Generation of WRF-ARW Background Errors (BE) For GSI

Syed RH Rizvi, Zhiquan Liu and Xiang-Yu Huang
NCAR/ESSL/MMM, Boulder, CO

1. Introduction:

Currently NCEP global background errors are used for running GSI for its regional applications. These global background errors are produced using NMC method with GFS forecasts. Keeping in view that GSI needs to be run with WRF-ARW forecast as its first guess, it is desired to generate background errors for WRF-ARW model. Depending upon domain configuration, since the characteristics of regional WRF-ARW forecasts is expected to be different than GFS forecast, it is a strong motivation for developing the WRF-ARW domain specific background errors instead of using globally generated GFS background errors.

This document describes the efforts made in developing the code for generation of WRF-ARW background error (BE) statistics to be used by GSI. This may also serve as a technical documentation for this code, which will hereafter be referred as "WRF-ARW-GSI_BE". For regional applications of GSI, the input background error (BE) information needs to be interpolated suitably to the desired regional domain configuration. For this purpose, the interpolation (both horizontal & vertical) strategy followed within GSI is also discussed. It also describes the role of various "tuning" parameters available within GSI to adjust the input BE at run time via its namelist parameters. "Remarks" are added to highlight some important differences in this code and "gen_be". It also highlights some special features in GSI code in regards to final utilization of BE statistics.

During the process of writing and testing this code, some useful graphics (NCL based) have also been developed to display the BE statistics used in GSI. Currently "oneobtest" utility of GSI uses single observation location in terms of Latitude/Longitude and pressure levels. For better understanding of BE structures, GSI has been updated to carry out single observation test located exactly at the model grid point. Keeping in view that if BE is generated for certain domain with some specific domain configuration and later GSI is run for the same domain then it might be more appropriate to utilize this BE as such without any vertical interpolation. To ensure this, a new namelist parameter (interpolate_reg_stats) has been introduced in GSI to bypass all the GSI built-in vertical interpolation procedure.

WRF-ARW-GSI_BE code has been tested extensively in CONUS (200 Km) and T44 (45 Km) domains both in single and multi-processor environment. Through a variety of single observation tests, response of BE produced with WRF-ARW-GSI_BE and NCEP global BE is compared for these two domains. Suitable tuning factors have also been worked out for these two domains.

Here it may be noted that currently for regional applications of WRFDA with "CV_OPTION=3", "be.cv_3" BE statistics is used. This BE statistics is generated using

NCEP GFS forecasts. It was provided by NCEP more than seven years back and so it is very old. In principle, WRF-ARW-GSI_BE code may be used to generate WRF-ARW BE for regional applications of WRFDA for its "CV_OPTION=3".

This document is organized as follows. Section 2 describes components of GSI BE statistics. Section 3 describes technical aspect related with the development strategy of "WRF-ARW-GSI_BE" code. Section 4 describes how BE statistics is utilized in GSI code with reference to its built-in strategy about interpolation and tuning its different control variables. Section 5 & 6 is devoted to results with CON200 (200 Km) and T44 (45 Km) domains. It gives a detailed comparison of the response of BE statistics produced by this code and NCEP global BE statistics. An appendix is also added to explain how this code, related scripts and graphics are organized within WRFDA. Main functions for some of the important subroutines used in this code are also discussed in this Appendix.

2. Components of GSI BE statistics:

GSI background errors (BE) statistics consist of the following.

a) Vertical sigma levels & latitude values for which the BE is available.
b) Horizontal length-scales for stream function (ψ), unbalanced velocity potential (χ_u), unbalanced temperature (t_u), relative humidity (rh) and unbalanced surface pressure (ps_u). Currently these are the basic analysis control variables (CVT) for GSI.
c) Vertical length-scales for stream function (ψ), unbalanced velocity potential (χ_u), unbalanced temperature (t_u) and relative humidity (rh).
d) Regression coefficients for χ_u, t_u and ps_u with ψ field.
e) Variance of ψ, χ_u t_u , ps_u and rh.
f) Variance of mean rh, which is eventually used for GSI "qoption=2". Variances are stored in 5% bins of rh value.

Regarding latitudinal and vertical variation for these BE parameters, as such GSI code do not have any restriction. Thus in principle, the BE for all these parameters may vary both in latitude and in vertical. For GSI "qoption=2", due to practical reasons, mean "rh" information is generated in 5% bins for the whole domain. While running GSI, depending upon the magnitude of background relative humidity, corresponding information is pulled out from the desired bin.

3. Development work for WRF-ARW-GSI_BE code:

"WRF-ARW-GSI_BE" and parallel NCL based graphics has been developed to compute and display various components of GSI background error statistics for WRF-ARW model. An end-to-end shell script has also been developed and integrated with WRFDA. Like WRFDA "gen_be" utility, "WRF-ARW-GSI_BE" procedure is run in three stages (stage0, staghe2 and stage2) to produce the desired BE. Role for these three stages subsequently will be discussed below.

Like “gen_be” utility of WRFDA, “WRF-ARW-GSI_BE” procedure is also capable of generating BE statistics both for “NMC” and “Ensemble” methods. For generating corresponding BE with ensembles, stage0” needs to be activated with its namelist option “BE_METHOD” set as “ENS”. Generally, since very little moisture is available in the top model layers and sometimes WRF-ARW forecasts also produces negative moisture in these layers (with some specific WRF-options), provision has been made to avoid computing moisture background error statistics for some of the top layers. The same may be achieved by setting nammelist variable “less_q_from_top” with the desired number of levels to be suppressed from the top of the model.

3.1 Stage0 for WRF-ARW-GSI_BE:

In “stage0” all the desired information like geographical information about the domain, map projection & the associated map factors, all meteorological information about the desired variables like surface pressure, wind, temperature, moisture etc are processed from 12 & 24 hour WRF-ARW forecasts (wrfout files) and written (unformatted) in an intermediate files. Stream function (ψ) and velocity potential (χ) are computed from u & v components of the wind field in two steps as follows.

Step 1: Compute horizontal divergence (D) and vorticity (ζ)

$$D = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \qquad (3.1.1)$$

$$\varsigma = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \qquad (3.1.2)$$

Step 2: Following Poisson equations are solved to get ψ and χ

$$\nabla^2 \psi = \varsigma \qquad (3.1.3)$$

$$\nabla^2 \chi = D \qquad (3.1.4)$$

WRF-ARW-GSI_BE uses standard Fast Fourier Transform (FFT) decomposition and Poisson solver codes, currently available with “gen_be” utility of WRFDA. Two options are available for solving Poisson equation (3.1.3-4), namely the “spectral” and “relaxation”. Default option is set as “spectral”.

Finally for “NMC” method, desired “perturbations” are formed by taking difference of 12 & 24 hour forecasts, valid for the same time. Such “perturbations” are generated for all the desired fields. In case of “Ensemble” method, ensemble mean is subtracted from each of the ensemble members to form the “perturbations”. Mean relative humidity is also processed for generating BE statistics corresponding to GSI “qoption=2”.

3.2 Stage1 for WRF-ARW-GSI_BE:

In this stage, time mean of all the perturbations generated in "stage0" are removed from each perturbation.

3.3 Stage2 for WRF-ARW-GSI_BE:

This stage mainly consists of two parts. First part computes the regression coefficients of velocity potential (χ), temperature (t) and surface pressure (ps) with respect to stream function (ψ) field. These regression coefficients are used to build-up unbalanced velocity potential (χ_u), unbalanced temperature (T_u) and unbalanced surface pressure (ps_u) by subtracting corresponding balanced part from the respective full field. Subsequently in the second part, variance and length-scales (both horizontal and vertical) are computed for each of the desired analysis control variables.

3.3.1 GSI Regression coefficients:

Following NCEP regional BE code, "WRF-ARW-GSI_BE" is designed to compute latitude dependent (ψ, χ) and (ψ,ps) regression coefficients. For each level, (ψ,t) regression coefficients are computed in an averaged sense over the whole horizontal domain. Steps followed in computing these regression coefficients are as follows**.**

a) Level wise, for each variable (ψ, χ, T & rh), corresponding horizontal mean is removed
b) For each latitudinal bin, compute variance of ψ ($<\psi,\psi>$)
c) For each latitudinal bin, compute covariance of ψ with χ ($<\psi,\chi>$)
d) Use $<\psi,\psi>$ and $<\psi,\chi>$, computed in steps b) & c), to get regression coefficients of (ψ, χ) as follows

$$\text{Regression coefficients of } (\psi, \chi) = <\psi, \chi> / < \psi, \psi >$$

e) Compute horizontally averaged vertical error covariance matrix (nz x nz, where nz is number of vertical sigma levels) for ψ
f) Normalize each elements of vertical error covariance matrix, computed in step e), with standard deviation (square root of diagonal terms)
g) Following Daley (1999, p110), compute vertical length-scale (s) at each sigma level and normalize it with layer thicknesses in log(sigma) units.
h) Using normalized length-scale (s) for ψ, computed in step g), vertical error covariance matrix for ψ (ψ_{vcorr}) is re-formulated as follows

$$\psi_{vcorr}(l,m) = \exp\{-\frac{1}{2}(x^2 / s^2)\}$$

Where, x is the thickness (in log(sigma) units) between l & m sigma levels.

i) Project ψ on the basis functions of vertical error covariance ψ_{vcorr}, computed in step h), to get coefficients (C_ψ) corresponding to these basis functions as follows

$$C_{\psi}(k) = \sum_{l=1}^{nz} \psi(l)\psi_{vcorr}(l,k)$$

j) Compute variance of C_{ψ} ($<C_{\psi}, C_{\psi}>$)
k) Compute covariance of C_{ψ} and surface pressure (ps) ($<C_{\psi}$, ps$>$)
l) For each level l, compute covariance of C_{ψ} and temperature ($<C_{\psi}$, t(l) $>$)
m) Use Chelosky decomposition method to solve (for X) following equations to get (C_{ψ}, ps) regression coefficients

$$<C_{\psi}, C_{\psi}> X = <C_{\psi}, ps> \quad (3.3.1.1)$$

n) For each level l, use Chelosky decomposition method to solve (for X) following equation to get (C_{ψ}, t(l)) regression coefficients

$$<C_{\psi}, C_{\psi}> X = <C_{\psi}, t(l)> \quad (3.3.1.2)$$

o) Finally (ψ,ps) and (ψ,t) regression coefficients are obtained by summing up contributions corresponding to each vertical error covariance basis function regression coefficients computed in step m) and n), respectively.

Remarks:

i) Currently the code produces surface pressure regression coefficients averaged over whole domain. Thus the surface pressure regression coefficients output is not latitude dependent.
ii) In "gen_be" utility of WRFDA, computation of (ψ,χ) regression coefficients is exactly the same like in this code. However, temperature and surface pressure regression coefficients are determined directly with the ψ field. Unlike as in WRF-ARW-GSI_BE, ψ field is not decomposed using its vertical error covariance in "gen_be" code.
iii) "gen_be" also employs eigen decomposition method to solve regression coefficient equations like 3.3.1.2 & 3.
iv) In GSI, for levels with sigma < 0.8, ψ and χ fields are decoupled.

