# WRF-Var Background Error Generation

WRF Tutorial

July 28th 2005

Dale Barker, Yong-Run Guo (MMM Division, NCAR)
Hui-Chuan Lin (CWB)
Mi-Seon Lee, Hyun-Cheol (Korean Meteorological Administration).

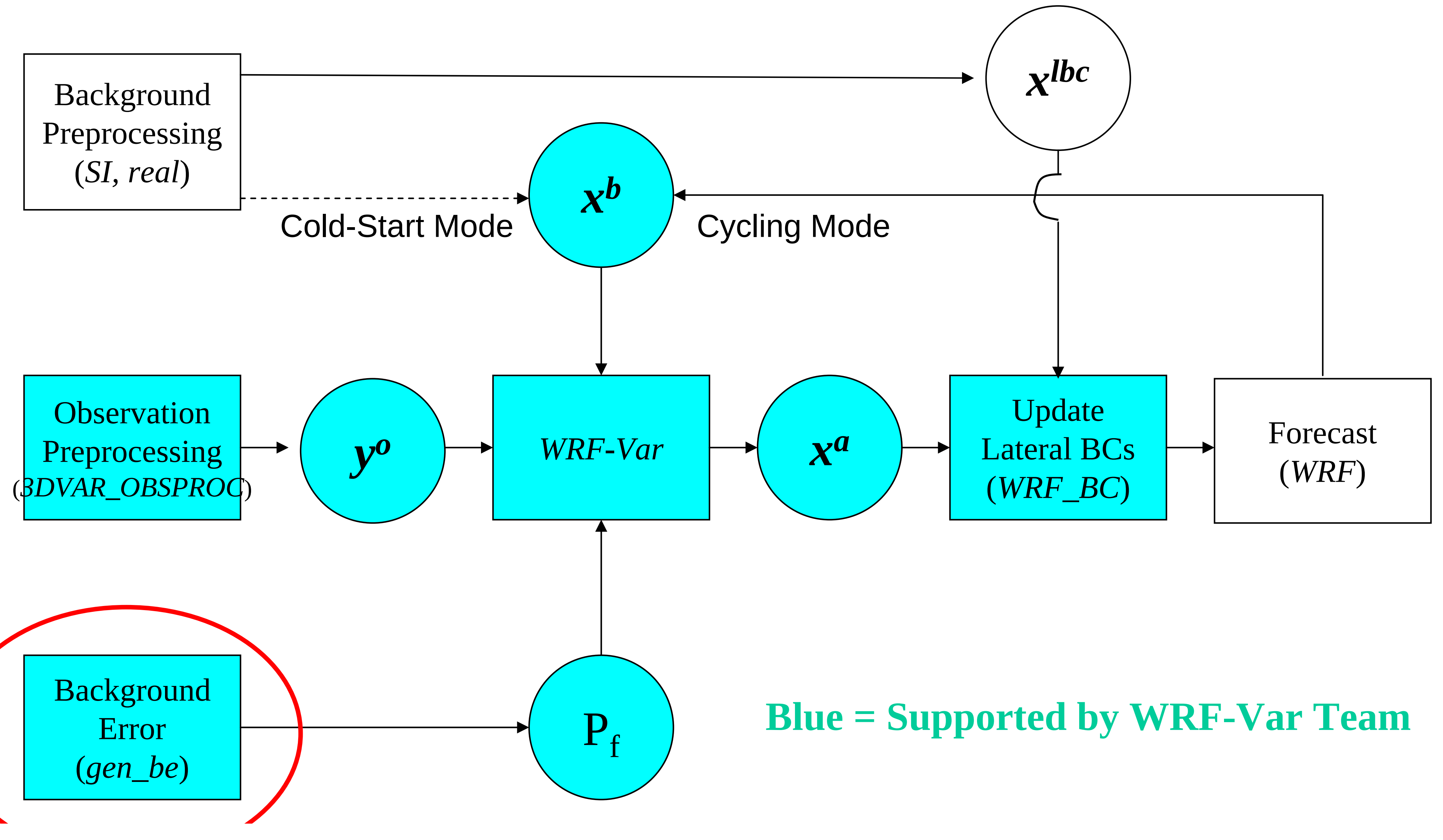
Email: dmbarker@ucar.edu

# Outline of Talk

1. Use of Background Errors in WRF-Var.
2. WRF-Var Background Error Covariance Generation (gen_be).
3. Error Tuning.
4. Future Work.

# 1. Use of Background Errors in WRF-Var

# WRF-Var in the WRF Modeling System



# Background Error (BE) Estimation in WRF-Var

The number 1 question from WRF-Var users is

**“What background error covariances are best for my application?”.**

Procedure:

1. Use default statistics files supplied with code (MM5, GFS-based).
2. Create you own, once you have run your system for ~a few weeks.
3. Implement, tune, and iterate.

