2015 GSI Community Tutorial
August 11-13, 2013, NCAR, Boulder

Background Error, Observation Error, and GSI Hybrid Analysis

Ming Hu

Developmental Testbed Center

Outlines

- GSI fundamentals (1): Setup and Compilation
- GSI fundamentals (2): Run and Namelist
- GSI fundamentals (3): Diagnostics
- GSI fundamentals (4): Applications
- Background Error, Observation Error, and GSI Hybrid Analysis
  - Background error covariance
  - Variation-ensemble hybrid analysis
  - Observation error

GSI input and output files

Analysis results; stdout and diagnostic files, …

$$J(x) = (x-x_b)^T \boldsymbol{B}^{-1}(x-x_b)+(y-\mathrm{H}[x])^T \boldsymbol{R}^{-1}(y-\mathrm{H}[x]) + J_c$$

Background file

Observations files

Background covariance file

CRTM coefficients

Observation error file

- **Control and state variable info file**
- Observation data control info files

Fixed files

- **Namelist**
- Radiance bias correction coefficient files

DTC Developmental Testbed Center

Background Error (BE) Covariance

What BE does and structure of a GSI BE ?

Options to tune a GSI BE?

Generate your own BE: GenBE

Background Error Covariance (BE)

- In 3dVar analysis, background error covariance (BE) is the most important and complex part, which includes three pieces of information:
  - **Variance**: the quality of the background, ratio of background error variance and observation error variance decide the fit of analysis to the observation.
  - **Horizontal and vertical impact scale**: decide the area and depth of the observation information should spread
  - **Balance**: relation between different analysis variables, such as temperature and wind

GSI BE structure

- BE is huge matrix that cannot be calculate explicitly. In GSI, the **B** matrix is decomposed into the following form

$$B = B_{balance} V B_Z (B_x B_y B_y B_x) B_Z V B^T_{balance}$$

Balance among different variables represented with pre-computed "regression coefficients"

Pre-calculated square root of variance

Horizontal impact is modeled as a Gaussian distribution with pre-computed "lengthscale" parameters in recursive filter

Vertical impact is modeled as a Gaussian distribution with pre-computed "lengthscale" parameters in recursive filter

GSI BE structure

$B_{balance}$

$B^T_{balance}$

B_z

$(B_x B_y B_y B_x)$

B_z

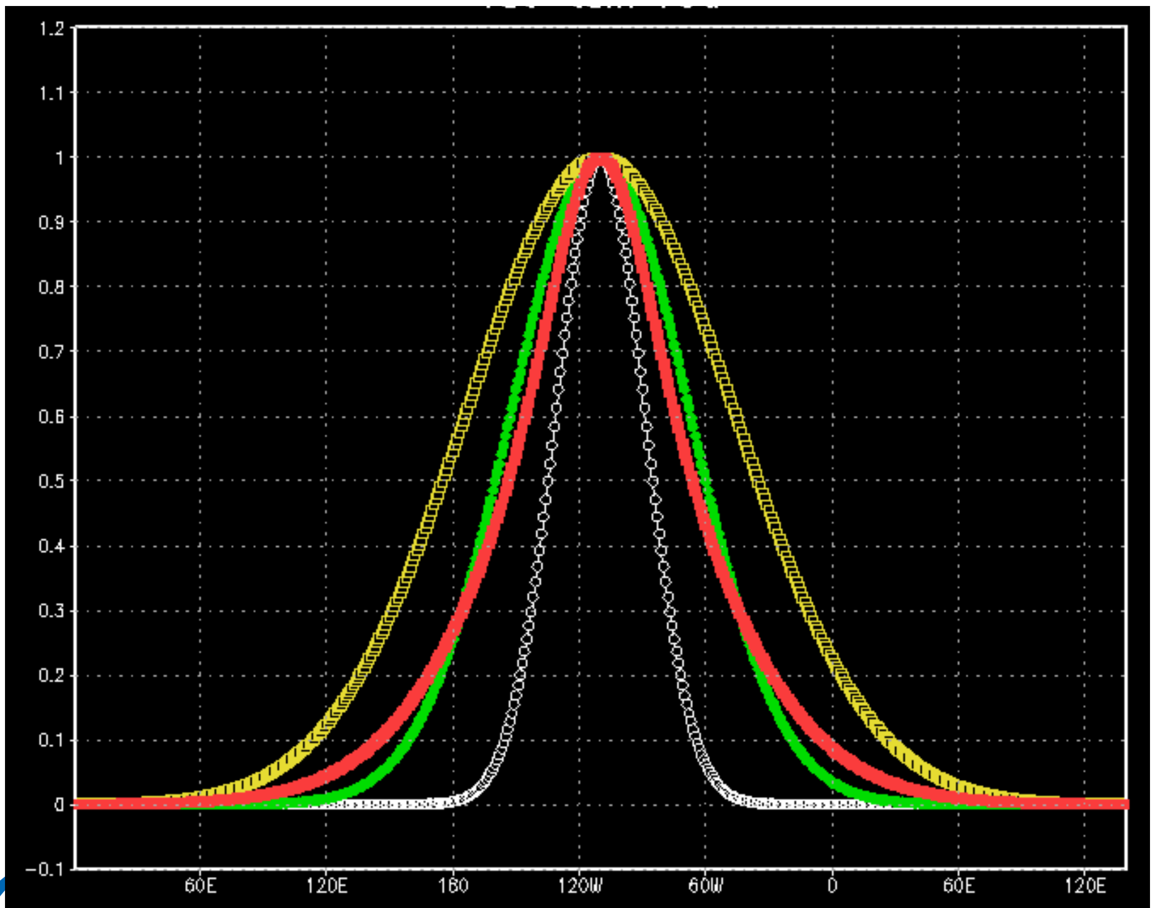
V

Category	Array name	Dimension	Content
Balance (Horizontal regression coefficients)	*agvi*	0:mlat+1,1:nsig,1:nsig	Regression coefficients for stream function and temperature
	wgvi	0:mlat+1,1:nsig	Regression coefficients for stream function and surface pressure
	bvi	0:mlat+1,1:nsig	Regression coefficients for stream function and velocity potential
Horizontal and vertical influence scale	*hwll*	0:mlat+1,1:nsig,1:nc3d	horizontal lengthscales for stream function, unbalanced velocity potential, unbalanced temperature, and relative humidity
	hwllp	0:mlat+1, nc2d	horizontal lengthscale for unbalanced surface pressure
	vz	1:nsig, 0:mlat+1, 1:nc3d	Vertical lengthscale for stream function, unbalanced velocity potential, unbalanced temperature, and relative humidity
variance	*corz*	1:mlat,1:nsig,1:nc3d	Square root of variance for stream function, unbalanced velocity potential, unbalanced temperature, and relative humidity
	corp	1:mlat,nc2d	Square root of variance for unbalanced surface pressure

Note: mlat = number of latitude in original background error coefficient domain,
nsig = number of vertical levels in analysis grid
nc3d = number of 3 dimensional analysis variables
nc2d = number of 2 dimensional analysis variables

Fat-tailed Power Spectrum for horizontal impact

Horizontal impact is modeled by the combination of three recursive filters that have different impact scales



Yellow: filter 1 (wide)

Green: filter 2 (middle)

White: filter 3 (narrow)

Red: combined to make fat-tail impact pattern

Courtesy of Wan-Shu Wu

GSI BE in release Package

- *nam_nmmstat_na.gcv* : contains the regional background error statistics
  - Computed using forecasts from the NCEP's NAM model covering North America.
  - Covers the northern hemisphere with 93 latitude lines from -2.5 degree to 89.5 degree
  - With 60 vertical sigma levels from 0.9975289 to 0.01364.