Provision has been made in "WRF-ARW-GSI_BE" (by introducing a namelist variable "fstat") to include proper weighting of coriolis parameter contributions in developing the regression coefficients between psi and temperature field. However, keeping in view that currently GSI default option does not support such statistics, default option for "fstat" is set to ".false."

3.3.2 Variance:

Procedure for computing variance for control variables dealing with full fields like ψ and rh is straight. For other control variables dealing with unbalanced variables like χ_u, t_u & ps_u, it proceeds as follows.

a) Use corresponding regression coefficients to compute unbalanced field.
b) For each case (date), accumulate variance (level-wise) in respective latitude bins for each control variable.
c) For each latitudinal bin, rescale accumulated variance computed in step b) with the number of cases (dates) and the number of respective bin points.

Remarks:

i) Square root of variance (standard deviation) is stored in BE file which is supplied as input to GSI
ii) For GSI "qoption=2", mean rh variance is also computed averaged over the whole domain. These variances are accumulated in equally spaced 5% bins of mean rh value.
iii) Moisture variance both for rh and mean rh is filled with very small value (=0.00000001) for all the desired levels as set by "less_q_from_top".
iv) In GSI, a fixed value of variance for ozone, sea surface temperature, land & ice is assigned. Currently, these values are 0.000008, 0.3 K and 1.0 K respectively.

3.3.3 Horizontal length-scale:

Following, Wan Shu et al (*MWR,* 2002) and Pannekoucke et al (*QJMRS,* 2008), for any variable X, length-scale (L) is computed using its variance and the variance of its Laplacian (∇^2X) as follows.

$$L = \left\{ \frac{8 * Variance(X)}{Variance\{\nabla^2 (X)\}} \right\}^{1/4} \tag{3.3.3.1}$$

For each control variable, procedure for computing length-scales proceeds as follows.

a) For each case (date), accumulate length-scale (level-wise) using (3.3.3.1) in respective latitude bins for each control variable.
b) For each control variable for each latitude bin, take average of length-scales (level-wise) computed in step a) for all the cases (dates)

Remarks:

i) In principle, this method computes horizontal length-scale at each grid point. However, these length-scales are accumulated in latitudinal bins and finally it is averaged in each bin.
ii) Selection of latitudinal bin-width (in degrees) is flexible and it may be dictated through a namelist parameter "lat_bins_in_deg". Thus if desired, with proper choice of "lat_bins_in_deg", the same code may be used to compute domain averaged BE.

iii) For moisture (rh) horizontal length-scales are computed only for the desired levels as dictated by “less_q_from_top”. For rest of the undesired levels it is filled with very small value (=0.00000001)

iv) Currently, WRFDA “gen_be” utility estimates horizontal scale-length by fitting following Gaussian curve with correlations (σ) accumulated in bins, which are arranged in increasing order of distance (r).

$$\rho(r) = \exp\left\{\frac{-r^2}{8L^2}\right\}$$

v) Currently in GSI code, following adjustments for horizontal length-scales are made for different control variables.

a) At higher levels (sigma < 0.15), for each control variable it uses a corresponding fixed value for its horizontal length-scale. This fixed value is the corresponding horizontal length-scale at the middle of the domain at model level which is just below the threshold sigma value = 0.15.
b) For sea surface temperature, corresponding values for ψ fields are used.
c) For ice and land temperature, it uses one-fourth of the corresponding first sigma-level ψ field.
d) For cloud hydrometeor, corresponding rh values are used.
e) For ozone a fixed value of 400 Km is used in the lower sigma levels (≤3/4 * nz). For higher sigma levels (> 3/4 * nz), it gradually increase linearly to 800 Km.

3.3.4 Vertical length-scale:

Following procedure is followed for computing vertical length-scales for ψ, χ_u, t_u and rh variables.

a) Compute variance at each level.
b) Compute vertical covariance between each level with adjacent levels just below
c) Normalize covariance computed in step b) with the respective square root of variance computed in step a)
d) Using normalized level-wise vertical covariance as computed in step c) and following Daly (1999, p110) compute vertical length-scale (VL) at any level l, as follows.

$$VL(l) = \left\{\frac{1}{abs[2 - vcor(l) - vcor(l+1)]}\right\}^{1/2} \qquad (3.3.4.1)$$

e) Convert vertical length-scale computed in step d) in log(sigma) units by rescaling it with the corresponding layer thickness (Δlog(sigma)).

Remarks:

i) Inverse of vertical length-scales computed in step e) is stored in BE file as input to GSI.

ii) In GSI code, a fixed value (=0.53333333) is used for ozone control variable.
iii) For cloud hydrometeor, corresponding values for moisture (rh) is used.

Following table (Table-1) gives a summary of the list of some useful namelist options, which are currently available with WRF-ARW-GSI_BE. Their default values are also listed.

Table-1: List of namelist options currently available with WRF-ARW-GSI_BE code

Variable name	Type	Default option	Description
BE_METHOD	character	NMC	Method of computing BE statistics "NMC" or "ENS", the ensemble based
POISSON_METHOD	Integer	1	Method for Poisson solver 1 – Spectral 2 – Relaxation
FFT_METHOD	Integer	2	Fast Fourier Transform 1 – Cosine 2 - Sine
FSTAT	Logical	False	It includes the contribution of coriolis parameter effect while computing temperature and stream function regression coefficients
LAT_BINS_IN_DEG	Real	1.0	Width of latitude bins in degrees
LESS_Q_FROM_TOP	Integer	0	Number of top model levels to eliminate moisture BE statistics
Debug	Integer	0	Flag for debugging the code Set > 0 for debugging

4. Utilization of BE statistics in GSI:

Since BE statistics is computed in latitudinal bins, it needs to be interpolated to the desired domain configuration both in horizontal and vertical. It is also well known (Bannister, 2008) that the BE statistics produced using either "NMC" or the "ensemble" methods has some shortfalls for its real time application. It needs to be tuned. Keeping in view that at higher model levels, there is very little balance between mass and wind field, accordingly the analysis procedure should relax the balance constraint. Rest of this section discusses how these aspects are currently tackled in GSI code.

4.1 BE Interpolation:

At each grid location, GSI code loads desired information about BE, which is originally available in latitudinal bins. First, the BE statistics gets interpolated from "BE sigma levels" (read from BE input file) to the desired analysis sigma levels corresponding to the analysis background field (first guess). For this purpose, first an average guess pressure profile (*ges_prslavg*) is computed with known WRF-ARW Eeta (η) values and guess surface pressure (*ges_psfcavg*) as 1013 hPa, as follows.

$$ges\_prslavg(k) = ptop + \eta(k) * (ges\_psfcavg - ptop) \quad , \quad 1 \le k \le nz$$

Where, *ptop* is the model top pressure

Using above guess pressure profile (*ges_prslavg*) model sigma levels (*sig*) are fixed as,

$$sig(k) = ges\_prslavg(k) / ges\_pafcavg \quad , \quad 1 \le k \le nz$$

Finally, vertical interpolation from "BE sigma levels" to the sigma levels (computed above) is done in log(sigma) co-ordinate. Interpolated fields are, all (three) regression coefficients, horizontal & vertical length-scales and variance.

Remarks:

It is not clear why BE is interpolated to the sigma levels as computed above based on fixed surface pressure (= 1013 hPa) and an average guess pressure profile. Instead, actual WRF-ARW sigma levels (available in WRF-ARW forecast as first guess) may be used.

At each grid point, horizontal interpolation is done linearly between BE bin latitudes. For this purpose in GSI code, a two-dimensional array (rlat) is build-up to store the location of analysis grid in BE statistics grid units. For all analysis grid points south of southern-most BE latitude, corresponding "rlat" is set as 1. Similarly for all analysis grid points north of northern-most BE latitude, "rlat" is assigned as the maximum of bin latitude index (mlat). For all analysis grid locations lying within BE latitudes, respective grid index is interpolated linearly with BE latitudes.

4.2 BE Tuning:

Through GSI namelist option, it is possible to adjust variance, horizontal & vertical length-scales. GSI has also a provision of analyzing three scales (Wan Shu et al (*MWR,* 2002). Corresponding length-scales and respective weights to be assigned to these three scales may also be tuned. Following table (Table 2) gives a summary of tuning parameters used in GSI.

Table 2: List of GSI namelist options, currently available for tuning BE statistics

Variable name	Used for tuning
AS (10 dimensioned array) Only first 8 are active	Variance for psi, chi_u, ps_u, t_u, moisture, ozone, (Sea, land & ice temperature), cloud condensate mixing ratio
VS	Vertical length-scales. Same is used for all the control variables
HZHCL	Horizontal length-scale corresponding to three scales.

(3-dimensioned array)	Same variable is used for all the control variables.
HSWGT (3-dimensioned array)	Weights to be assigned to three scales

Remarks: Currently in GSI code, there is no way that horizontal length-scales for different control variables may be tuned differently. The same is true for vertical length-scales also.

5. Experiment with CONUS (200 Km) domain:

WRF-ARW-GSI_BE has been tried to generate WRF-ARW BE for GSI, using "NMC method" option for CONUS 200 Km domain. For this run, July 2007 12 & 24 hour WRF-ARW forecasts,, initialized both at 00 and 12 UTC are used. Thus in all 60 pairs of perturbations are utilized to generate WRF-ARW BE. Figure 5.1 displays (top-left panel) vertical profile for the contribution of balance part of χ and temperature field averaged over the whole domain. Contribution of surface pressure is also shown (bottom-left) panel. Parallel result from "gen_be" utility of WRFDA is also shown (right panel). No serious discrepancies are seen between the two outputs from two completely independent codes, namely the WRF-ARW-GSI_BE and "gen_be". This gives confidence that the regression coefficients computed by "WRF-ARW-GSI_BE" code is consistent with WRFDA "gen_be" utility. However, slight difference is noticed, especially in the contribution of balanced part of temperature field. These differences are mainly due the computation of basic inputs for computing regression coefficients, namely the variance & covariance matrixes. Procedure for binning, averaging and solving linear regression coefficients etc. are slightly different in the two codes. Moreover, method for computing regression coefficients for temperature and surface pressure is totally different in WRF-ARW-GSI_BE and "gen_be" utility of WRFDA. As described in (3.3.1), in case of "WRF-ARW-GSI_BE", temperature regression is first computed with ψ-EOF's coefficients. Next, the contributions of each of these regression coefficients are summed up in ψ-EOF-space to get the final regression coefficient. However in case of "gen_be" level-wise, corresponding regression coefficients are computed directly with the ψ-field.