A new utility *gen_be* has been developed at NCAR to calculate BEs.

# Incremental WRF-Var and Preconditioning

- *Define* ***analysis increments:*** $x^a = x^b + I\,x'$

- *Solve* ***incremental*** cost function:

$$J(\mathbf{x}) = \frac{1}{2} x'^T \mathbf{P_b}^{-1} \mathbf{x}' + \frac{1}{2} \sum_n [\mathbf{d} - \mathbf{y}']_n^2 / \sigma_{on}^2$$

where $y' = \mathbf{H}x'$, $\mathbf{d} = \mathbf{y} - H(\mathbf{x}^b)$

- *Define* ***preconditioned control variable*** $v$ space transform $x' = Uv$ where U transform CAREFULLY chosen to satisfy $\mathbf{P_b} = UU^T$.

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

$$J(\mathbf{x}) = \frac{1}{2} \sum_i v_i^2 + \frac{1}{2} \sum_n [\mathbf{d} - \mathbf{y}']_n^2 / \sigma_{on}^2$$

# Background Error Estimation for WRF-Var

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

$$\mathbf{P}_f = \overline{(\mathbf{x}^b - \mathbf{x}^t)(\mathbf{x}^b - \mathbf{x}^t)^T} \approx \overline{\mathbf{x}' \, \mathbf{x}'^T}$$

Two ways of defining *x'* in utility *gen_be*:

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

$$\mathbf{P}_f = \overline{\mathbf{x}' \, \mathbf{x}'^T} \approx A\overline{(\mathbf{x}^{t2} - \mathbf{x}^{t1})(\mathbf{x}^{t2} - \mathbf{x}^{t1})^T}$$

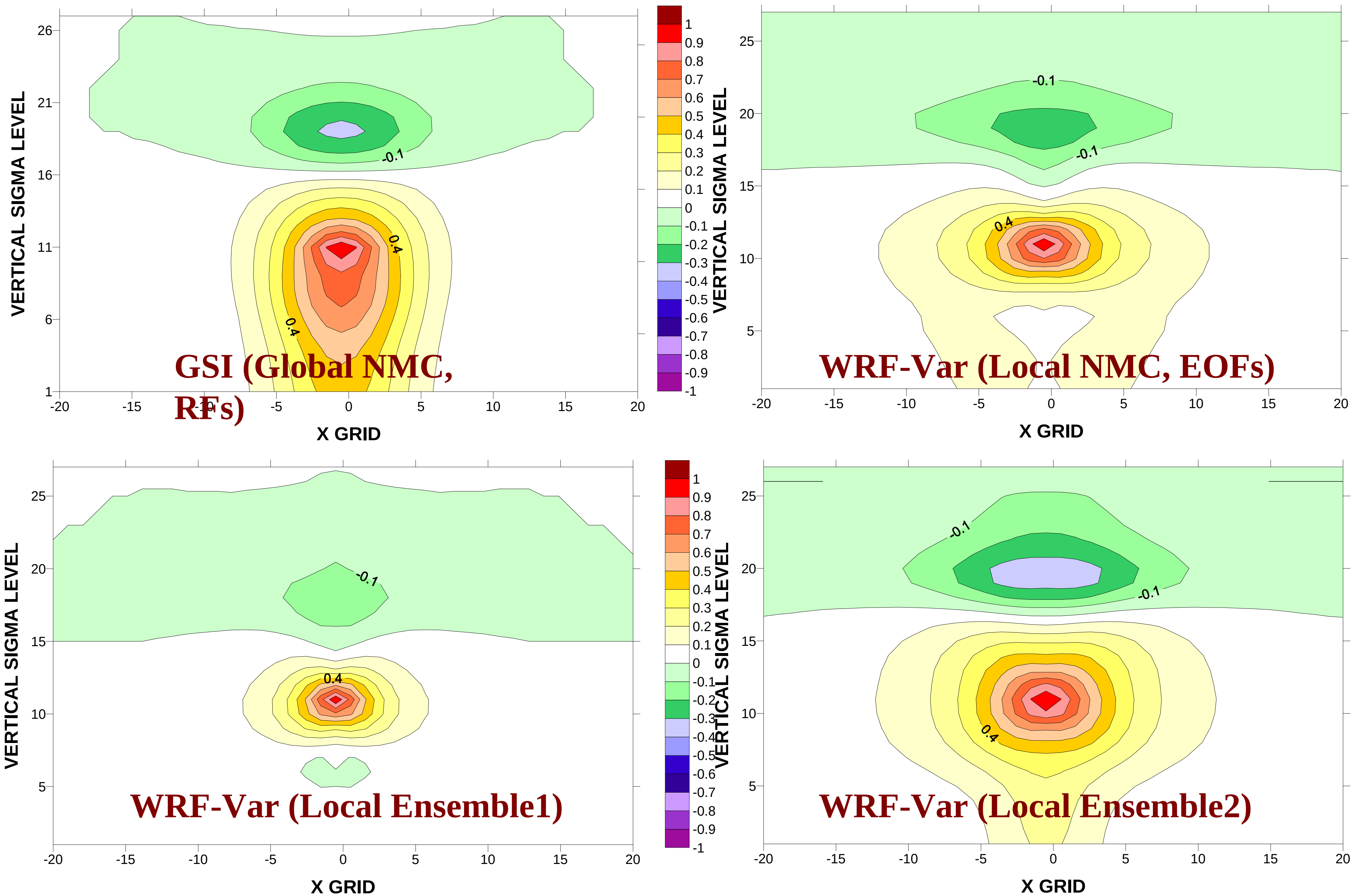
  where e.g. t2=24hr, t1=12hr forecasts…

