_k - \mathbf{y}_k)\right]$$

$$\begin{aligned} J_c(\mathbf{x}_0) &= \frac{\gamma_{df}}{2}\left[\left(\delta\mathbf{x}_{N/2} - \delta\mathbf{x}^{df}_{N/2}\right)^T \mathbf{C}^{-1}\left(\delta\mathbf{x}_{N/2} - \delta\mathbf{x}^{df}_{N/2}\right)\right] \\ &= \frac{\gamma_{df}}{2}\left[\left(\delta\mathbf{x}_{N/2} - \sum_{i=0}^{N} f_i\delta\mathbf{x}_i\right)^T \mathbf{C}^{-1}\left(\delta\mathbf{x}_{N/2} - \sum_{i=0}^{N} f_i\delta\mathbf{x}_i\right)\right] \\ &= \frac{\gamma_{df}}{2}\left[\left(\sum_{i=0}^{N} h_i\delta\mathbf{x}_i\right)^T \mathbf{C}^{-1}\left(\sum_{i=0}^{N} h_i\delta\mathbf{x}_i\right)\right] \end{aligned}$$

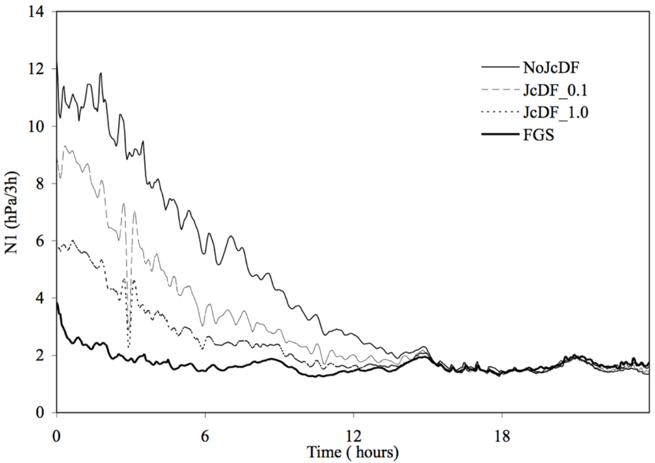
where:

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





Weak constraint with digital filter (domain averaged surface pressure variation)



Consider lateral boundary condition as control variable

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

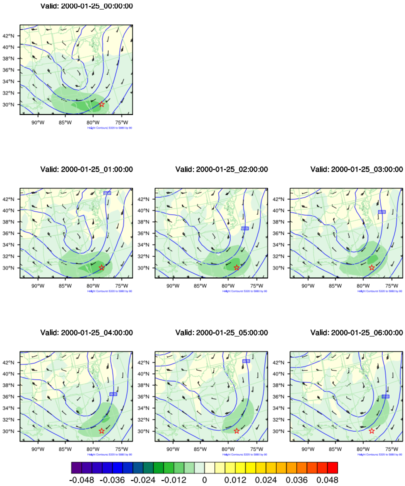
$$\begin{aligned} J_{lbc} &= \frac{1}{2}(\mathbf{x}(t_k) - \mathbf{x}_b(t_k))^T \mathbf{B}^{-1}(\mathbf{x}(t_k) - \mathbf{x}_b(t_k)) \\ &= \frac{1}{2}\delta\mathbf{x}(t_k)^T \mathbf{B}^{-1}\delta\mathbf{x}(t_k) \end{aligned}$$

J_{lbc} is the J_b at the end of the assimilation window
lateral boundary control is obtained through

$$\frac{\partial \delta \mathbf{x}_{lbc}}{\partial t} = \frac{\delta\mathbf{x}(t_k) - \delta\mathbf{x}(t_0)}{t_k - t_0}$$

Single observation experiment

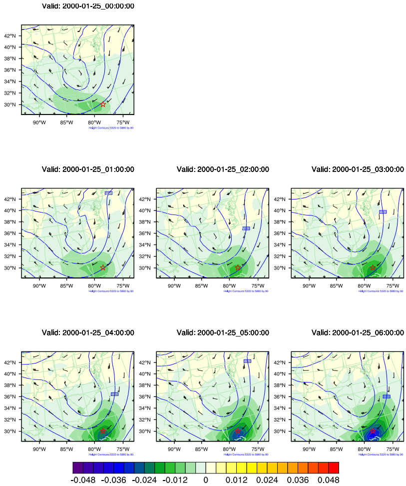
To investigate the impact of including boundary condition control in data assimilation, a 6h observation close to boundary is put at the downstream of the boundary inflow, we expect that the major analysis increments response at 0h should be in boundary condition and outside of domain.