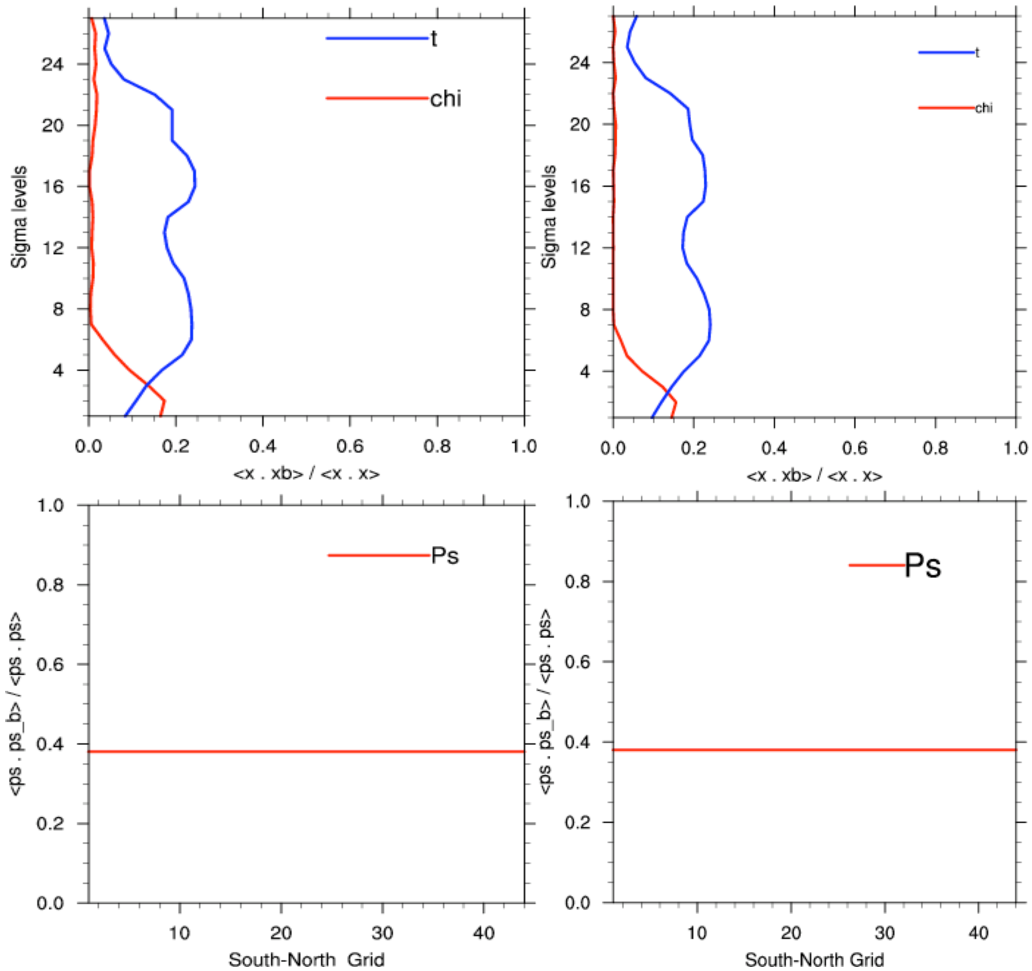

Figure 5.1: Domain averaged vertical profile of the contribution of balanced part of velocity potential (χ) and temperature (top-left panel). Contribution of balanced part of surface pressure is shown in bottom-left panel. Right panel displays the corresponding results from "gen_be" utility of WRFDA.

Figures 5.2-4 display average profile of variance, horizontal & vertical length-scales for ψ, χ_u, t_u and rh, respectively. In Figure 5.2 (bottom), variance for unbalanced surface pressure (ps_u) is also displayed. In these Figures, left panel is for the output from WRF-ARW BE from WRF-ARW-GSI_BE and the right panel is the parallel output from NCEP global BE statistics ("nam_nmmstat_na_glberror"). In Figure 5.3 (bottom), horizontal length-scale for unbalanced surface pressure (ps_u) is also displayed.

In all these figures, because of differences in horizontal grid resolution, number of vertical sigma levels & its placement in vertical, the model top etc. for CONUS and global GFS domain configuration, it is hard to compare the parallel profiles level-by-level for the two BE output. Nevertheless, variable-wise its characteristic looks similar. Important observations are as follows.

a) Stream function and unbalanced velocity potential variance is one order smaller for WRF-ARW BE as compared to NCEP global BE. At jet level ($\cong$ 200 hPa), a distinct peak is observed both in ψ and unbalanced χ field in WRF-ARW BE. Moisture variance is almost half in WRF-ARW BE, compared to global BE.

b) In troposphere, not much variation is observed in the horizontal length-scales for almost all the fields in global BE. Whereas, in WRF-ARW BE except for ψ, a rising trend in the horizontal length-scales for all other control variables is seen. Horizontal length-scales magnitude for different control variable is almost same except for unbalanced velocity potential (It is almost two and half times bigger in NCEP global BE).

c) In general, vertical length-scale decreases with height, both in global and WRF-ARW BE statistics. Regarding magnitude, no definite conclusions can be drawn because of varying number and mismatch of the placement of sigma levels in the vertical.

d) Characteristic of surface pressure field is similar both in global and WRF-ARW BE statistics.

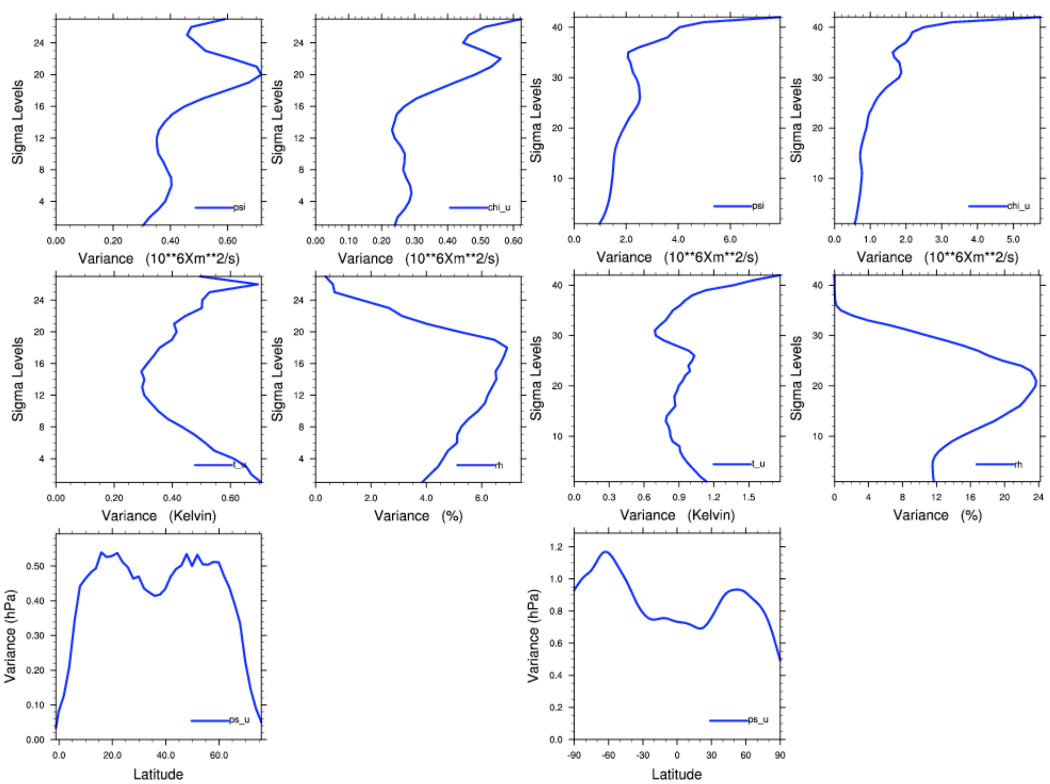

Figure 5.2: Average variance profile for ψ, χ_u, t_u and rh. Left and right panels are corresponding to WRF-ARW and NCEP global BE ("nam_nmmstat_na_glberror"), respectively. In both the panels, variance of unbalanced surface pressure (ps_u) is also displayed at the bottom.

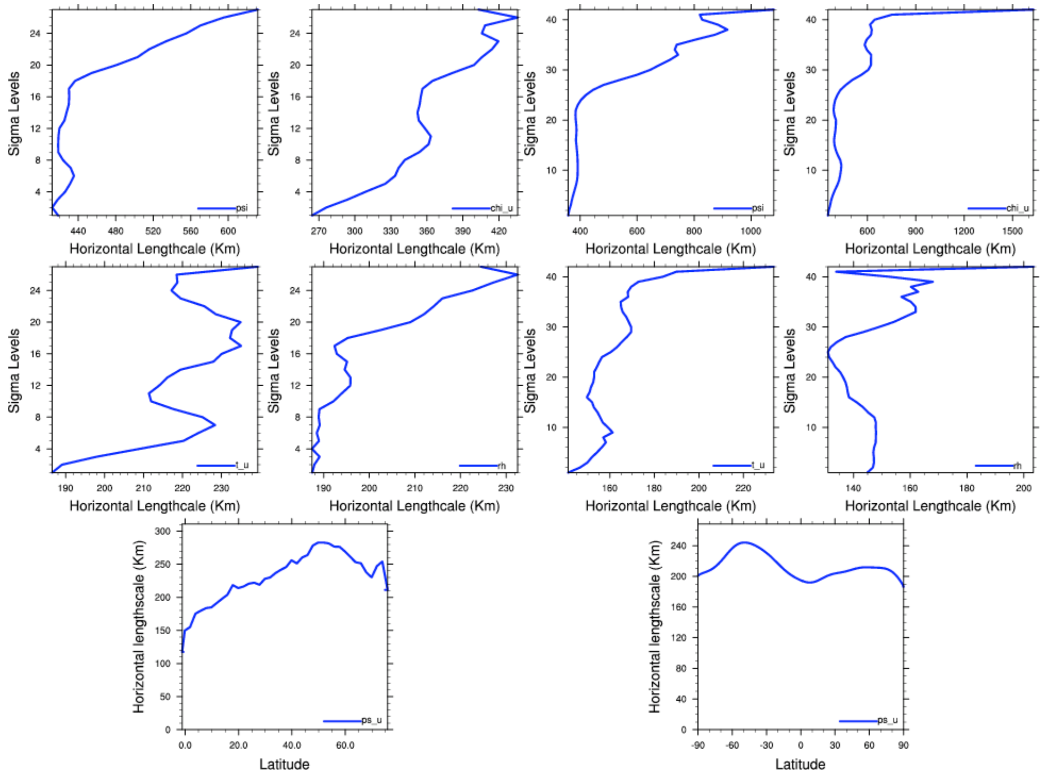

Figure 5.3: Average profile of horizontal length-scales for ψ, χ_u, t_u and rh. Left and right panels are corresponding to WRF-ARW and NCEP global BE ("nam_nmmstat_na_glberror"), respectively. In both the panels, horizontal length-scale for unbalanced surface pressure (ps_u) is also displayed at the bottom.