- …or ensemble perturbations (Fisher 2003):

$$\mathbf{P}_f = \overline{\mathbf{x}' \, \mathbf{x}'^T} \approx C\overline{(\mathbf{x}^k - \langle \mathbf{x} \rangle)(\mathbf{x}^k - \langle \mathbf{x} \rangle)^T}$$

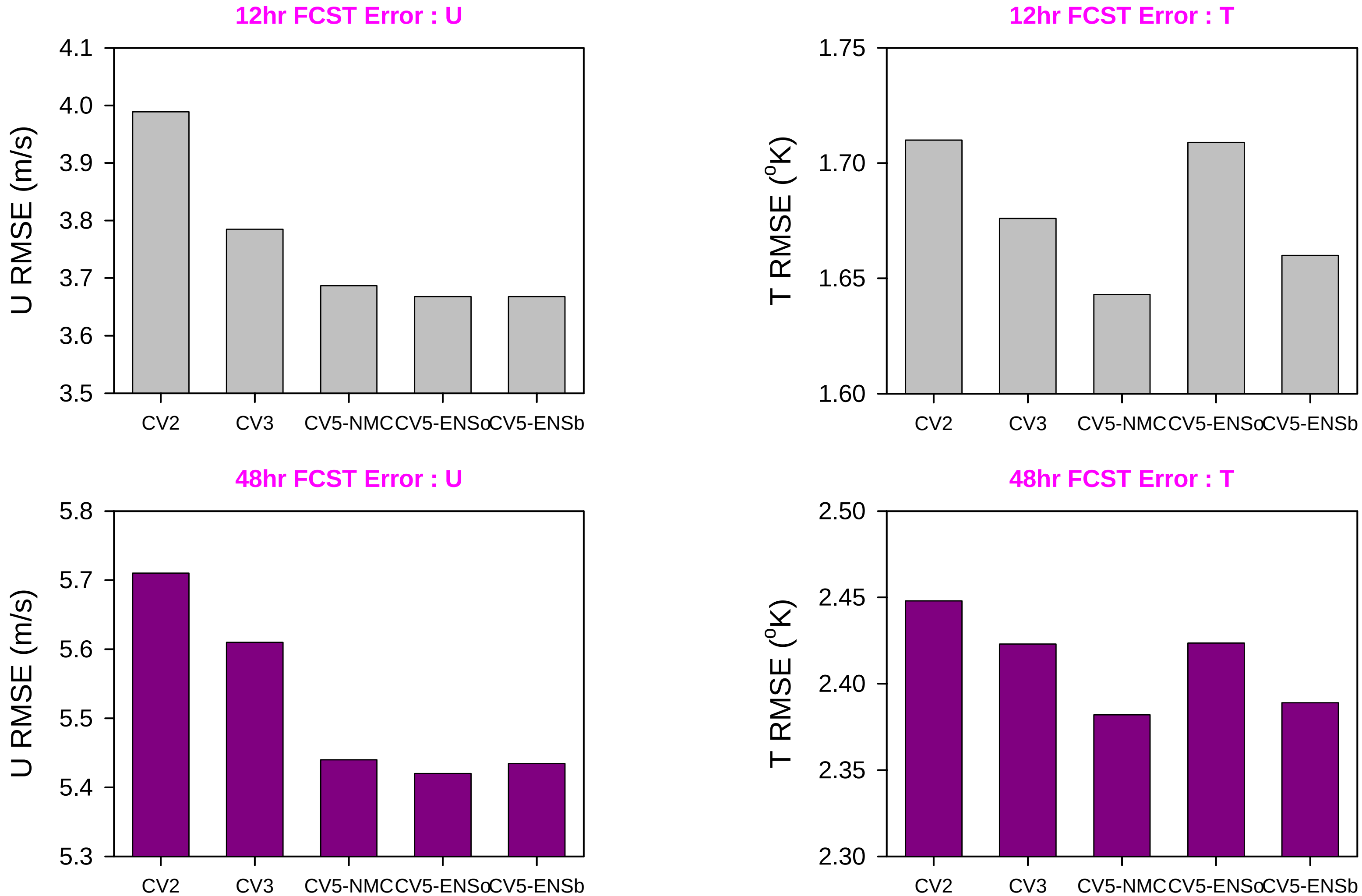
- Tuning via innovation vector statistics and/or variational methods.