Remarks

Forecasted 500mb T difference (DA forecast - reference forecast)

- $\star$ is the location of obs. at the ending time (6h).
- $O - B = -0.95K$
- LBC control is turned off



Remarks

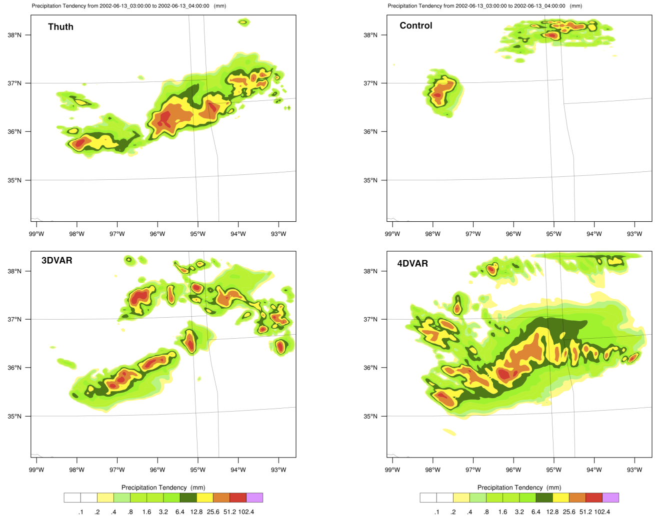
Forecasted 500mb T difference (DA forecast - reference forecast)

- ⋆ is the location of obs. at the ending time (6h).
- LBC control is turned on

An OSSE radar data assimilation with WRF 4D-Var

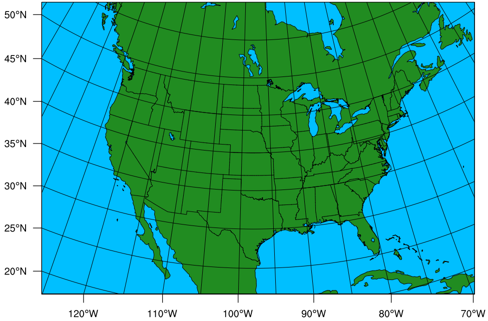
- TRUTH —– Initial condition from TRUTH (13-h forecast initialized at 2002061212Z from AWIPS 3-h analysis) run cutted by ndown, boundary condition from NCEP GFS data.
- NODA —– Both initial condition and boundary condition from NCEP GFS data.
- 3DVAR —–3DVAR analysis at 2002061301Z used as the initial condition, and boundary condition from NCEP GFS. Only Radar radial velocity at 2002061301Z assimilated (total data points = 97,033), 3 outer loops.
- 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 data points = 384,304), 3 outer loops.