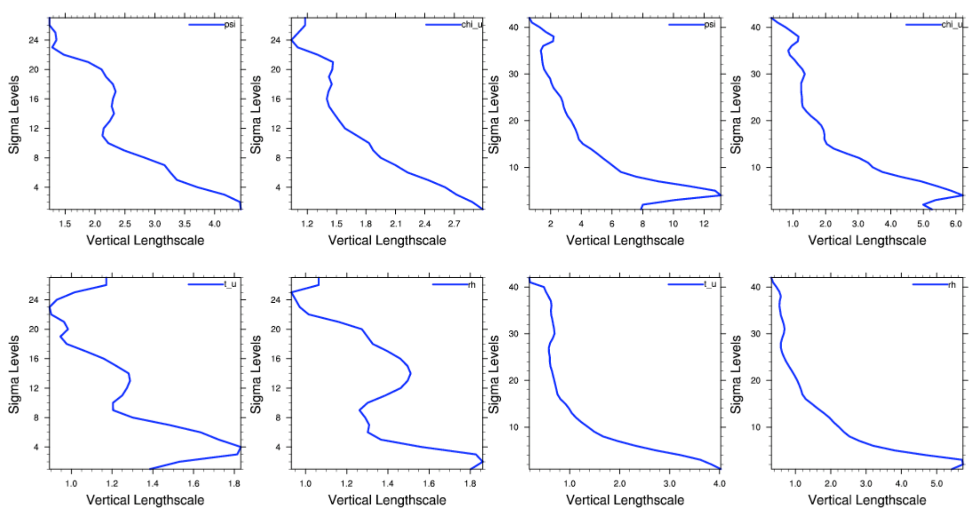

Figure 5.4: Average profile of vertical length-scales for ψ, χ_u, t_u and rh. Left and right panels are corresponding to WRF-ARW and NCEP global BE ("nam_nmmstat_na_glberror"), respectively.

Single Observation results with CONUS (200 Km) Domain:

For understanding structure of BE statistics, single observation tests are made with GSI for wind (U-component), temperature and moisture with WRF-ARW-GSI_BE generated BE for CONUS 200 Km domain. Parallel runs are also made with NCEP global BE ("nam_nmmstat_na_glberror"). For each of these runs, horizontal location of observation is at the middle of the domain and vertical position is at 12th sigma level (close to 500 hPa). In each of these runs, innovations input is 1 unit and observation errors are the default values as specified with "oneobtest" utility of GSI. With NCEP global BE, GSI default BE "tuning" options are used. Whereas with WRF-ARW BE, two parallel runs are made one without any BE "tuning" (all tuning parameters are set to 1.0) and the other with horizontal length-scale tuning factor =1.25. All single observation parallel runs are made with "qoption=1".

Here it may be noted that due to resolution mismatch (both in horizontal and vertical), BE tuning factors etc., when GSI is run, accordingly the desired BE info is extracted from the input BE and gets adjusted. Thus before discussing single observations test results, let us see the impact of tuning and vertical interpolation on BE. For vertical-length scales, the same is displayed in Figure 5.5. Side by side, it displays the interpolated BE corresponding to WRF-ARW (left panel) and NCEP global BE (right panel). By comparing Figure 5.4 & 5.5, one can see that the contents of NCEP global BE has been adjusted a lot. It is seen that the same is true for all other control variables also (not shown). Bulk of these changes are contributed by the default values of BE "tuning" factors. For example, the default value for vertical length-scale tuning parameter is 0.7, accordingly the final interpolated vertical length-scales for all the control variables at each level will be reduced 30%. Whereas, in case of WRF-ARW BE, because it is custom made for this domain, accordingly the GSI run was made with "interpolate_reg_stats=.false.". Moreover, vertical length-scale tuning factor is also set as 1.0. Due to these reasons, WRF-ARW BE did not undergo any change.

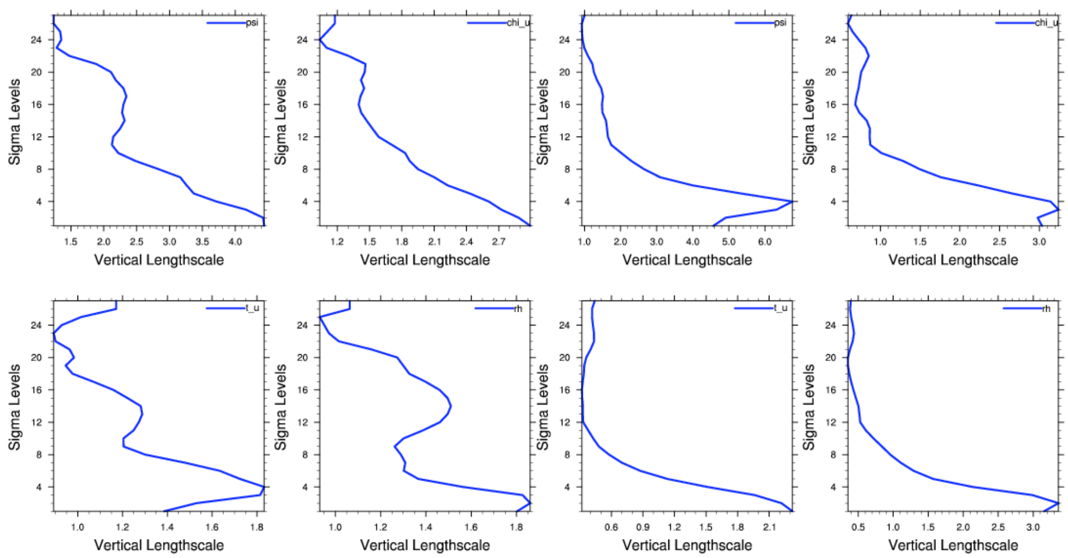

Figure 5.5: GSI interpolated average profile of vertical length-scales for ψ, χ_u, t_u and rh. Left and right panels are corresponding to WRF-ARW and NCEP global ("nam_nmmstat_na_glberror") BE, respectively.

In Figure 5.6-8, results corresponding to single observation tests with U-component of wind, temperature and moisture observations are shown respectively. It displays horizontal (XY at12th sigma level), West-East (XZ) and South-North (YZ) cross-sections for U, T and Q analysis increments respectively. In these figures, left and middle panels correspond to GSI runs with WRF-ARW and NCEP global BE respectively. The right-most panel corresponds to GSI run with WRF-ARW BE with horizontal length-scale tuning factor =1.25. It can be seen that the quality of analysis increments are very similar with two BE's. With U-single observation, U-increments are elongated more in West-East direction with global BE, whereas with WRF-ARW BE these increments are mainly circular in shape. In general with WRF-ARW BE, analysis increments penetrate more vertically. Horizontal spread for wind increments is more with global BE both in West-East and South-North directions. With temperature single observation, horizontal spread of temperature analysis increments is similar with two BE's. Compared to WRF-ARW BE since variance of temperature is more in global BE, accordingly the magnitude of temperature increments are more with global BE. Vertical spread of temperature with global BE is very-very small as compared to WRF-ARW BE. Results with moisture as single observation suggests that like temperature, the vertical scales with global BE is very-very small as compared to WRF-ARW BE.

Multivariate response with two BE's is shown in Figure 5.9. It displays (XY cross-section) the response on U, V, T & Q with 1.0 Kelvin temperature innovations. Left & right panels correspond to WRF-ARW and NCEP global BE respectively. Clearly, it is seen that the structure of analysis increments for all the variables is similar with two BE's. However, overall the magnitude is higher with global BE as compared to WRF-ARW BE.

Here it may be worth mentioning that with proper selection GSI BE tuning parameters it is always possible to achieve the desired spread in analysis increments both in horizontal & vertical. Increase in horizontal spread of analysis increments with WRF-ARW BE is shown (right-most panel) with horizontal length-scales tuning factor =1.25. However, since these scale-lengths tuning parameters (both horizontal & vertical) are common for all the control variables, sometimes it may not be possible to adjust the desired spread for wind consistent with temperature and vise-versa.

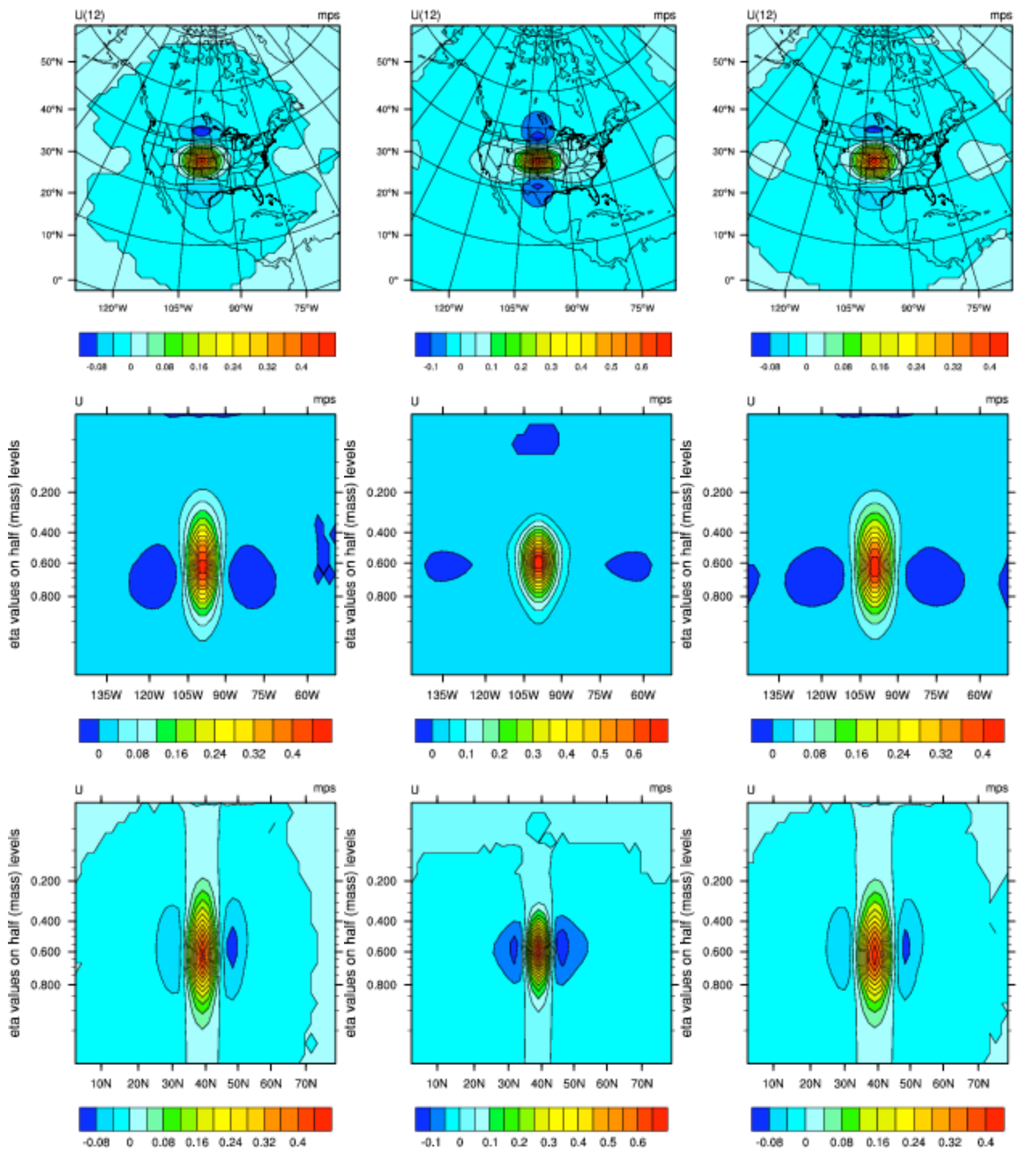

Figure 5.6: Horizontal (XY at 12th sigma level), West-East (XZ) and South-North (YZ) cross-sections of analysis increments for U with 1.0 mps wind (U-comp.) innovations at 12th sigma level in the middle of the domain. Middle panels correspond to GSI runs with NCEP global BE. Left & right panels correspond to WRF-ARW BE without tuning and with horizontal length-scale tuning = 1.25.