# Analysis Sensitivity to Background Error Model




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

# Forecast Error (vs. sonde) Sensitivity to BE Stats




Courtesy Mi-Seon Lee

CV2 = Global statistics (MM5), CV3 = Global Statistics (GFS), CV5 = Local statistics (WRF)

NMC = NMC-method, ENS = Perturbed observation ensemble.

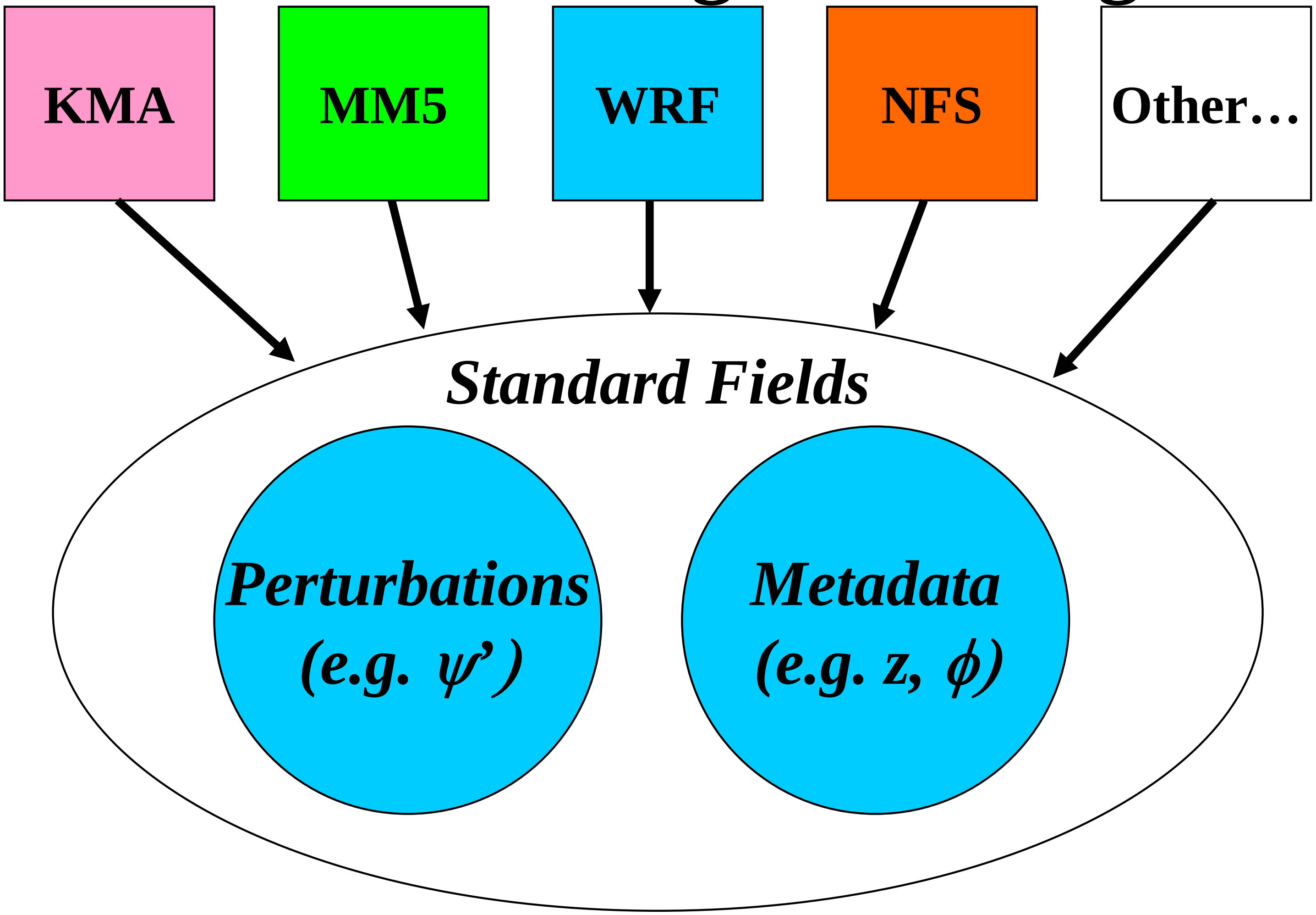
# 2. WRF-Var Background Error Covariance Generation (gen_be)

# Background Error (BE) Estimation in WRF-Var

A new utility *gen_be* has been developed at NCAR to calculate BEs. The new *gen_be* code is split into a number of stages:

- **Stage0: Convert model-specific data to "standard fields".**
- Stage1: Remove time-domain mean from fields.
- Stage2: Calculate regression coefficients, and use them to define "unbalanced" control variables.
- Stage3: Calculate vertical error covariances (eigenvectors and eigenvalues) for control variables.
- Stage 4: Calculate horizontal error correlations: lengthscales (in regional domains), and "power spectra" in global domains.