OSSE 3rd hour precipitation simulation

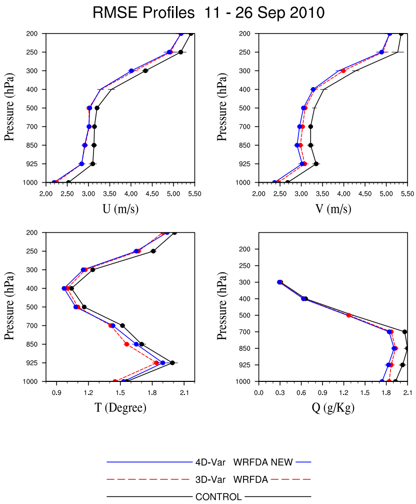


Experiment configuration

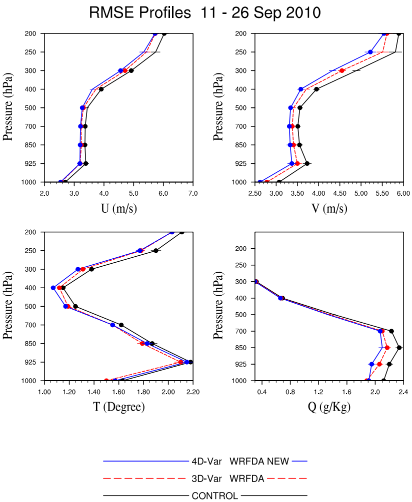
- Grids: 105x72x28L
- Resolution: 60km
- Period: 2010091100-2010092600 @0Z,6Z,12Z,18Z
- First guess is the 12h forecast from NCEP FNL
- 48h forecasts from FG, 3DVAR and 4DVAR
- Verified against NCEP GDAS prepbufr data



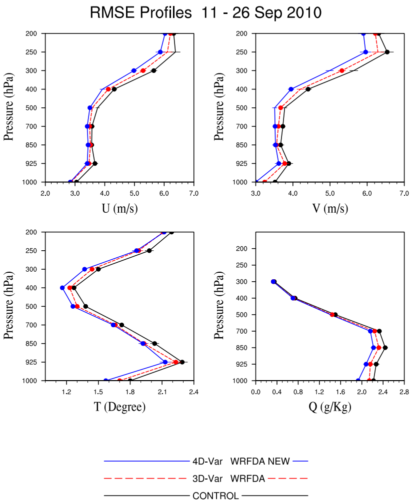
Averaged RMSE of 24H forecast verification



Averaged RMSE of 36H forecast verification



Averaged RMSE of 48H forecast verification



Download and setup test dataset for this tutorial

- download the WRFDA and WRFPLUSV3 codes from :

  `http://www.mmm.ucar.edu/wrf/users/wrfda/download/get_source.html`
- enter into WRFDA/var/test/4dvar
- get the test dataset from :

  `ftp://ftp.ucar.edu/pub/mmm/xinzhang/WRF4DVar_V3.3_Tutorial/2008020521`
- $ln\ -fs\ wrfinput\_d01\ fg$
- $ln\ -fs\ ../../build/da\_wrfvar.exe\ .$
- $ln\ -fs\ ../../run/be.dat.cv3\ be.dat$

Installation

- Install WRFPLUS V3.3
  - $./configure\ (-d)\ wrfplus$
  - $./compile\ em\_real$
  - $wrf.exe$ should be generated under $main$ directory.
- for csh, tcsh : $setenv\ WRFPLUS\_DIR\ path\_of\_wrfplusv3$
- for bash, ksh : $export\ WRFPLUS\_DIR = path\_of\_wrfplusv3$
- Install WRFDA V3.3
  - $./configure\ (-d)\ 4dvar$
  - $./compile\ all\_wrfvar$
  - $da\_wrfvar.exe$ should be generated under $var/build$ directory.

Tips for compilation

- Speed up the compilation with parallel make —gnu make: $setenv\ J\ "-j\ 6"$
- $setenv\ BUFR$ 1 to assimilate prepbufr observation.
- $setenv\ CRTM$ 1 to assimilate radiance bufr data with CRTM.

Portability

We have tested the WRF 4D-Var V3.3 on following systems:

- IBM with XLF compiler V12.1
- Linux with PGI compiler V8.0-4 64-bit
- Linux with INTEL compiler V11.1
- Mac with PGI compiler V10.3-0 64-bit
- Mac with GFORTRAN compiler V4.4.0

Common problems in compilation

- To use gfortran (default 32-bit) on Mac snow leopard system, "-m64" need to be manually appended after "SFC" in $configure.wrf$
- Enough memory is needed to compile some subroutines in WRFPLUS with default optimization level (-O3). Manually reduce the compilation optimization level for some subroutines when system can not allocate enough memory to perform compilation with higher level optimization.
- On some platforms, some compilers might not be able to compile WRF with real_size=8. Usually, upgrading the compiler is the easiest way to solve this problem.

Test for tangent linear model and adjoint model

- After WRFPLUS compilation, It is a good practice to run tangent linear model test and adjoint model test with you own case IC and BC.
- Under $WRFPLUSV3/test/em\_real$ directory, a test case are setup to let users test the tangent linear model and adjoint model.
- In *namelist.input*, turn on $check\_TL$ or $check\_AD$ in *&perturbation* to run tangent linear check or adjoint check.