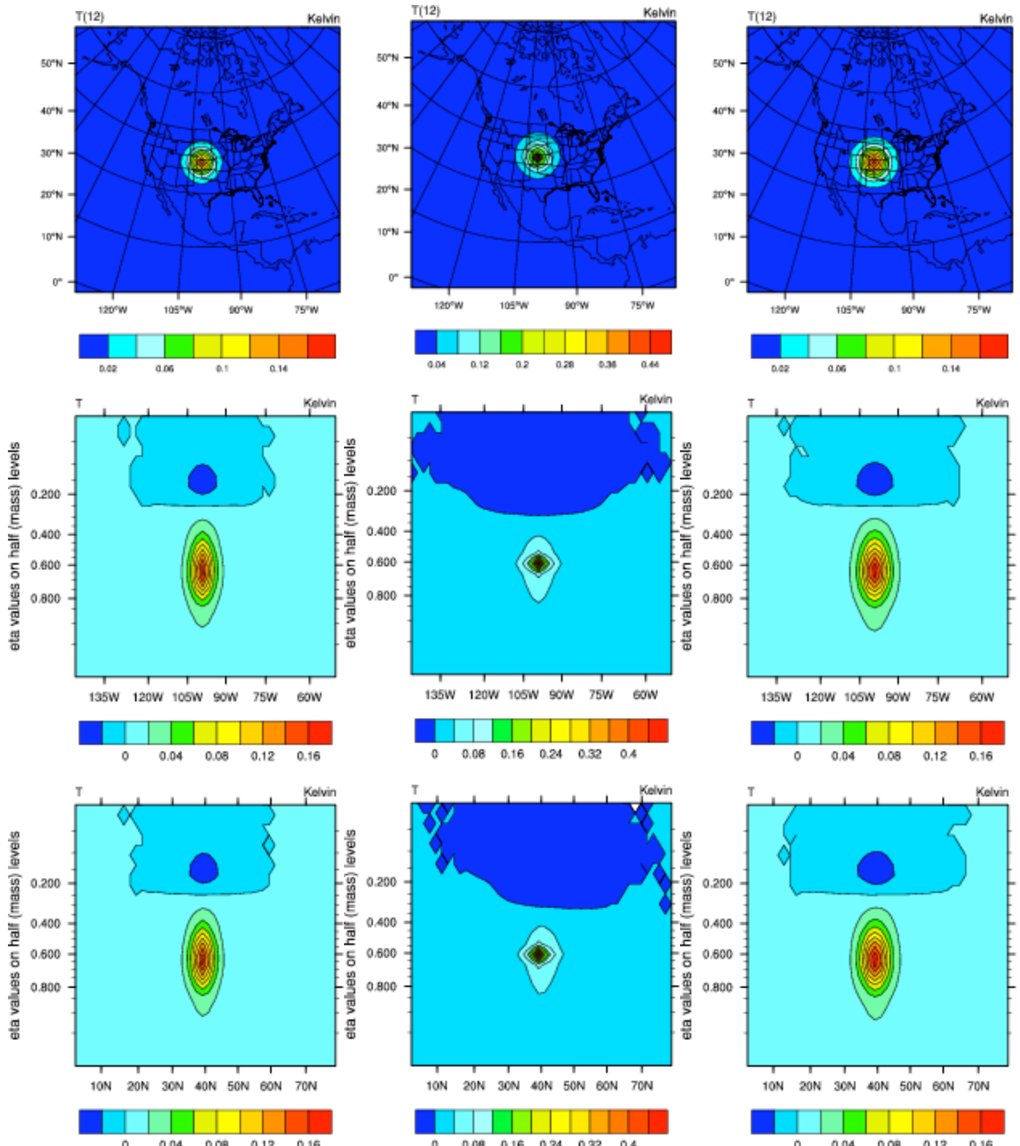

Figure 5.7: Horizontal (XY at 12th sigma level), West-East (XZ) and South-North (YZ) cross-sections of analysis increments for temperature with 1.0 Kelvin temperature innovations at 12th sigma level in the middle of the domain. Middle panels correspond to GSI runs with NCEP global BE. Left & right panels correspond to WRF-ARW BE without tuning and with horizontal length-scale tuning = 1.25.

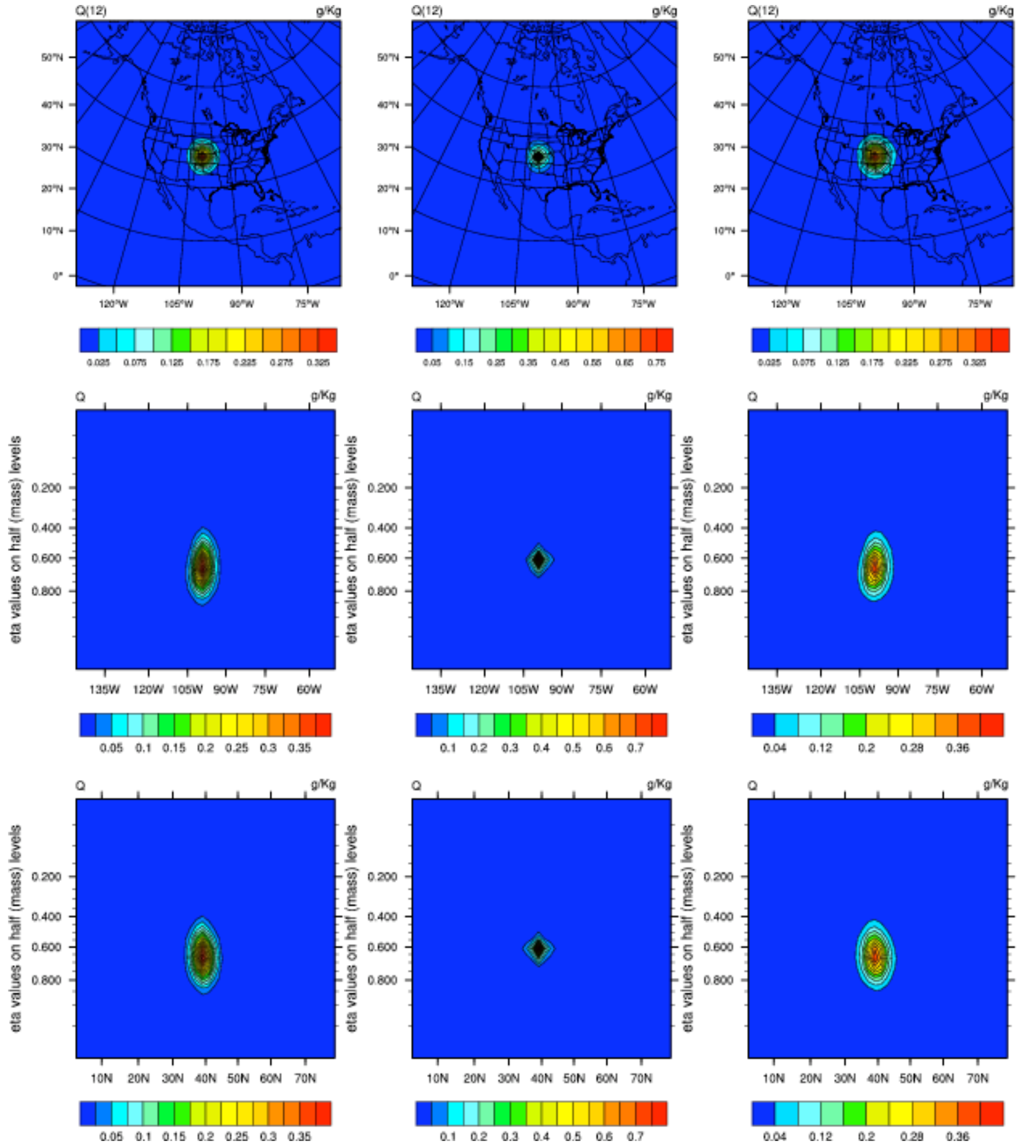

Figure 5.8: Horizontal (XY at 12th sigma level), West-East (XZ) and South-North (YZ) cross-section of analysis increments for moisture with 1.0 g/Kg moisture innovations at 12th sigma level in the middle of the domain. Middle panels correspond to GSI runs with NCEP global BE. Left & right panels correspond to WRF-ARW BE without tuning and with horizontal length-scale tuning = 1.25.

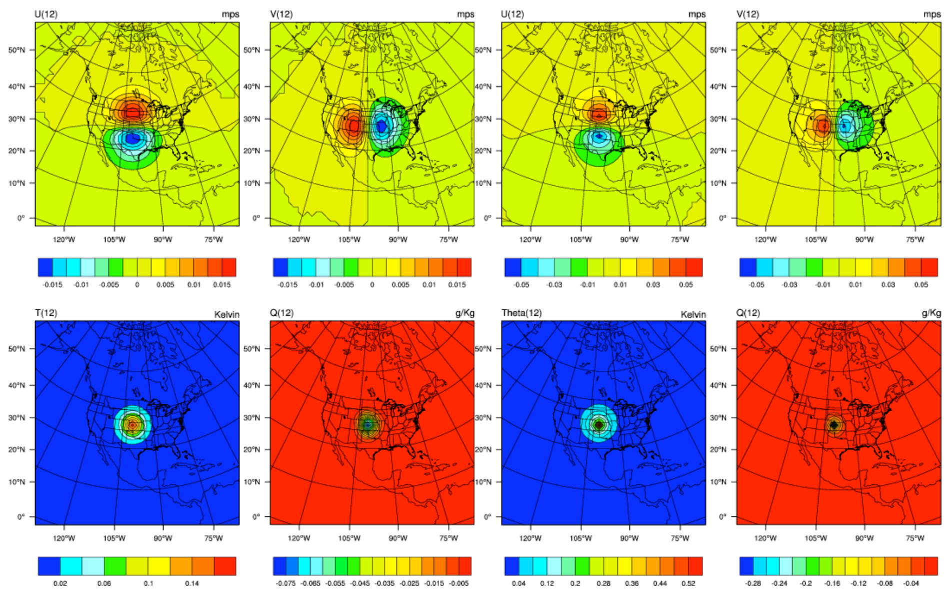

Figure 5.9: Horizontal (XY at 12th sigma level) cross-section of U, V , T and Q with 1.0 Kelvin temperature innovations at 12th sigma level in the middle of the domain. Left & right panels correspond to WRF-ARW BE with horizontal length-scale tuning = 1.25 and with global NCEP BE respectively.