- *nam_glb_berror.f77.gcv* : contains the global background errors
  - Based on the NCEP's GFS model, a global forecast model.
  - Covers global with 192 latitude lines from -90 degree to 90 degree and
  - With 42 vertical sigma levels from 0.99597 to 0.013831.

- Global BE

BE Tuning Options

- In GSI namelist:

`vs`	scale factor for vertical correlation lengths for background error
`nhscrf`	number of horizontal scales for recursive filter
`hzscl(3)`	scale factor for horizontal smoothing. Specifies factor by which to reduce horizontal scales (i.e. 2 would then apply 1/2 of the horizontal scale)
`hswgt(3)`	empirical weights to apply to each horizontal scale

- In anavinfo
  - Column "**as**" under section "**control_vector::**"

```
control_vector::
!var       level   itracer as/tsfc_sdv  an_amp0   source   funcof
 sf         40       0       1.00        -1.0     state     u,v
 vp         40       0       1.00        -1.0     state     u,v
 ps          1       0       0.50        -1.0     state     prse
 t          40       0       0.70        -1.0     state     tv
 q          40       1       0.70        -1.0     state     q
```

Developmental Testbed Center

Example values for BE Tuning options

- Run the GSI executable

	Global	Regional
fixed B matrix	nam_glb_berror.f77.gcv	nam_nmmstat_na.gcv
vs	0.7	1.0
hzscl	1.7, 0.8, 0.5	0.373,0.746,1.50
hswgt	0.45, 0.3, 0.25	0.45, 0.3, 0.25
ss/tsfc_sdv	control_vector:: !var as/tsfc_sdv sf 0.60 vp 0.60 ps 0.75 t 0.75 q 0.75 oz 0.75 sst 1.00 cw 1.00 stl 3.00 sti 3.00	control_vector:: !var as/tsfc_sdv sf 1.00 vp 1.00 ps 0.50 t 0.70 q 0.70 oz 0.50 sst 1.00 cw 1.00 stl 1.00 sti 1.00

BE tuning examples

- Through **single observation test**, we can how those BE tuning option works
  - Check fundamental talks on setting up single observation test
- Regional BE "*nam_nmmstat_na.gcv*" is used
- In GSI namelist:
  - Horizontal impact scale of BE before tuning:
    hzscl_op='0.373,0.746,1.50,'
  - Vertical impact scale of BE before tuning:
    vs_op='1.0,'

Courtesy of Min Sun

Single Obs(U) Test with BE tuning

Single observation test parameters

Parameter	Value
ΔU	1m/s
Observation Error	2m/s
Horizontal Location	Domain Center
Vertical Location	700hPa

Horizontal impact scale tuning

Horizontal Test	'hzscl_op' Value
Original	0.373,0.746,1.50
1/2	0.1865, 0.373, 0.75
1/4	0.09325,0.1865,0.375
1/8	0.046625,0.09325,0.1875
1/16	0.0233125,0.046625,0.09375
1/32	0.01165625,0.0233125,0.046875