# BE Generation: gen_be_stage0

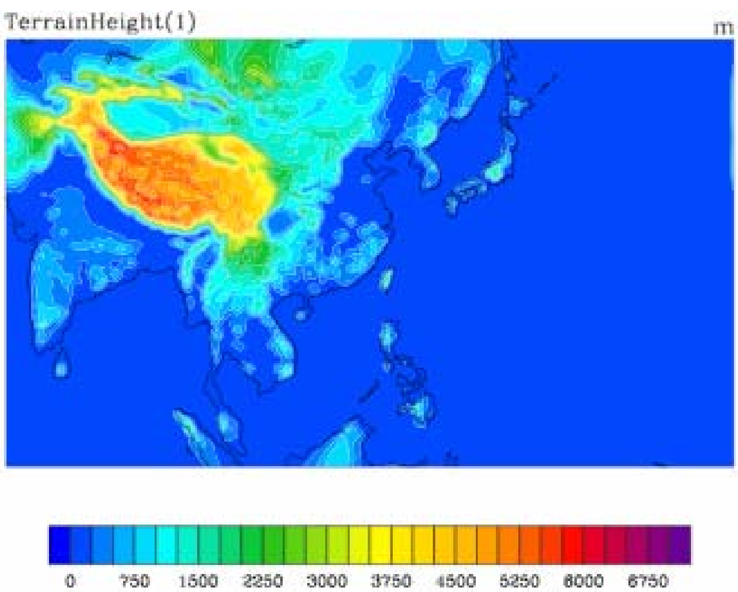



- Stage0 is the only stage that is model specific (e.g. gen_be_stage0_nfs).
- Perturbed fields are either forecast differences (NMC-Method) or ensemble perturbations (ensemble method).
- Standard fields are perturbations of streamfunction, velocity potential, temperature, relative humidity, and surface pressure.

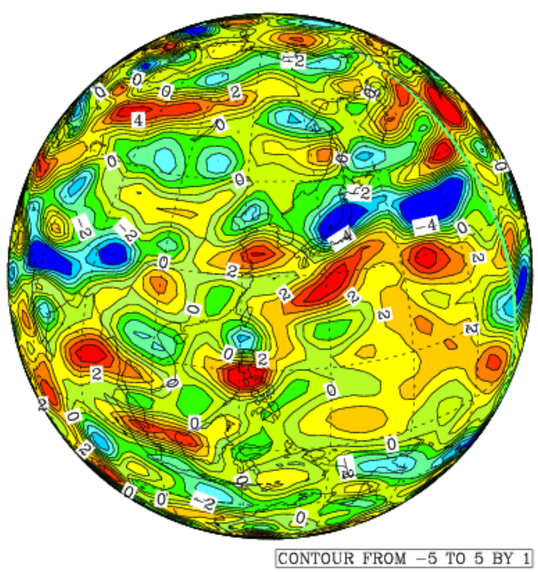
# Background Error (BE) Estimation for different domains/models/times

In the following results, we apply gen_be stages 1-4 to a variety of models to see robustness of results:

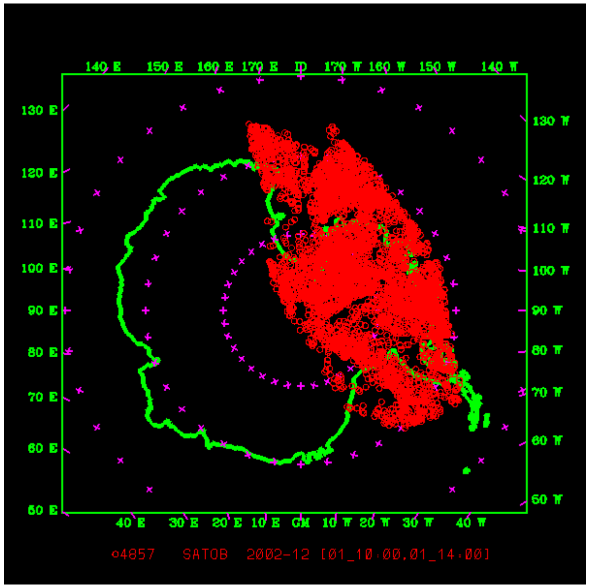
CWB 45km WRF




Korean T213 Global:




AMPS 30km WRF



# Background Error (BE) Estimation in WRF-Var

A new utility *gen_be* has been developed at NCAR to calculate BEs. The new *gen_be* code is split into a number of stages:

- Stage0: Convert model-specific data to "standard fields".
- Stage1: Remove time-domain mean from fields.
- **Stage2: Calculate regression coefficients, and use them to define "unbalanced" control variables.**
- Stage3: Calculate vertical error covariances (eigenvectors and eigenvalues) for control variables.
- Stage 4: Calculate horizontal error correlations: lengthscales (in regional domains), and "power spectra" in global domains.