6. Experiment with T44 (45 Km) domain:

For understanding issues associated with resolution, parallel runs are made with T44 (45 Km) domain. Since this domain consisted of 161 and 151 grid points in West-East and South-North direction respectively and 57 vertical sigma levels, it was ideal for testing this code for MPI environment. Accordingly, BE statistics was generated with12 & 24 hour WRF-ARW forecasts, initialized both at 00 and 12 UTC for the period 15th August to 15th September, 2007. Thus in this case also 60 pairs of perturbations are used as input to WRF-ARW BE for generating WRF-ARW BE. Figure 6.1 displays (top panel) average vertical profile of the contribution of balance part of χ and temperature field averaged over the whole domain. Contribution of surface pressure is shown in bottom panel.

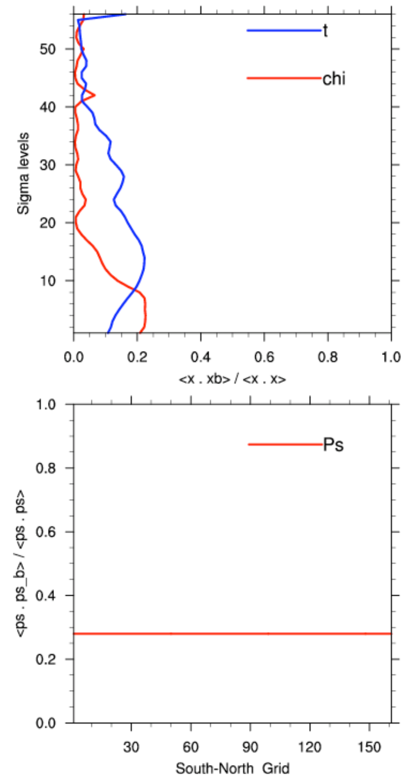

Figure 6.1: Domain averaged vertical profile of the contribution of balanced part of velocity potential (χ) and temperature (top panel). Contribution of balanced part of surface pressure is shown in bottom.

Figures 6.2-4, display average profile of variance, horizontal & vertical length-scales for ψ χ, t_u,and rh respectively. In Figure 6.2 variance for unbalanced surface pressure is also displayed (bottom panel). Left panel is corresponding output for WRF-ARW BE from WRF-ARW-GSI_BE and the right panel is the parallel output from NCEP global BE statistics (nam_nmmstat_na_glberror). In Figure 6.3 horizontal length-scale for unbalanced surface pressure is also displayed (bottom). Here it may be noted that in Figures 6.2-4, right panels are exactly the same as shown on the right panel in Figure 5.2-4. For easy comparisons, these are displayed side-by-side.

In general variance for all control variables are small with WRF-ARW BE. In particular, it is one order small for ψ and χ_u fields. Horizontal length-scales for all the control variables are small (almost half) with WRF-ARW BE. For χ_u it is almost four times smaller. In lower troposphere, horizontal length scales increases with height in WRF-ARW BE, whereas in NCEP global BE it does not vary much. Decreasing trend with height for vertical length-scales is seen in both the BE's.

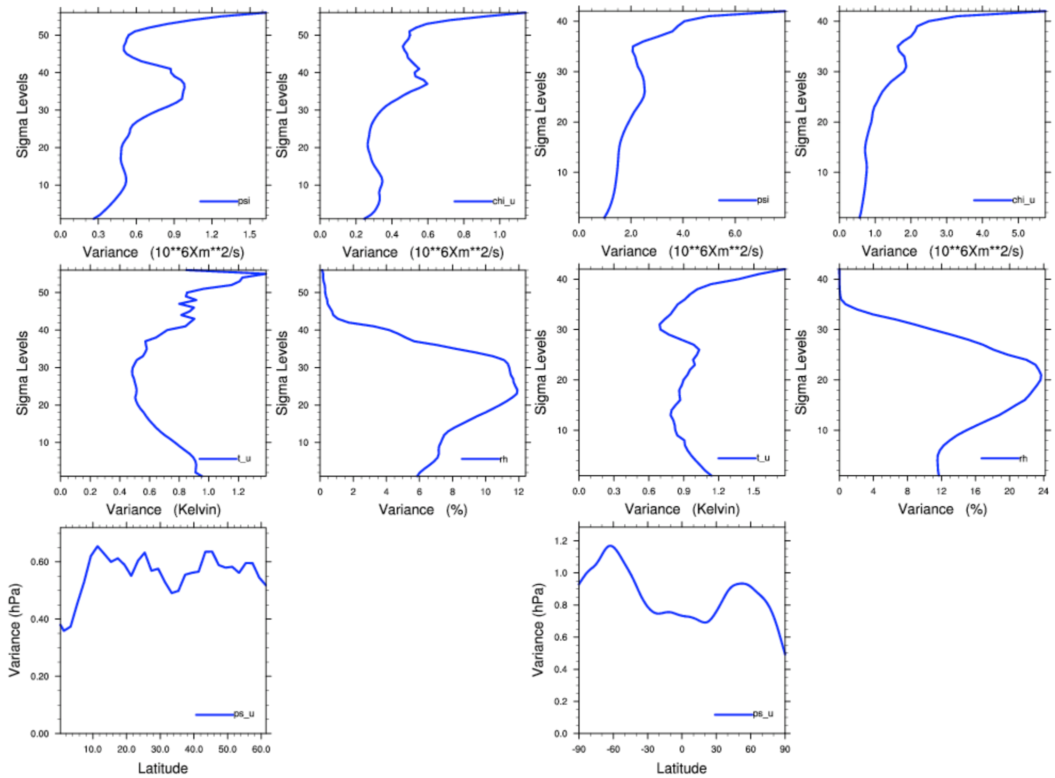

Figure 6.2: Average variance profile for ψ, χ_u, t_u and rh. Left and right panels are corresponding to WRF-ARW and NCEP global ("nam_nmmstat_na_glberror") BE, respectively. In both the panels, variance of unbalanced surface pressure (ps_u) is also displayed at the bottom.

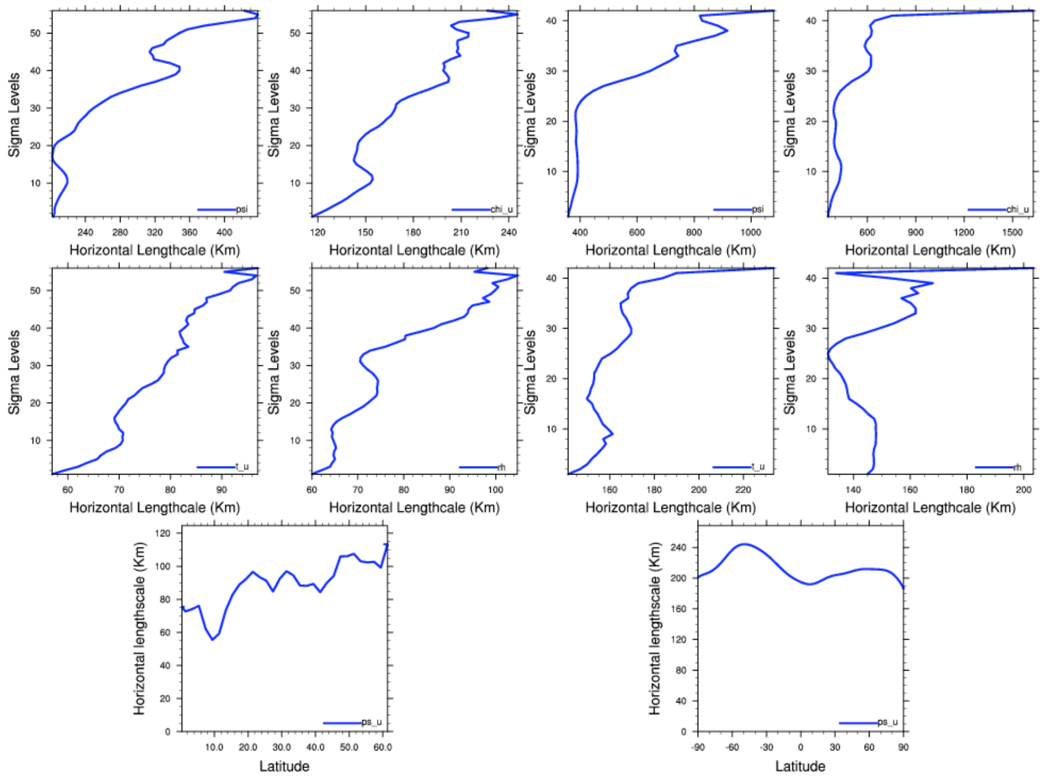

Figure 6.3: Average profile of horizontal length-scales for ψ, χ_u, t_u and rh. Left and right panels are corresponding to WRF-ARW and NCEP global ("nam_nmmstat_na_glberror") BE, respectively. In both the panels, horizontal length-scale for unbalanced surface pressure (ps_u) is also displayed at the bottom.