Vertical impact scale tuning

Vertical Test	'vs' Value
original	1.0
1/2	0.5
1/4	0.25

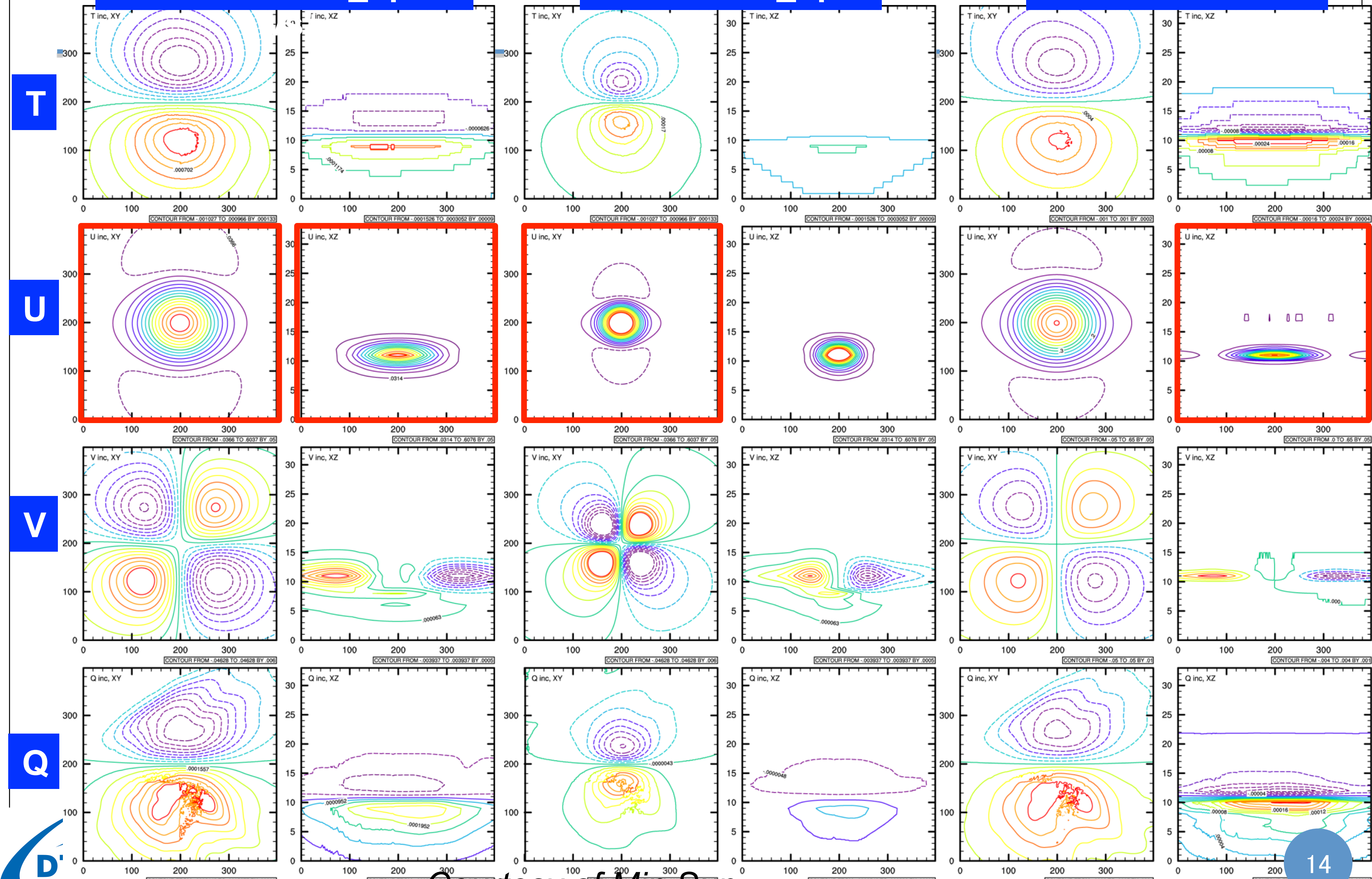
Courtesy of Min Sun

Analysis increment

Default ‘hzscl_op’ &

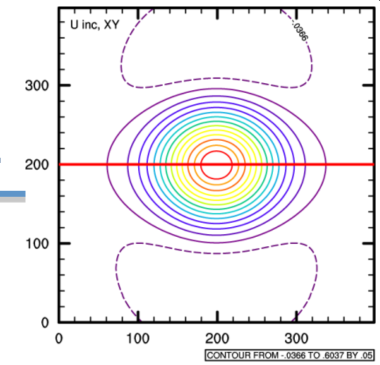
1/2 ‘hzscl_op’

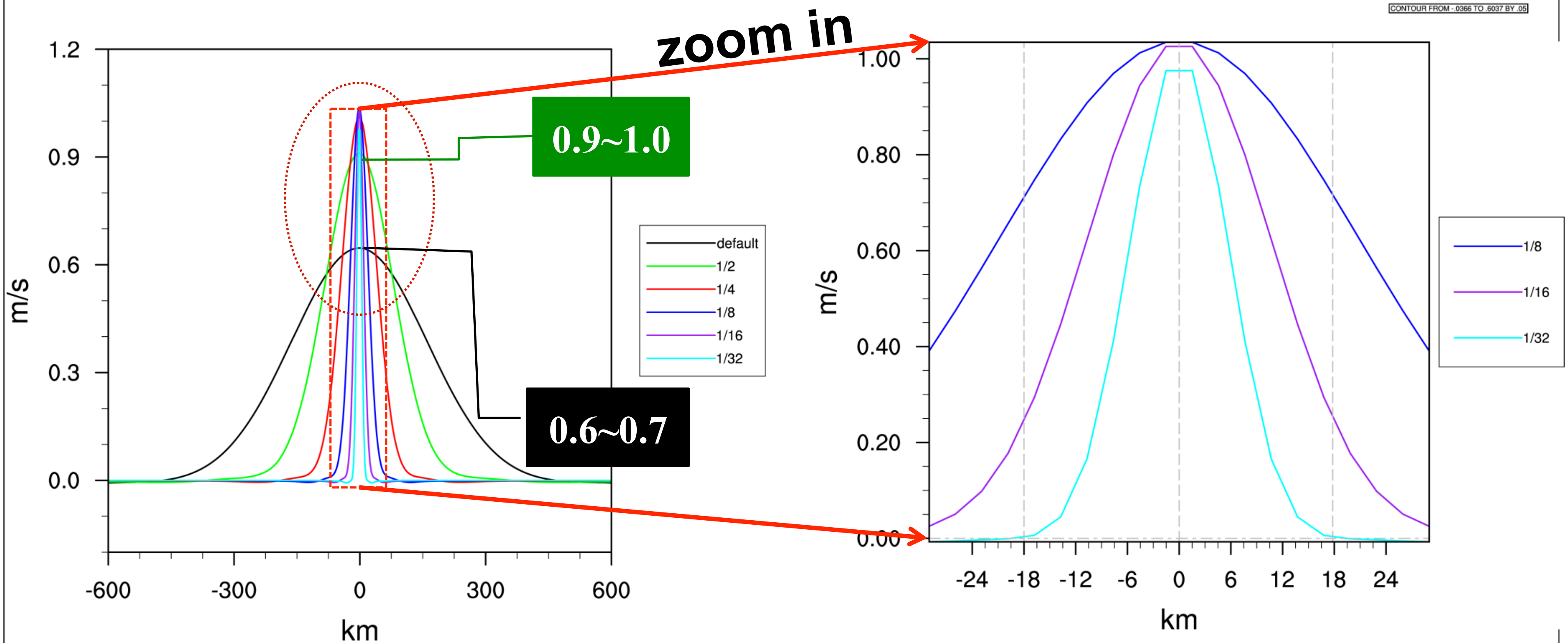
1/2 ‘vs’