# WRF-Var Control Variable Transform

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

$$\mathbf{x}' = U\mathbf{v} = U_p U_v U_h \mathbf{v}$$

Define control variables:

$$\psi'$$

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

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

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

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

# Balance Via Statistical Regression

**Regression Coefficients** after Wu et al (2002):

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

# Balance Via Statistical Regression

**Regression Coefficients** after Wu et al (2002):

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

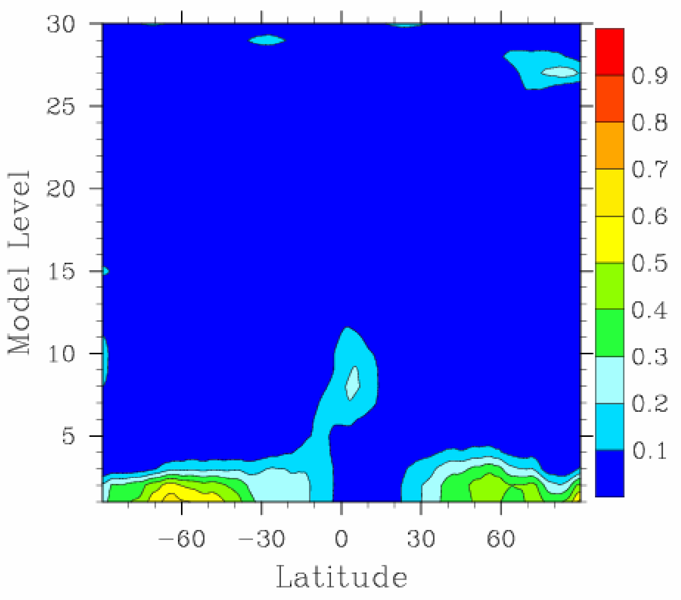
# Balance Via Statistical Regression

**Regression Coefficients** after Wu et al (2002):

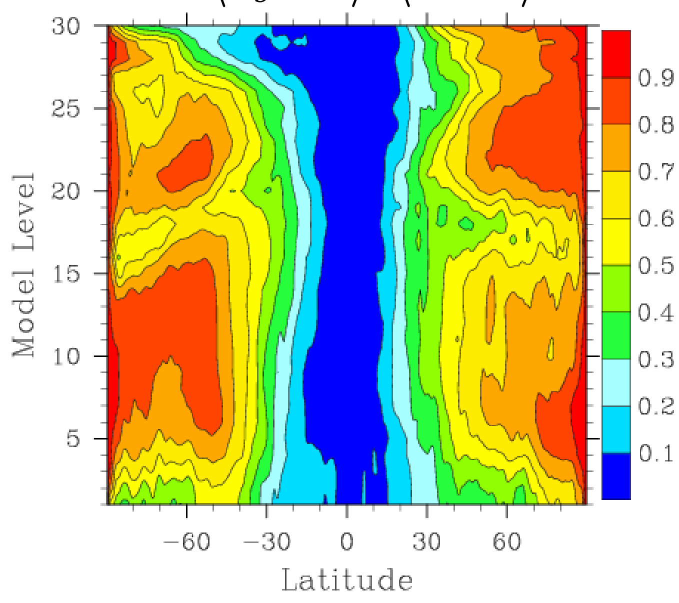
$$\chi_b' = c\psi' \qquad T_b'(k) = \sum_{k1} G(k,k1)\psi'(k1) \qquad p_{sb}' = \sum_{k} W(k)\psi'(k)$$