Test for tangent linear model

Taylor formula:

$$\lim_{\alpha \to 0} \frac{M(x + \alpha\delta\mathbf{x}) - M(x)}{M'(\alpha\delta\mathbf{x})} = 1$$

check results

```
 =============== Tangent Linear check ===================
 check==== U === V === W == PH === T == MU == MOIST =====
 check     T     T     T     T     T     T     T
alpha_m=.1000E+00  coef=   0.98250076417818E+00  val_n= 0.3628649E+11  val_l= 0.3693279E+11
alpha_m=.1000E-01  coef=   0.99781045126907E+00  val_n= 0.3685192E+09  val_l= 0.3693279E+09
alpha_m=.1000E-02  coef=   0.99949153238165E+00  val_n= 0.3691401E+07  val_l= 0.3693279E+07
alpha_m=.1000E-03  coef=   0.10002560538015E+01  val_n= 0.3694225E+05  val_l= 0.3693279E+05
alpha_m=.1000E-04  coef=   0.99981685944643E+00  val_n= 0.3692603E+03  val_l= 0.3693279E+03
alpha_m=.1000E-05  coef=   0.10000972073298E+01  val_n= 0.3693638E+01  val_l= 0.3693279E+01
alpha_m=.1000E-06  coef=   0.99996624597337E+00  val_n= 0.3693154E-01  val_l= 0.3693279E-01
alpha_m=.1000E-07  coef=   0.99999992233716E+00  val_n= 0.3693279E-03  val_l= 0.3693279E-03
alpha_m=.1000E-08  coef=   0.10000017668820E+01  val_n= 0.3693285E-05  val_l= 0.3693279E-05
alpha_m=.1000E-09  coef=   0.10000050602279E+01  val_n= 0.3693298E-07  val_l= 0.3693279E-07
alpha_m=.1000E-10  coef=   0.10000451984913E+01  val_n= 0.3693446E-09  val_l= 0.3693279E-09
```

Test for adjoint model

adjoint identity:

$$\forall \mathbf{x}, \forall \mathbf{y} : \langle M^{'}.\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{M}^{*}.\mathbf{y} \rangle$$

check results

```
ad_check: VAL_TL:     0.41466174569087E+11
ad_check: VAL_AD:     0.41466174569088E+11
```

- Although the tangent linear model might be imperfect.
- The adjoint test must be perfect. otherwise, there are bugs in the adjoint model.

WRF 4D-Var observation preparation

- Conventional observation — LITTLE_R format

  `http://www.mmm.ucar.edu/wrf/users/wrfda/Tutorials/2010_Aug/docs/WRFDA_obsproc.pdf`
- OR Conventional observation — prepbufr format
  - near real-time data : `ftp://ftp.ncep.noaa.gov/pub/data/nccf/com/gfs/prod`
  - history archives : `http://dss.ucar.edu/dataset/ds337.0`
- Satellite radiance bufr data
  - near real-time data : `ftp://ftp.ncep.noaa.gov/pub/data/nccf/com/gfs/prod`
  - history archives : `http://dss.ucar.edu/dataset/ds735.0`
- Ascii formated Radar data

  `http://www.mmm.ucar.edu/wrf/users/wrfda/Tutorials/2010_Aug/docs/WRFDA_radar.pdf`

Tips for using prepbufr and bufr data on non-IBM platforms

On non-IBM platforms,the prepbufr and bufr formats observation downloaded from NCEP ftp server or NCAR achives should be converted. This conversion was conducted using the C code ssrc.c located in the *utils* directory of the GSI distribution.