Courtesy of Min Sun

Horizontal Cross Section of U Increment

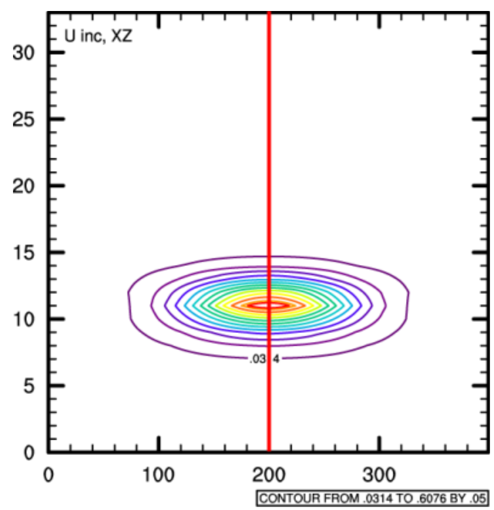


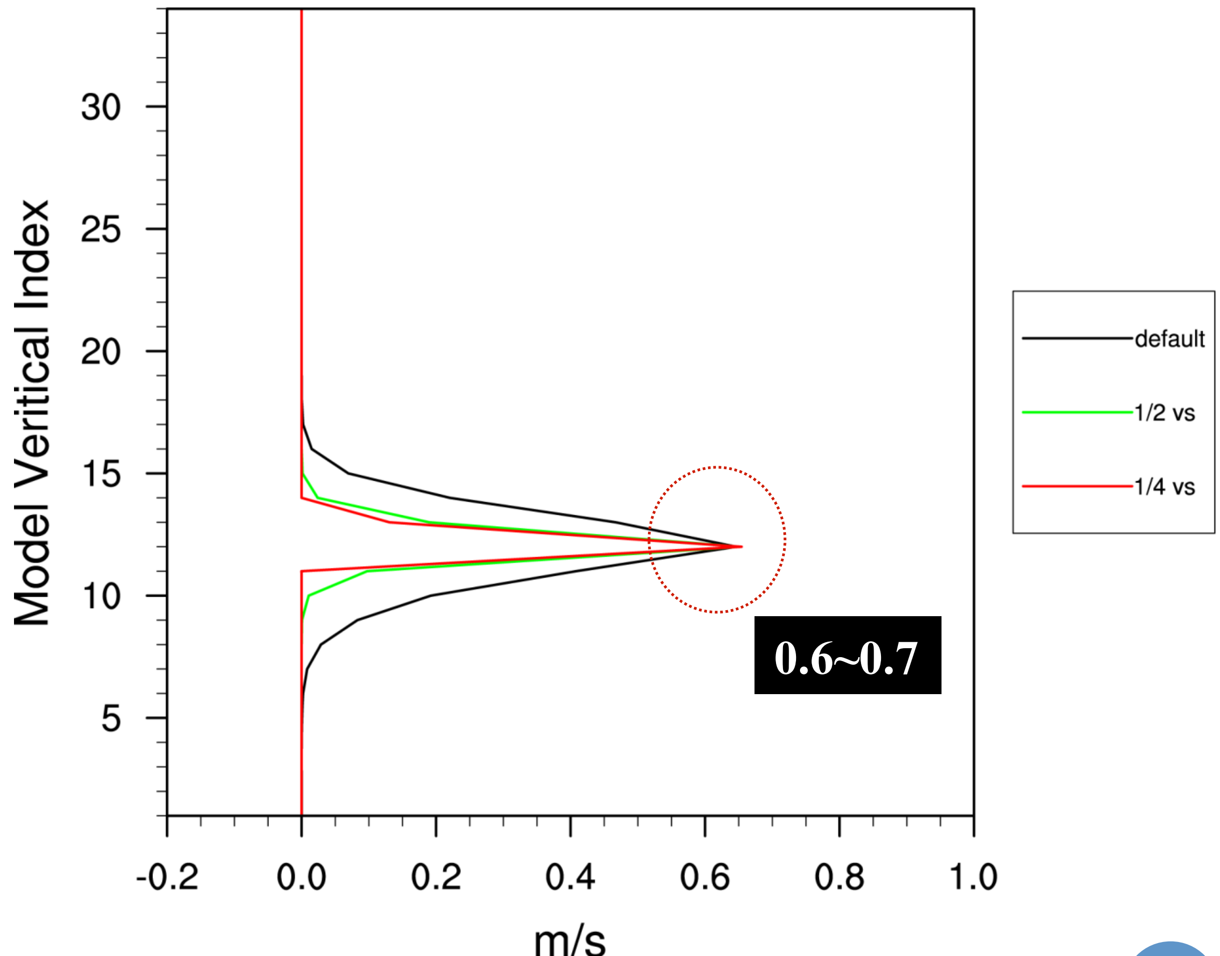


Change 'hzscl_op' not only change the horizontal influence scale, but also the weight (how much analysis results fit to the observations)!

Courtesy of Min Sun

Vertical Cross Section of U Analysis Increment

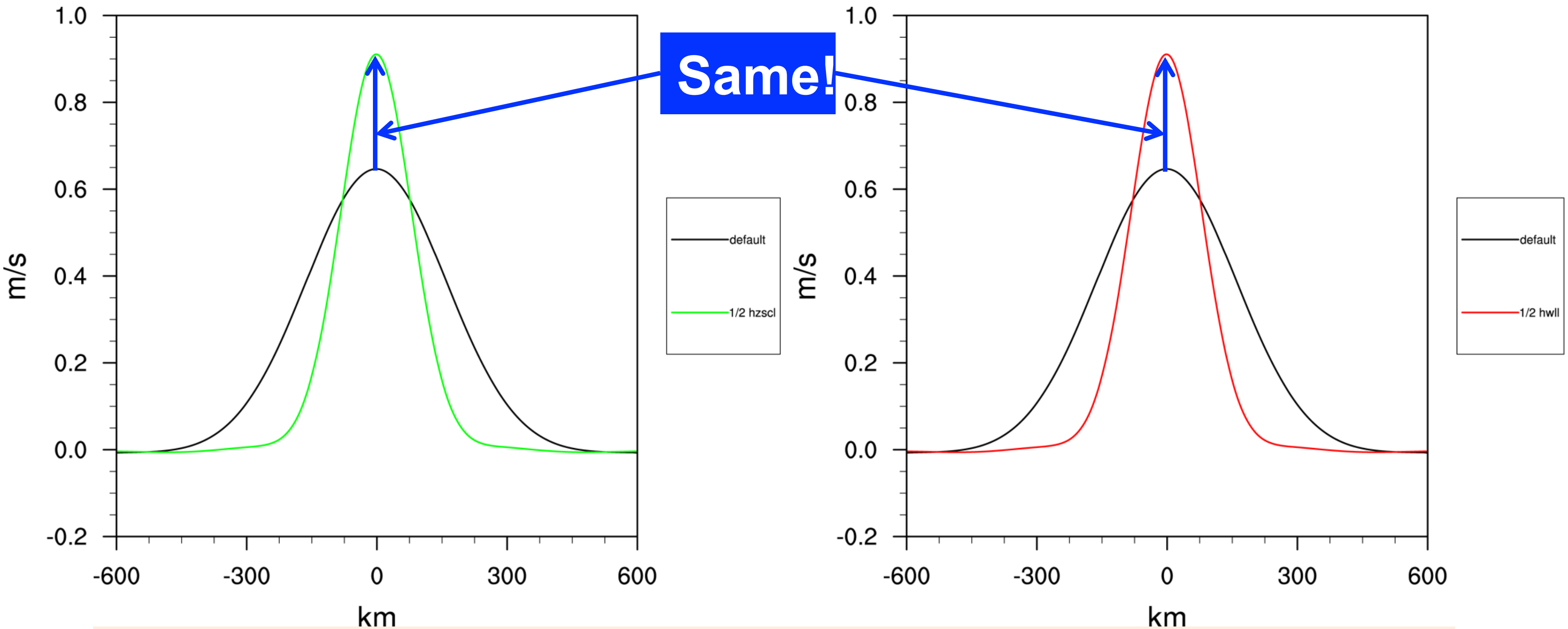




Courtesy of Min Sun

Two Ways to Change Horizontal Influence Scale

- **By changing parameter 'hzscl_op' in namelist**
- **By reading BE matrix directly and changing 'hwll' array which stores horizontal influence scales**