**Explained covariance due to balance constraints (Jan 2004 global data - NMC-method applied)**:

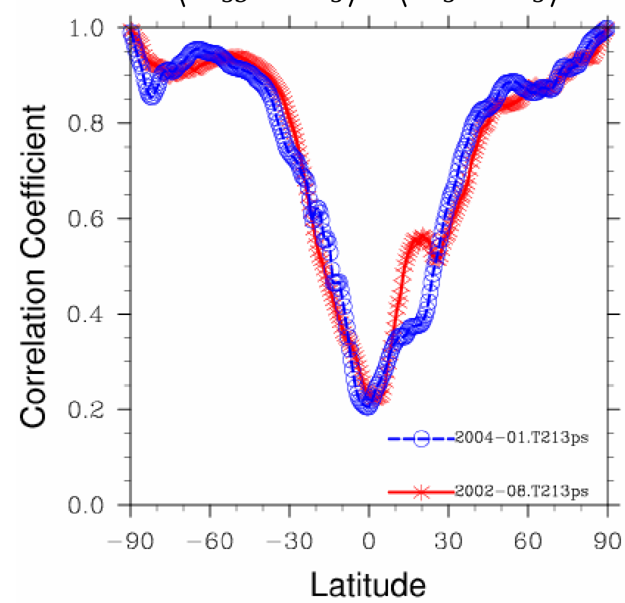
$\langle \chi_b \bullet \chi \rangle / \langle \chi \bullet \chi \rangle$




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




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

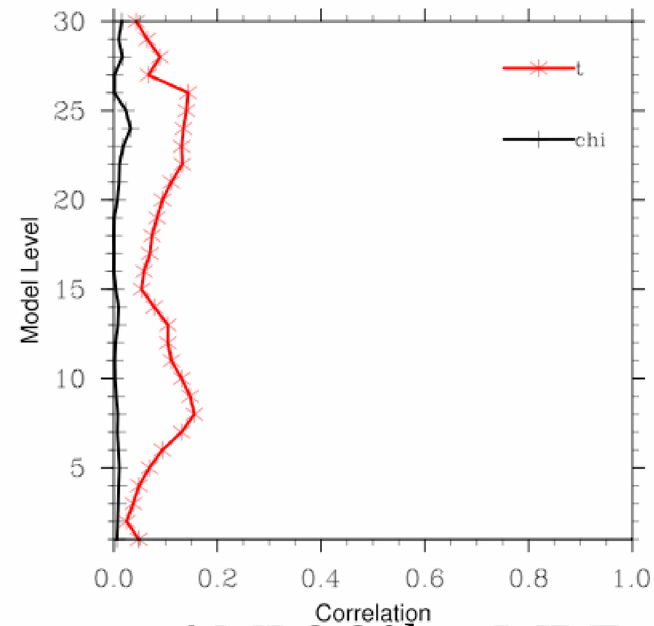


NCAR

Mesoscale & Microscale Meteorology @ NCAR

# Stage2: CWB/KMA/AMPS Comparison

CWB 45km WRF:

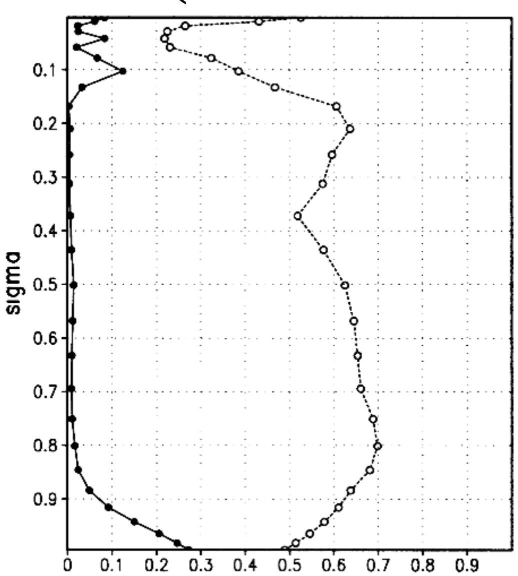



Calculate domain averaged correlations:

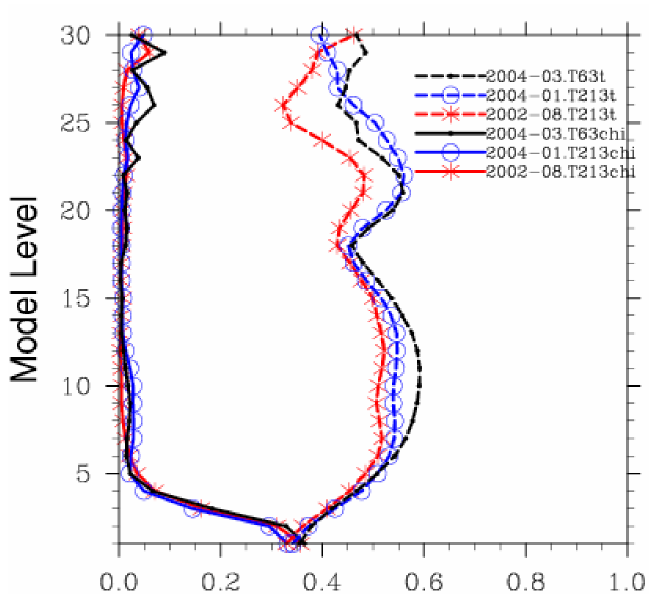
$$\langle T_b \bullet T \rangle / \langle T \bullet T \rangle \qquad \langle \chi_b \bullet \chi \rangle / \langle \chi \bullet \chi \rangle$$

For a variety of models using NMC-method.