More detail information and GSI codes download, please refer to

`http://www.dtcenter.org/com-GSI/users/support/faqs/index.php`

- How to compile ssrc.c:

  pgcc -o ssrc.exe ssrc.c

- How to convert :

  ssrc.exe < prepbufr.gdas.2008020600.wo40 > ob.bufr

  ssrc.exe < gdas.1bamua.t00z.20080206.bufr > amusa.bufr

Important namelist variables for 4D-Var run

- *&wrfvar1*
  - *var4d*: logical, if run 4D-Var
  - *var4d_lbc* : logical, if include lateral boundary condition control in 4D-Var
  - *var4d_bin*: integer, seconds, length of sub-window to group observations in 4D-Var
- *&perturbation*
  - *trajectory_io*: logical, do not change, testing purpose
  - *enable_identity* : logical, if run TL/AD model with identity model, testing purpose
  - *jcdfi_use*: logical, if turn on the digital filter as a weak constraint.
  - *jcdfi_diag*: integer, 0/1, J_c term diagnostics
  - *jcdfi_penalty*: real, weight to jcdf term

Important namelist variables for 4D-Var run, cont'd

- *&physics*
  - all physics options must be consistent with which used in wrfinput or fg
- $\&wrfvar18, 21, 22$
  - *analysis_date* is the start time of the assimilation window
  - *time_window_min* is the start time of the assimilation window
  - *time_window_max* is the end time of the assimilation window
- *&time_control*
  - *run_xxxs* must be consistent with the length of the assimilation window
  - *start_xxxx* must be consistent with the start time of the assimilation window
  - *end_xxxx* must be consistent with the end time of the assimilation window

Adjoint check before 4D-Var run

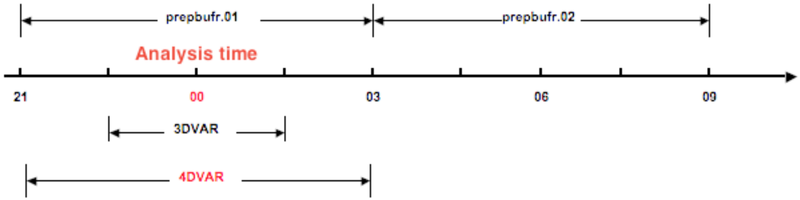
It is always a good practice to run adjoint check before the product run. How:

- `&wrfvar10`
  `test_transforms=true,`
- run
  `da_wrfvar.exe`

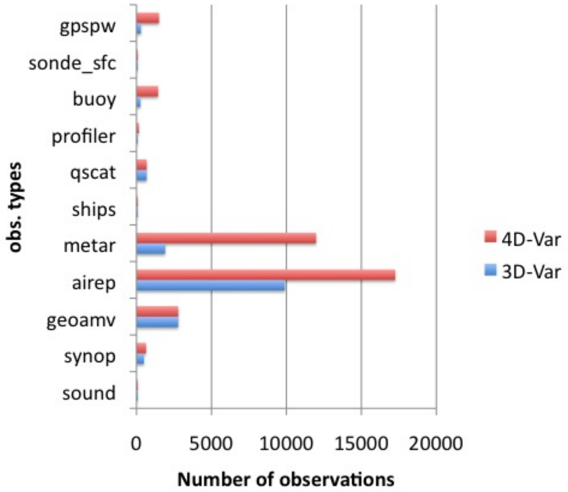
Check results

```
...
 wrf: back from adjoint integrate
 d01 2008-02-05_21:00:00 read nonlinear xtraj time stamp:2008-02-05_21:00:00
 Single Domain < y, y     > =   2.15435506772433E+06
 Single Domain < x, x_adj > =   2.15435506772431E+06

 Whole Domain < y, y     > =   2.15435506772433E+06
 Whole Domain < x, x_adj > =   2.15435506772431E+06

da_check_xtoy_adjoint: Test Finished:

 *** WRF-Var check completed successfully ***
```