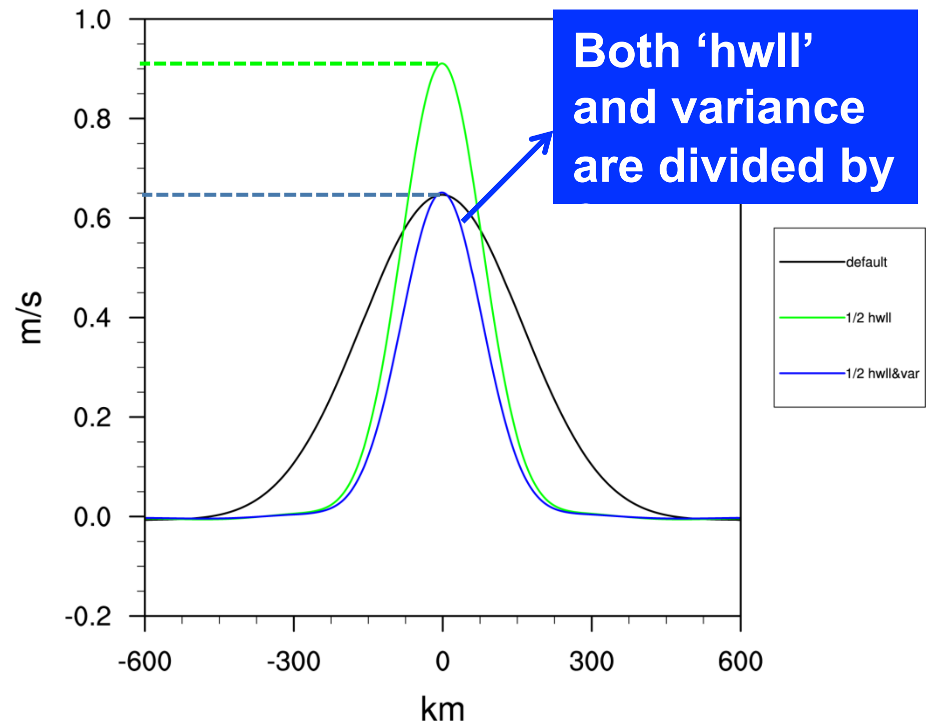
How to change the horizontal influence scale, meanwhile do not affect the weight ?

Courtesy of Min Sun

Change Variance in BE

How to change the horizontal influence scale, meanwhile do not influence the weight ?

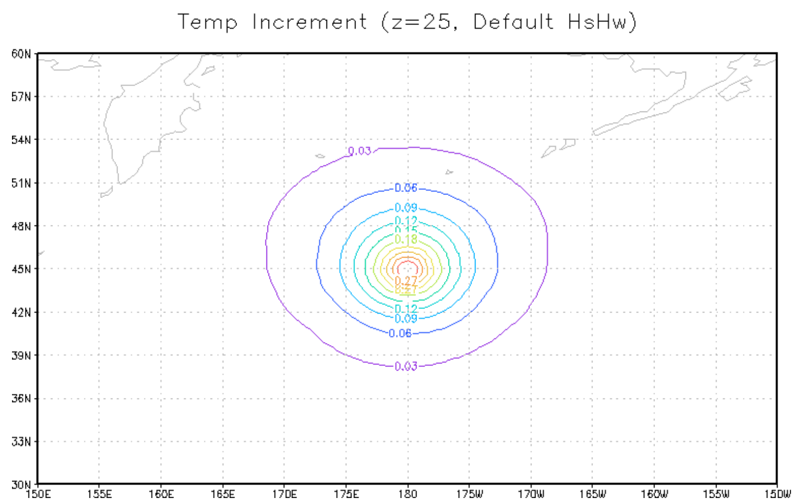
Change the arrays which store horizontal influence scales ‘hwll’ and variance of BE simultaneously!

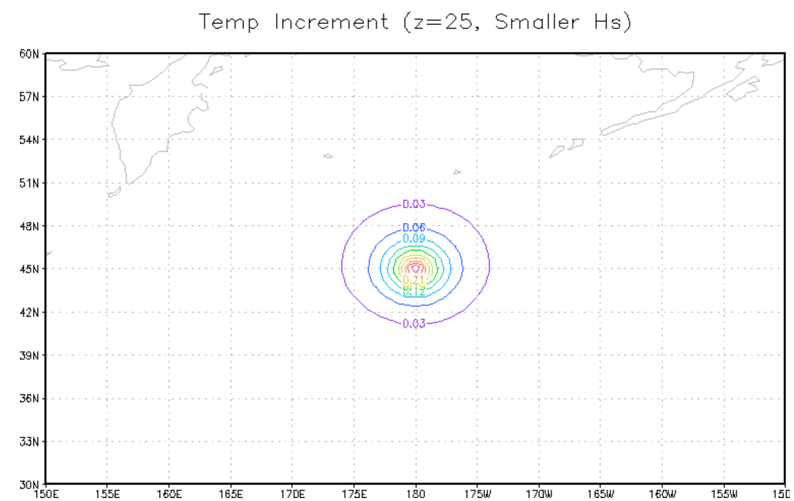


Courtesy of Min Sun

Example from Daryl Klest's talk

Tuning Example (Scales)





Hzscl = 1.7, 0.8, 0.5

Hswgt = 0.45, 0.3, 0.25

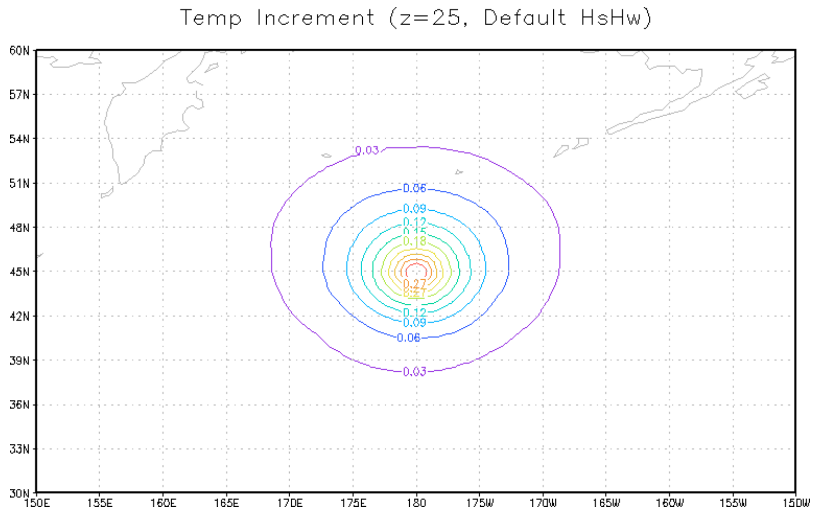
Hzscl = 0.9, 0.4, 025

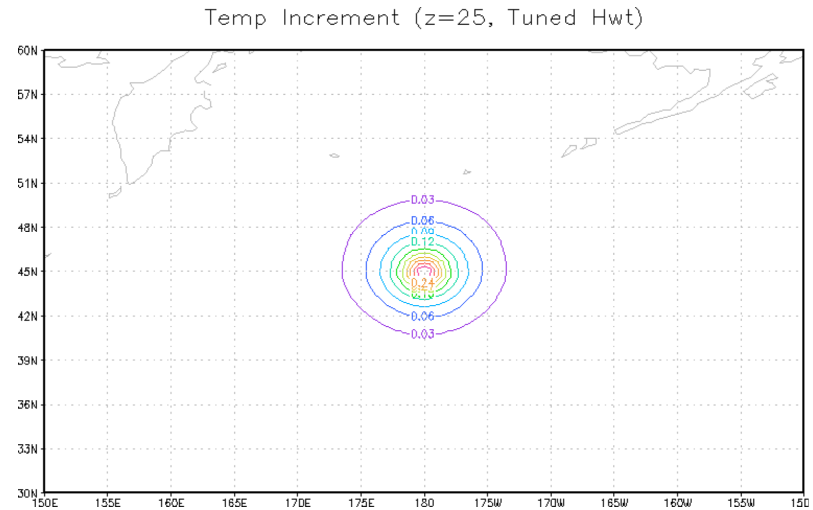
Hswgt = 0.45, 0.3, 0.25

500 hPa temperature increment (K) from a single temperature observation utilizing GFS default (left) and tuned (smaller scales) error statistics.