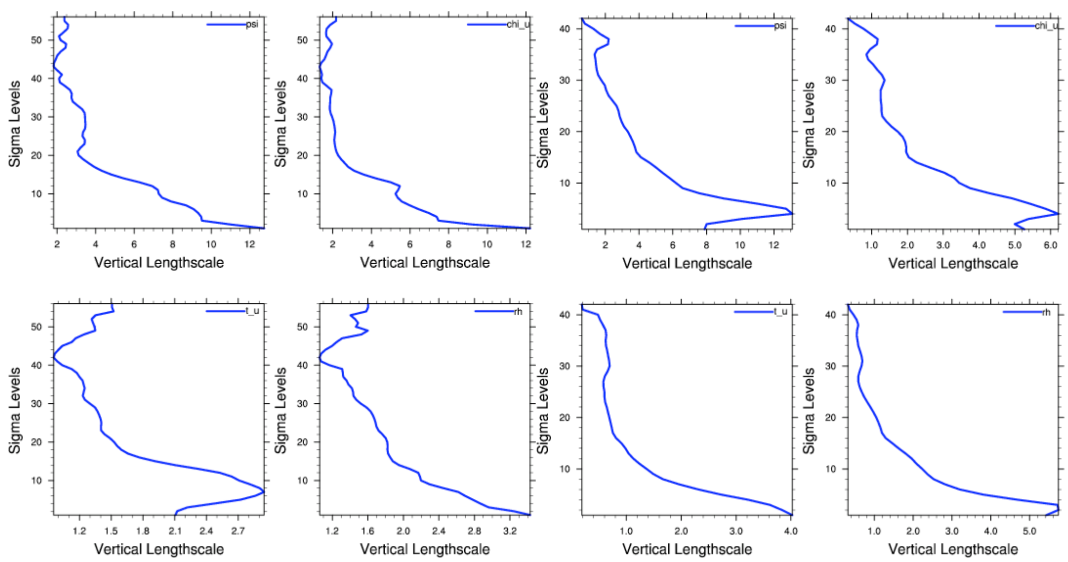

Figure 6.4: Average profile of vertical length-scales for ψ, χ_u, t_u and rh. Left and right panels are corresponding to WRF-ARW and NCEP global ("nam_nmmstat_na_glberror") BE, respectively.

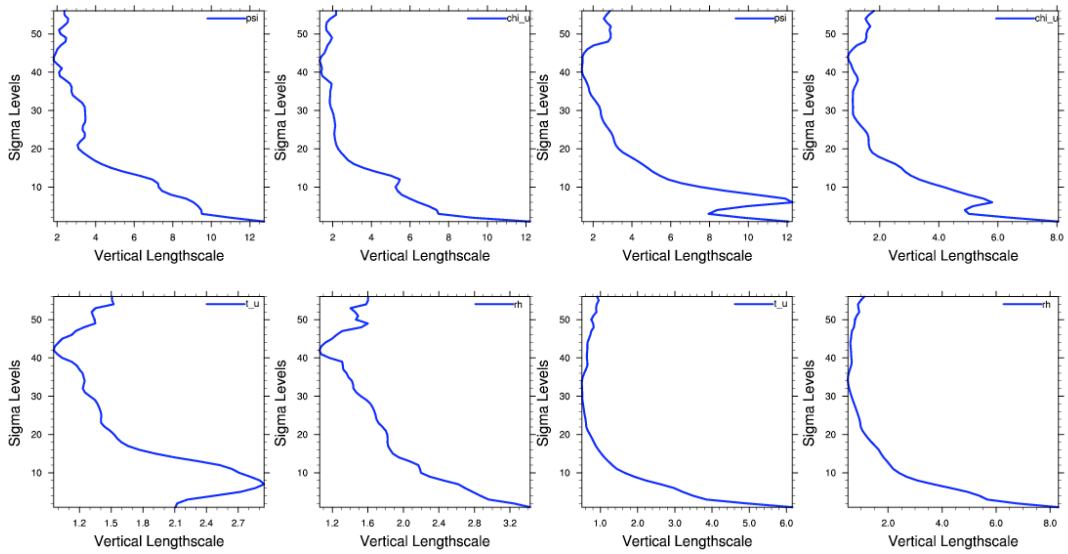

Figure 6.5: GSI interpolated average profile of vertical length-scales for ψ, χ_u, t_u and rh. Left and right panels are corresponding to WRF-ARW and NCEP global (“nam_nmmstat_na_glberror”) BE, respectively.

Single Observation results with T44 (45 Km) Domain:

Single observation results, like shown in Figure 5.6a-c for CONUS 200 Km domain, for T44 domain it is shown in Figure 6.6-8. Here in case of WRF-ARW BE, horizontal length-scale tuning factor is set 2.0. It may be seen (right panel) that by inflating horizontal length-scales by a factor of 2.0, spread of analysis increments resembles better with the corresponding with global NCEP BE (middle panel).

Multivariate response with two BE’s is shown in Figure 6.9. In this case GSI run with WRF-ARW BE was made without any tuning it. The quality of analysis increments is same. Horizontal spread with WRF-ARW BE is small and the same may be adjusted with proper selection of horizontal length-scale tuning parameter.

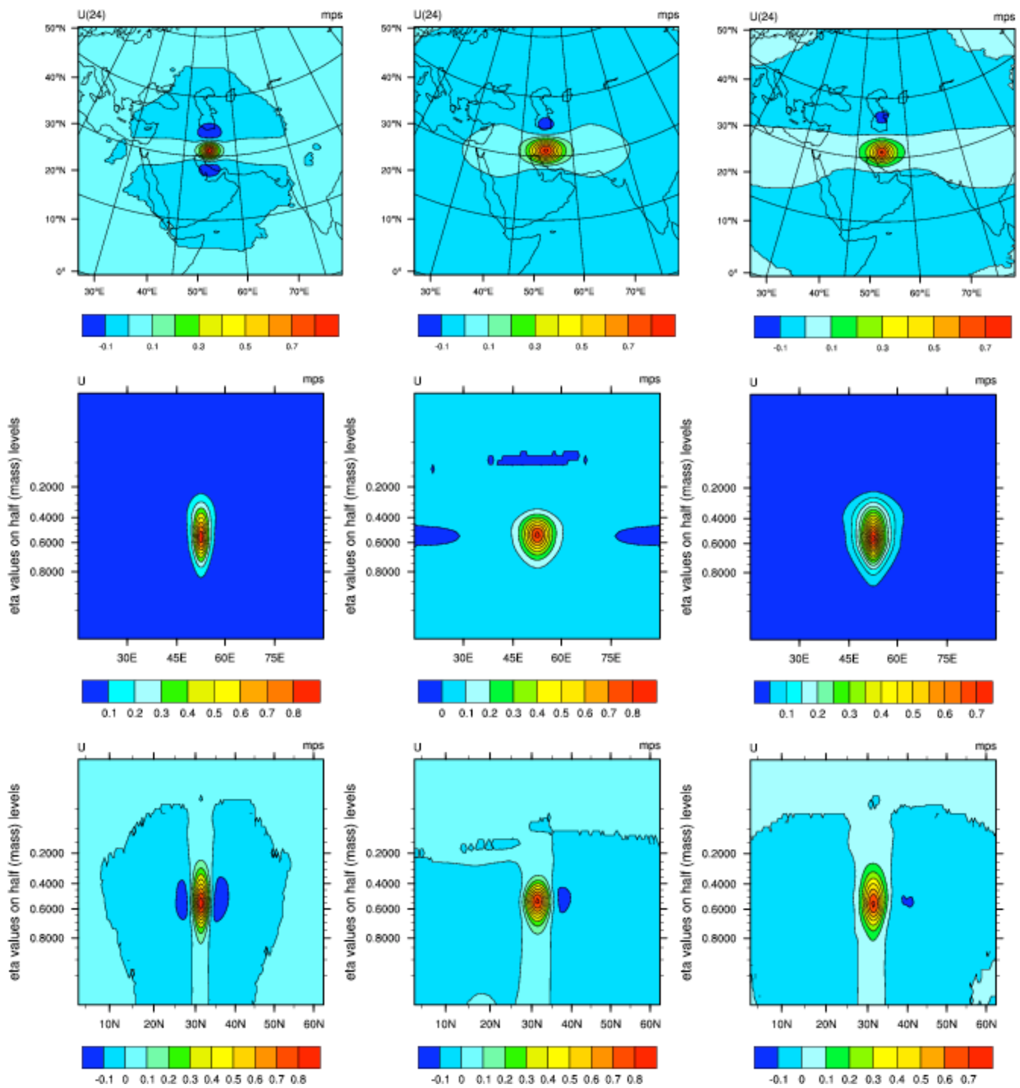

Figure 6.6: Horizontal (XY at 12th sigma level), West-East (XZ) and South-North (YZ) cross-sections of analysis increments for U with 1.0 mps wind (U-comp.) innovations at 12th sigma level in the middle of the domain. Middle panels correspond to GSI runs with NCEP global BE. Left & right panels correspond to WRF-ARW BE without tuning and with horizontal length-scale tuning = 2.0.

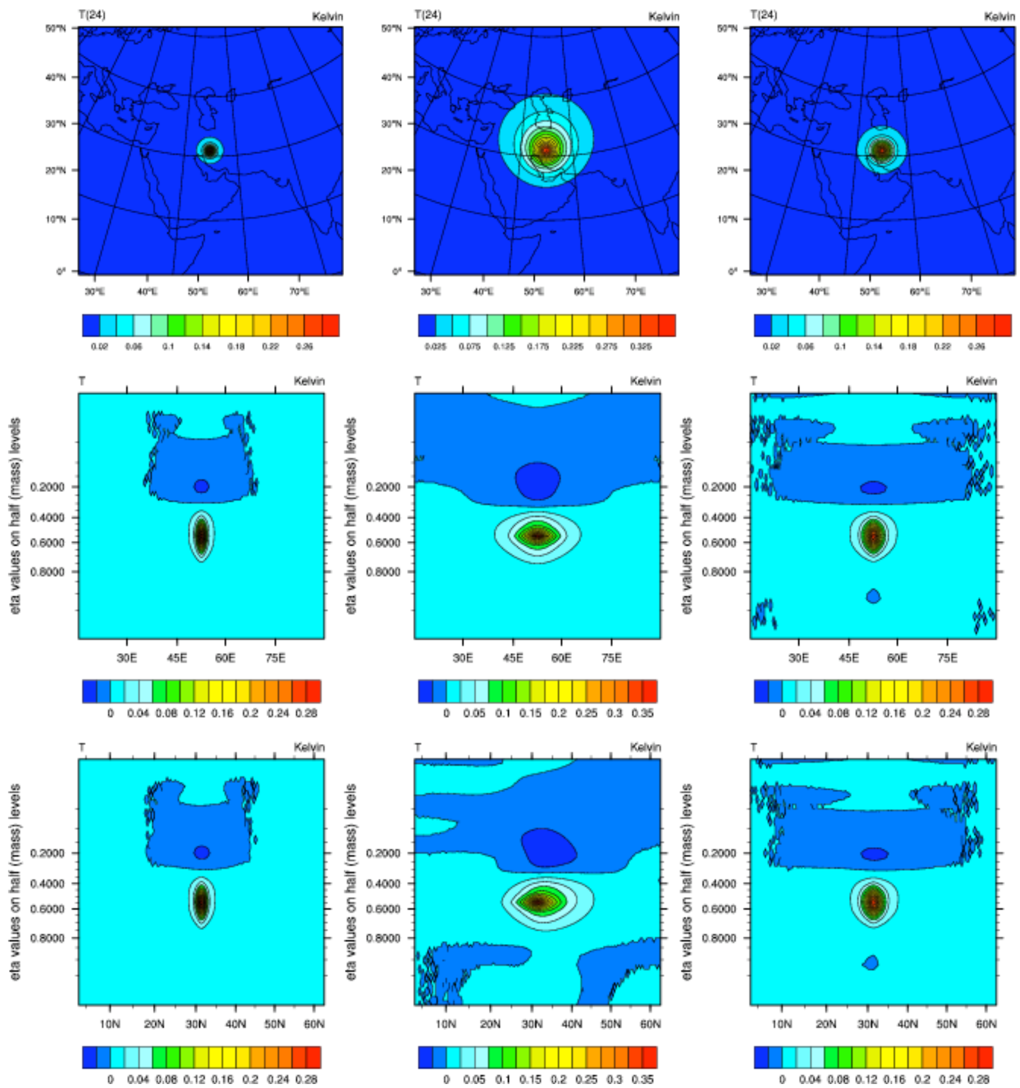

Figure 6.7: Horizontal (XY at 12th sigma level), West-East (XZ) and South-North (YZ) cross-sections of analysis increments for temperature with 1.0 Kelvin temperature innovations at 12th sigma level in the middle of the domain. Middle panels correspond to GSI runs with NCEP global BE. Left & right panels correspond to WRF-ARW BE without tuning and with horizontal length-scale tuning = 2.0.