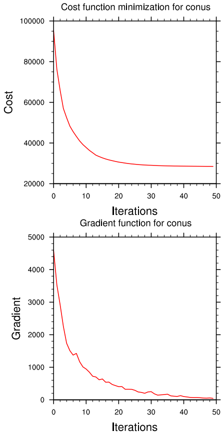
Symmetric 4D-Var window



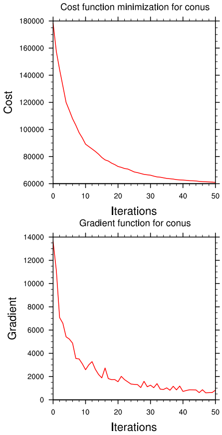
Comparison of obs. usage on 2008020600



Minimization comparison

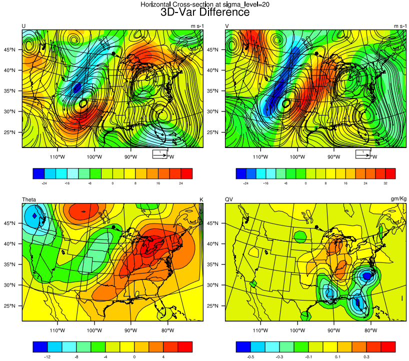


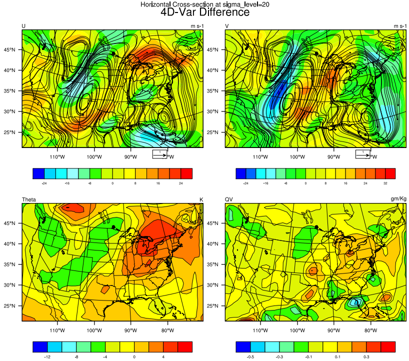
3D-Var



4D-Var

Sample analysis increments valid on 2008020600





Assimilate satellite radiance data

refer to WRFDA Users' guide Chapter 6:
http://www.mmm.ucar.edu/wrf/users/wrfda/Docs/user_guide_V3.3/users_guide_chap6.htm#_Radiance_Data_Assimilations

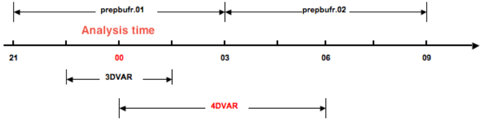
- link/copy amsua data as *amsua.bufr*
- link/copy amsub data as *amsub.bufr*
- Modify namelist.input for radiance data :

```
&wrfvar4
use_amsuaobs=true,
use_amsubobs=true,
&wrfvar14
rtminit_nsensor=6,
rtminit_platform=1,1,1,1,1,1,
rtminit_satid=15,16,18,15,16,17,
rtminit_sensor=3,3,3,4,4,4,
thinning_mesh=120.0,120.0,120.0,120.0,120.0,120.0,
thinning=true,
qc_rad=true,
rtm_option=2,
use_varbc=true,
use_crtm_kmatrix=true,
```

Additional links for radiance assimilation

- link/copy amsua data as $amsua.bufr$
- link/copy amsub data as $amsub.bufr$
- $link -fs\ WRFDA/var/run/radiance\_info\ radiance\_info$
- $link -fs\ WRFDA/var/run/crtm\_coeffs\ crtm\_coeffs$

One-side 4D-Var window



- link/copy prepbufr data at 00Z as *ob01.bufr*
- link/copy prepbufr data at 06Z as *ob02.bufr*
- link/copy amsua data at 00Z as *amsua01.bufr*
- link/copy amsua data at 06Z as *amsua02.bufr*
- ...

Common problems in WRF 4D-Var run

Error message

```
**************BUFR ARCHIVE LIBRARY ABORT*****************
BUFRLIB: OPENBF - ERROR READING INPUT FILE CONNECTED TO UNIT  96 WHEN CHECKING
FOR 'BUFR' IN FIRST 4 BYTES OF RECORD
**************BUFR ARCHIVE LIBRARY ABORT*****************
```

- Solution: prepbufr and/or bufr data should be converted for non-IBM platforms.

Error message, PGI compiler only

```
0: ALLOCATE: 18446744072053605056 bytes requested; not enough memory
```

- Solution: It might be a bug, under investigation. Please try with another compiler.

Developments after V3.3

Finished:

- 3 physics schemes were added in WRF tangent linear model and adjoint model.
  - surface drag (bl_pbl_physics=98)
  - large scale condensate (mp_physics=98)
  - a simplified cumulus scheme (cu_physics=98)
- Parallelization of WRF tangent linear model is done.

Under development:

- Parallelization of WRF adjoint model.
- Add precipitation observation to forcing term.
- Different resolutions in outer loops and inner loops.

Thank You

The NESL Mission is:
To advance understanding of weather, climate, atmospheric composition and processes;
To provide facility support to the wider community; and,
To apply the results to benefit society.

NCAR is sponsored by the National Science Foundation