Example from Daryl Klest's talk

Tuning Example (Weights)





Hzscl = 1.7, 0.8, 0.5	Hzscl = 1.7, 0.8, 0.5
Hswgt = 0.45, 0.3, 0.25	**Hswgt = 0.1, 0.3, 0.6**

500 hPa temperature increment (K) from a single temperature observation utilizing GFS default (left) and tuned (weights for scales) error statistics.

GEN-BE

- Reference 2012 GSI tutorial

- A paper on New GEN-BE:
  - Gael Descombes and Tom Auligne (2015)

Variation-Ensemble hybrid

Setup hybrid runs

Namelist options for hybrid

Results from RAP experiments

Hybrid Variational Ensemble Data Assimilation

$$J(x) = (x-x_b)^T B^{-1}(x-x_b)+(y-H[x])^T R^{-1}(y-H[x])$$
$$= J_b + J_o$$

- ***B*** is the background error covariance matrix.
- 3D-Var uses static ("climate") background errors
- Hybrid DA uses flow dependent background error information from an ensemble in a variational DA system
- Replace B by a weighted sum of 3D-VAR B and ensemble covariance

$$\mathbf{B} = \alpha_1 \mathbf{B}_1 + \alpha_2 \mathbf{B}_2$$
$$\alpha_1 = 1 - \alpha_2$$

DTC Developmental Testbed Center

Courtesy of Terra Ladwig

Setup GSI hybrid

- *Step 1: Link the ensemble members to the GSI run directory*

regional_ensemble_option	explanation	GSI recognized ensemble file names
1	GFS ensemble internally interpolated to hybrid grid	*filelist : a text file include path and name of ensemble files*
2	ensembles are in WRF NMM (HWRF) format	*d01_en001,* *d01_en002,* …
3	ensembles are in ARW netcdf format	*wrf_en001,* *wrf_en002,* …
4	ensembles are in NEMS NMMB format	*nmmb_ens_mem001,* *nmmb_ens_mem002,* …

Example in GSI run script:

```
if [ -r  ${mempath}/wrfout_d01_${iiimem} ]; then
   ln -sf ${mempath}/wrfout_d01_${iiimem} ./wrf_en${iiimem}
else
   echo "member ${mempath}/wrfout_d01_${iiimem} does not exit”
fi
```

Setup GSI hybrid: basic options

Options	explanation
l_hyb_ens	if true, turn on hybrid ensemble option
n_ens	number of ensemble members
beta1_inv	the weight given to the static background error covariance: 0 <= *beta1_inv* <= 1; *1 - beta1_inv* is the weight given to the ensemble derived covariance =1, ensemble information turned off =0, static background errors turned off
regional_ensemble_option	integer, used to select type of ensemble to read in for regional applications. Currently takes values from 1 to 4: =1: use GEFS internally interpolated to ensemble grid. =2: ensembles are in WRF NMM format =3: ensembles are in ARW netcdf format. =4: ensembles are in NEMS NMMB format.

Setup GSI hybrid: further tuning

Options	explanation
s_ens_h	homogeneous isotropic horizontal ensemble localization scale (km)
s_ens_v	vertical localization scale: • If positive, in grid units; • if negative, in lnp unit.
grid_ratio_ens	for regional runs, ratio of ensemble grid resolution to analysis grid resolution. If turned on and specified an appropriate value, could increase the computational efficiency.

RAP GSI hybrid with GFS Ensemble

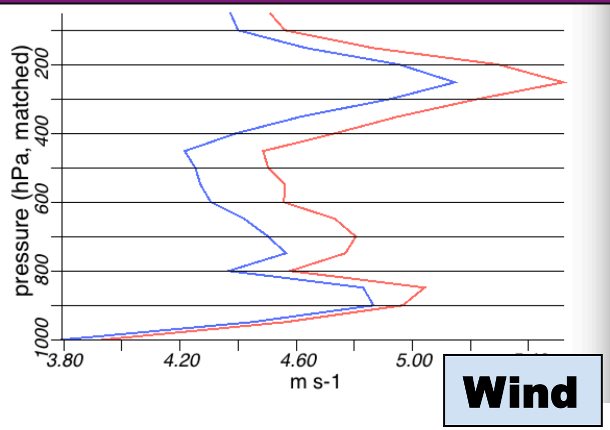
- RAPv2 hybrid configurations:
  - With **half** Ensemble BE and **half** Static BE
  - Ensemble grid is **3 times coarser** than background grid
  - Ensemble forecasts are available every **6-hour**
  - Horizontal localization scale is **110** km
  - Vertical localization scale is **3** grid levels
  - Use GFS ensemble

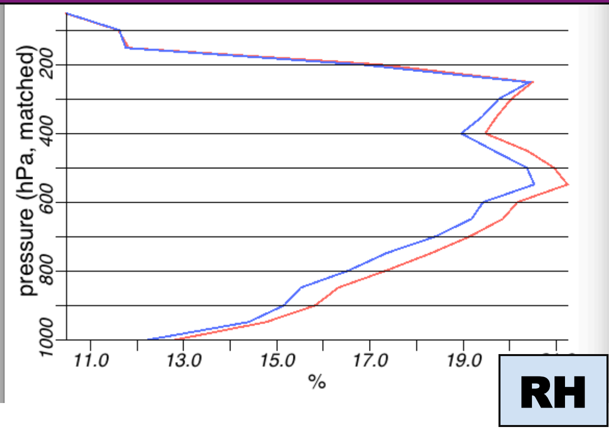
- Baseline retrospective tests
  - May 28th to June 4th, 2012
  - Only difference are analysis: 3DVar versus Hybrid