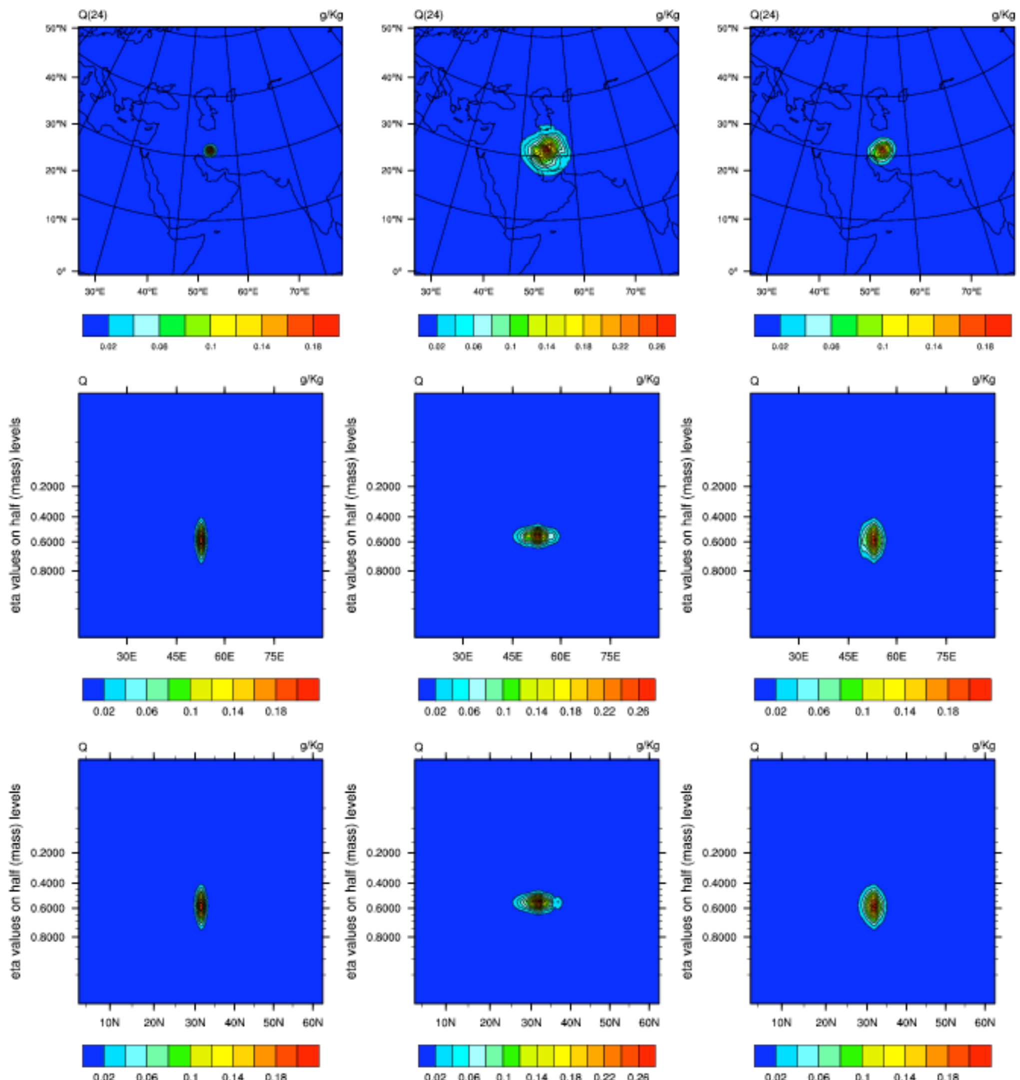

Figure 6.8: Horizontal (XY at 12th sigma level), West-East (XZ) and South-North (YZ) cross-section of analysis increments for moisture with 1.0 g/Kg moisture innovations at 12th sigma level in the middle of the domain. Middle panels correspond to GSI runs with NCEP global BE. Left & right panels correspond to WRF-ARW BE without tuning and with horizontal length-scale tuning = 2.0.

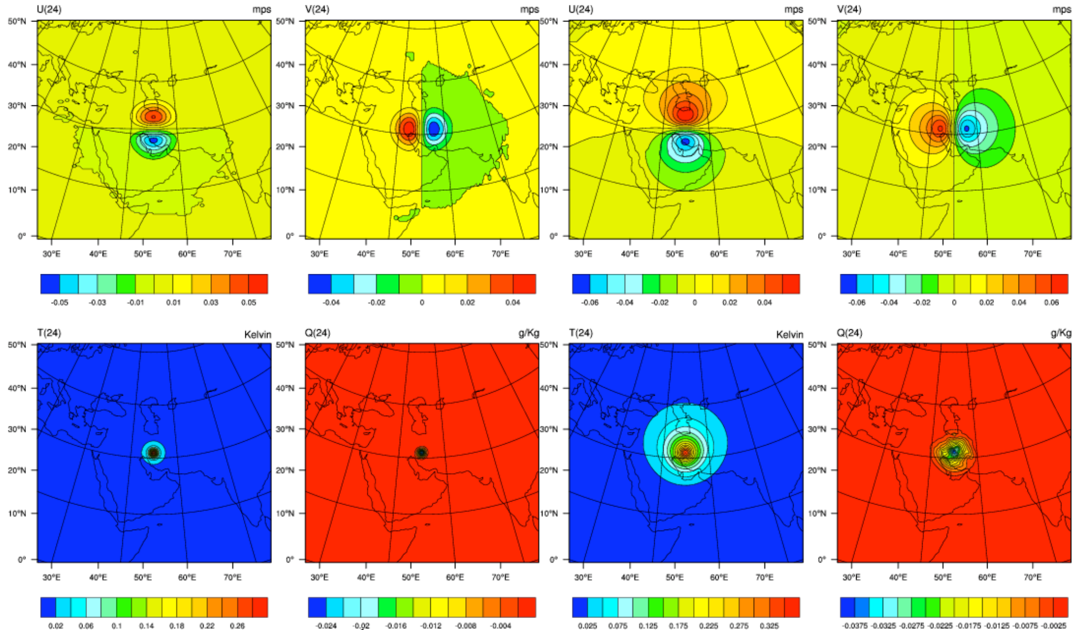

Figure 6.9: Horizontal (XY at 12th sigma level) cross-section of U, V , T and Q with 1.0 Kelvin temperature innovations at 12th sigma level in the middle of the domain. Left & right panels correspond to WRF-ARW BE without tuning and with global NCEP BE respectively.

Acknowledgements: Authors may like to acknowledge NCEP for providing GSI and the regional background error generation code. NCEP regional BE code was very useful in developing "stage2". Most of the code in "stage2" has been used as such from NCEP regional BE code. Help provided by Hui-Shao at DATC, Ming at NOAA, Boulder and especially Wan Shu at NCEP are highly appreciated. This work was supported by AFWA.

References:

Bannister, R.N. (2008): A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances, *Q.J.R.Meteorol.Soc.*, **134**, 1951– 1970.

Bannister, R. N. (2008): A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the fore-cast error covariance statistics, *Q.J.R.Meteorol.Soc.*, **134**, 1971–1996.

Daley, R. (1991): *Atmospheric Data Analysis*. Cambridge University Press, 457 pp.

Pannekoucke, O., L. Berre and G. Desroziers (2008): Background error correlation length-scale estimates and their sampling statistics. *QJRMS.* **134**, 497-511.

Wan-Shu Wu, R. James Purser and David F Parrish (2002): Three-Dimensional Variational Analysis with Spatially Inhomogeneous Covariances. *Mon. Wea. Rev*. **130**, pp. 2905-16.

Appendix:

The code is organized as follows. Source code for all the three stages resides under "gen_be" sub-directory of WRFDA. Parallel scripts for running this code are under "var/scripts/gen_gbe" sub-directory. For display of BE statistics, graphics are organized under "graphics/ncl/gen_be". All the desired executables gets generated with normal "*compile all_wrfvar*" command along with all other WRFDA executables. For any desired domain, corresponding BE file may be produced by executing a suitable wrapper script, like normal "*gen_be*" utility. A sample "*wrapper_gen_be_gsi.ksh*" script is created under "var/scripts/gen_be" subdirectory as an example to run this code for CONUS 200 Km domain. Some of the important "subroutines" used in this code and there functions are explained in the following Table.

Name	Description	Used in
da_uv_to_div_c	Given U & V in Arakawa-C grid, it computes horizontal divergence on mass grid	Stage0
da_uv_to_vor_c	Given U & V in Arakawa-C grid, it computes horizontal vorticity on mass grid	Stage0
da_del2a_to_a	Solves Poisson equation using "spectral" method	Stage0
da_sor	Solves Poisson equation using "relaxation" method	Stage0
da_psichi_to_uv_c	Given stream function and velocity potential it computes U & V wind components. It is activated for "inverse" transform tests.	Stage0
get_fixed_fields	Gathers geo-graphical info like, lat/lon of grid locations, sigma values, map-projection related info etc. Given "lat_bins_in_deg" it generates all the necessary horizontal bins info for producing BE.	Stage2
compute_reg_coeff	Computes χ , T and ps_u regression coefficients	Stage2
compute_variance_and_len_scales	Computes variance and length-scales (both horizontal and vertical) for all the control variables.	Stage2
vert_corr	Computes vertical correlations of 3D-fields	Stage2
read_wrf_arw	Reads and processes all 2 & 3D fields written in stage1	Stage2
psi_tilde	Does EOF's decomposition of stream function	Stage2
choldc	Does Chelosky decomposition for a	Stage2

	symmetric matrix	
cholsl	Given Chelosky decomposed symmetric matrix A and a matrix B, solves AX =B for X.	Stage2
horz_lenscale	Computes horizontal length-scale for a 2D filed.	Stage2
get_basis	Computes EOF's for vertical error correlation matrix	Stage2