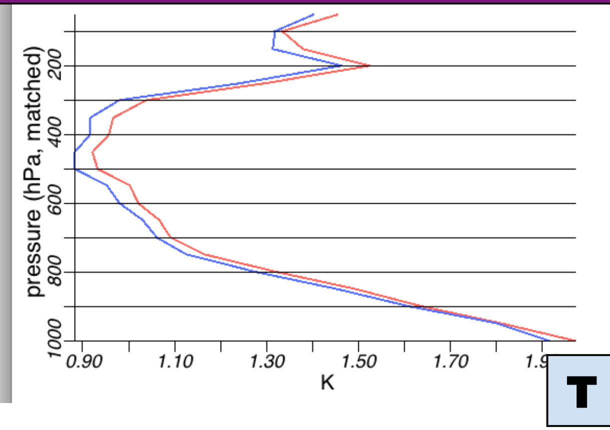
RAPv2 baseline test results

RAP Hybrid (with GFS Ens) **RAP No Hybrid (3D-VAR)**

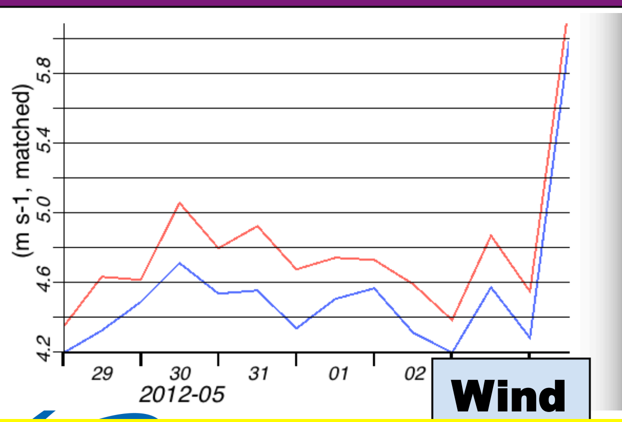
Upper Air RMS Vertical Profile for 6 hour forecast

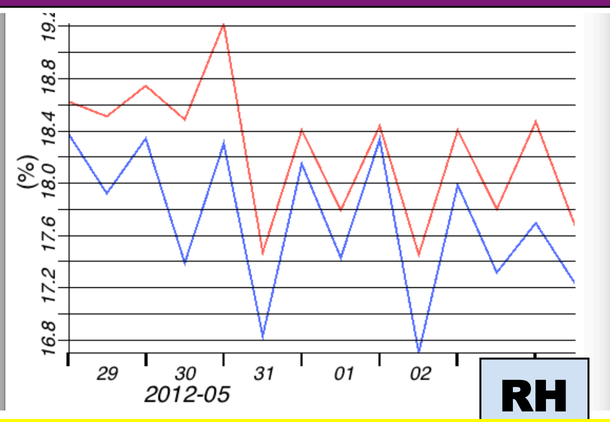


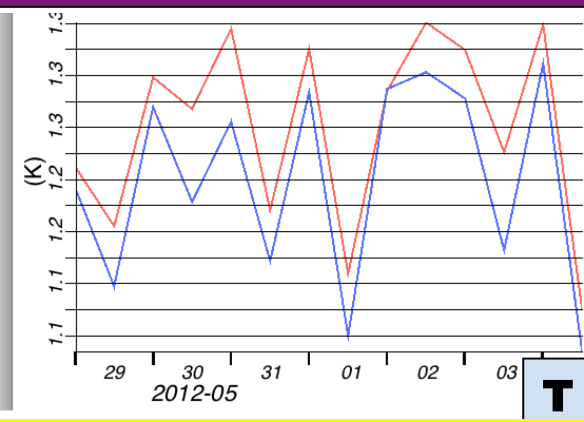




Upper Air RMS Time Series for 6 hour forecast





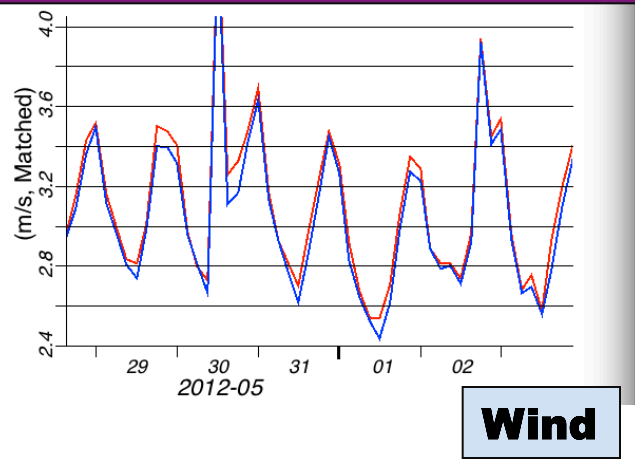


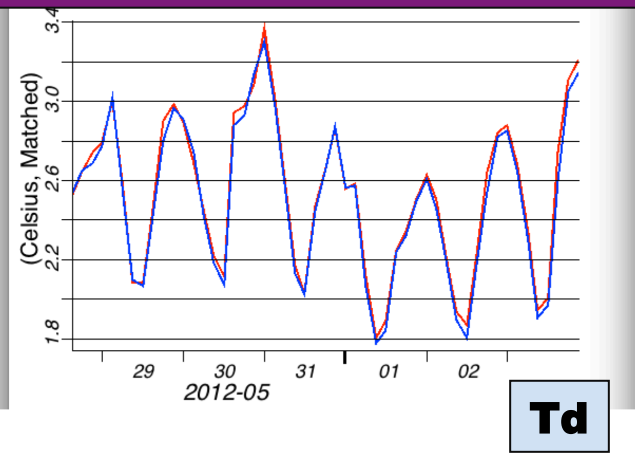
Consistent improved upper-air environment
Little impact to the ceiling forecast, surface forecast, precip forecast

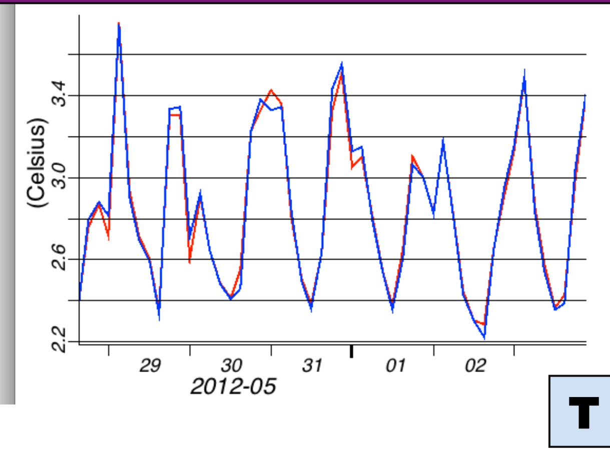
RAPv2 baseline test results

RAP Hybrid (with GFS Ens) **RAP No Hybrid (3D-VAR)**

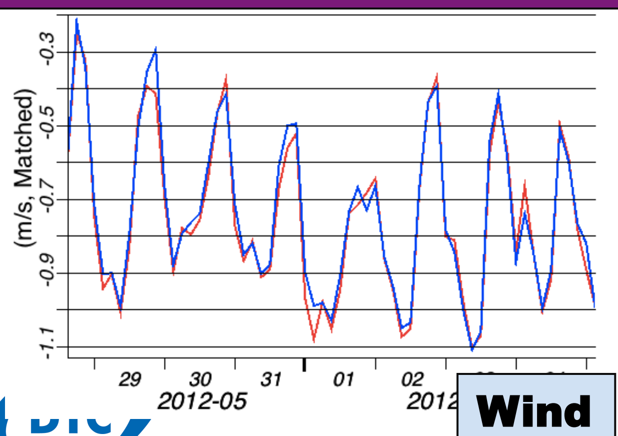
Surface RMS Time Series for 6 hour forecast

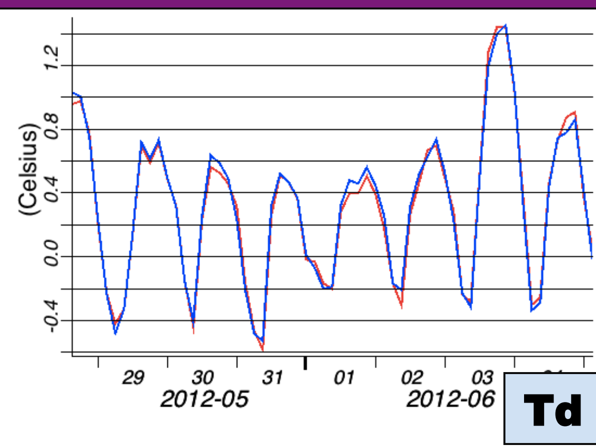


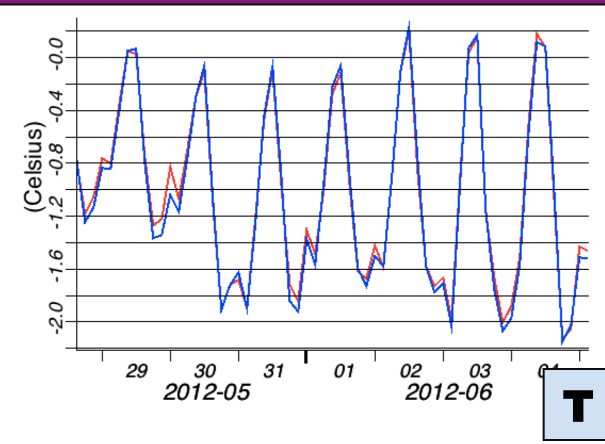




Surface BIAS Time Series for 6 hour forecast







Observation Error

Observations errors

External observation error table for conventional obs

Satellite radiance observation error

Radar radial wind observations error

Adaptive Tuning of Observation Error

Observation errors

- Each observation has to have observation error to be used in the variational analysis
- Observation error is unrelated and its variance represents the quality of this observation.
- The ratio of observation error and background error decide how much analysis results fit to the observation value.
- Each type of observations has their own way to find error:
  - Conventional observation:
    - Errors are read in from the PrepBUFR file for Global analysis and when "*oberrflg*" set to false.
    - Errors are generated based on an external observation error table for regional analysis or for global analysis and when "*oberrflg*" set to true.
  - Satellite radiance
  - Radar radial wind
  - …

External observation error table

```
Column #
      1           2           3           4           5           6
  120 OBSERVATION TYPE
  0.11000E+04 0.12671E+01 0.56103E+00 0.10000E+10 0.68115E+00 0.10000E+10
  0.10500E+04 0.13302E+01 0.63026E+00 0.10000E+10 0.68115E+00 0.10000E+10
  0.10000E+04 0.14017E+01 0.73388E+00 0.10000E+10 0.68115E+00 0.10000E+10
  0.95000E+03 0.14543E+01 0.86305E+00 0.10000E+10 0.71307E+00 0.10000E+10
  0.90000E+03 0.14553E+01 0.99672E+00 0.10000E+10 0.74576E+00 0.10000E+10
  0.85000E+03 0.13865E+01 0.11210E+01 0.10000E+10 0.77845E+00 0.10000E+10

  220 OBSERVATION TYPE
  0.11000E+04 0.10000E+10 0.10000E+10 0.17721E+01 0.10000E+10 0.10000E+10
  0.10500E+04 0.10000E+10 0.10000E+10 0.20338E+01 0.10000E+10 0.10000E+10
  0.10000E+04 0.10000E+10 0.10000E+10 0.22927E+01 0.10000E+10 0.10000E+10
  0.95000E+03 0.10000E+10 0.10000E+10 0.24559E+01 0.10000E+10 0.10000E+10
  0.90000E+03 0.10000E+10 0.10000E+10 0.25377E+01 0.10000E+10 0.10000E+10
  0.85000E+03 0.10000E+10 0.10000E+10 0.25705E+01 0.10000E+10 0.10000E+10
```

Column #	1	2	3	4	5	6
Content	Pressure	T	q	UV	Ps	Pw
Unit	hPa	degree C	percent/10	m/s	mb	kg/m^2(or mm)

For each observation, error is from a vertical interpolation based on error table

Satellite radiance observation error

- In *satinfo* file, observation error is set based on sensor_satellite and channel:

```
!sensor/instr/sat        chan  iuse   error   error_cld   ermax    var_b     var_pg   cld_det
 amsua_n15                  1     1   3.000     9.100     4.500   10.000     0.000      -2
 amsua_n15                  2     1   2.000    13.500     4.500   10.000     0.000      -2
 amsua_n15                  3     1   2.000     7.100     4.500   10.000     0.000      -2
 amsua_n15                  4     1   0.600     1.300     2.500   10.000     0.000      -2
...
 amsua_n15                 15     1   3.000    10.000     4.500   10.000     0.000      -2
 hirs3_n17                  1    -1   2.000     0.000     4.500   10.000     0.000      -1
 hirs3_n17                  2    -1   0.600     0.000     2.500   10.000     0.000       1
...
 hirs3_n17                  9    -1   1.100     0.000     3.500   10.000     0.000      -1
```

Observation error for clear radiance

Observation error for cloudy radiance

Radar radial wind observation error

- Level II radial wind observations are superobed as new radial wind observations and the observation error for new radial velocity is:

$$error = \sqrt{\left| \overline{V_r^2} - \overline{V_r}^2 \right|}$$

where, Vr is a vector includes all level-II radial wind observations in a superob box.

- The observation error can be inflated through a namelist variable "*erradar_inflate*" in section */obsqc/*. The default value is 1.

Adaptive Tuning of Observation Error

- Talagrand (1997) on $E(J(X^a))$
- Desroziers & Ivanov (2001)

  $E(J^o) = \frac{1}{2} \mathrm{Tr}(I_p - HK)$

  $E(J^b) = \frac{1}{2} \mathrm{Tr}(KH)$

  where I_p is identity matrix with order p

  K is Kalman gain matrix

  H is linearized observation forward operator
- Chapnik et al.(2004): robust even when B is incorrectly specified

More details please see Wan-Shu Wu's talk in 2013 GSI summer tutorial:
"*Background and Observation Error Estimation and Tuning*"

Courtesy of Wan-Shu Wu

Questions? …

gsi_help@ucar.edu