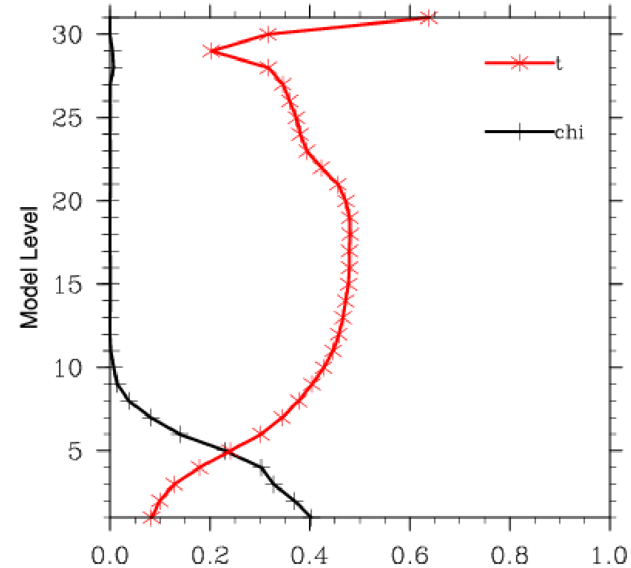
NCEP (Wu et al 2002):




Korean T213 Global:




AMPS 30km WRF:



# Background Error (BE) Estimation in WRF-Var

A new utility *gen_be* has been developed at NCAR to calculate BEs. The new *gen_be* code is split into a number of stages:

- Stage0: Convert model-specific data to "standard fields".
- Stage1: Remove time-domain mean from fields.
- Stage2: Calculate regression coefficients, and use them to define "unbalanced" control variables.
- **Stage3: Calculate vertical error covariances (eigenvectors and eigenvalues) for control variables.**
- Stage 4: Calculate horizontal error correlations: lengthscales (in regional domains), and "power spectra" in global domains.

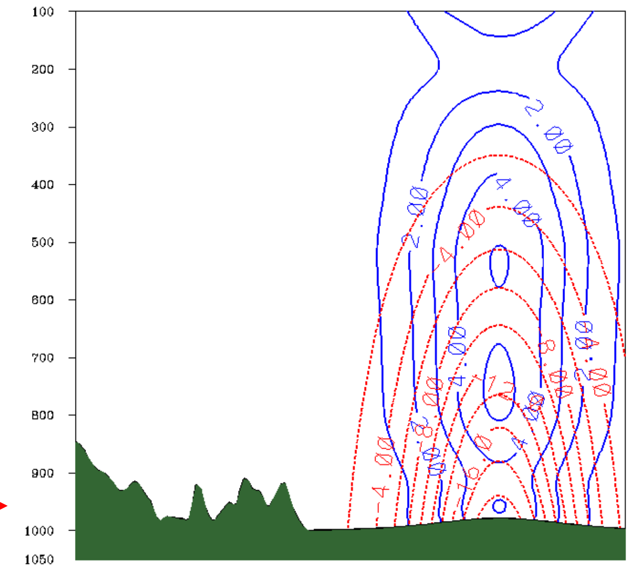
NCAR

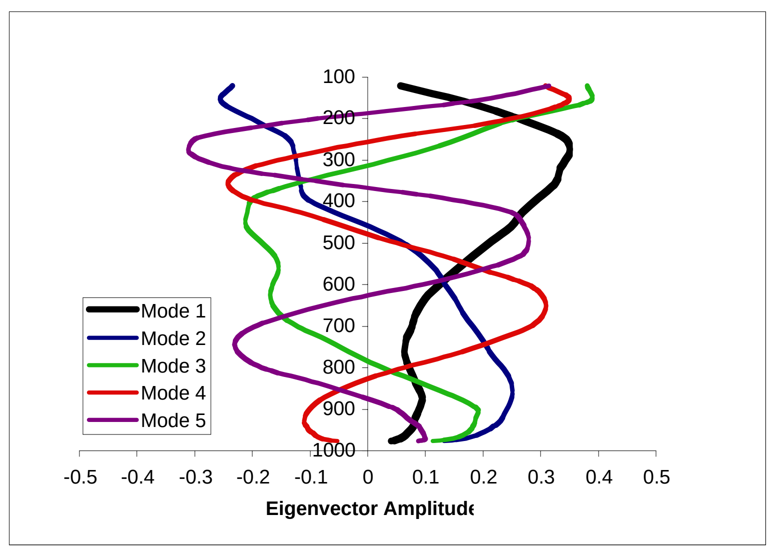
Mesoscale & Microscale Meteorology @ NCAR

# WRF-Var Control Variable Transform

$$x' = Uv = U_p U_v U_h v$$

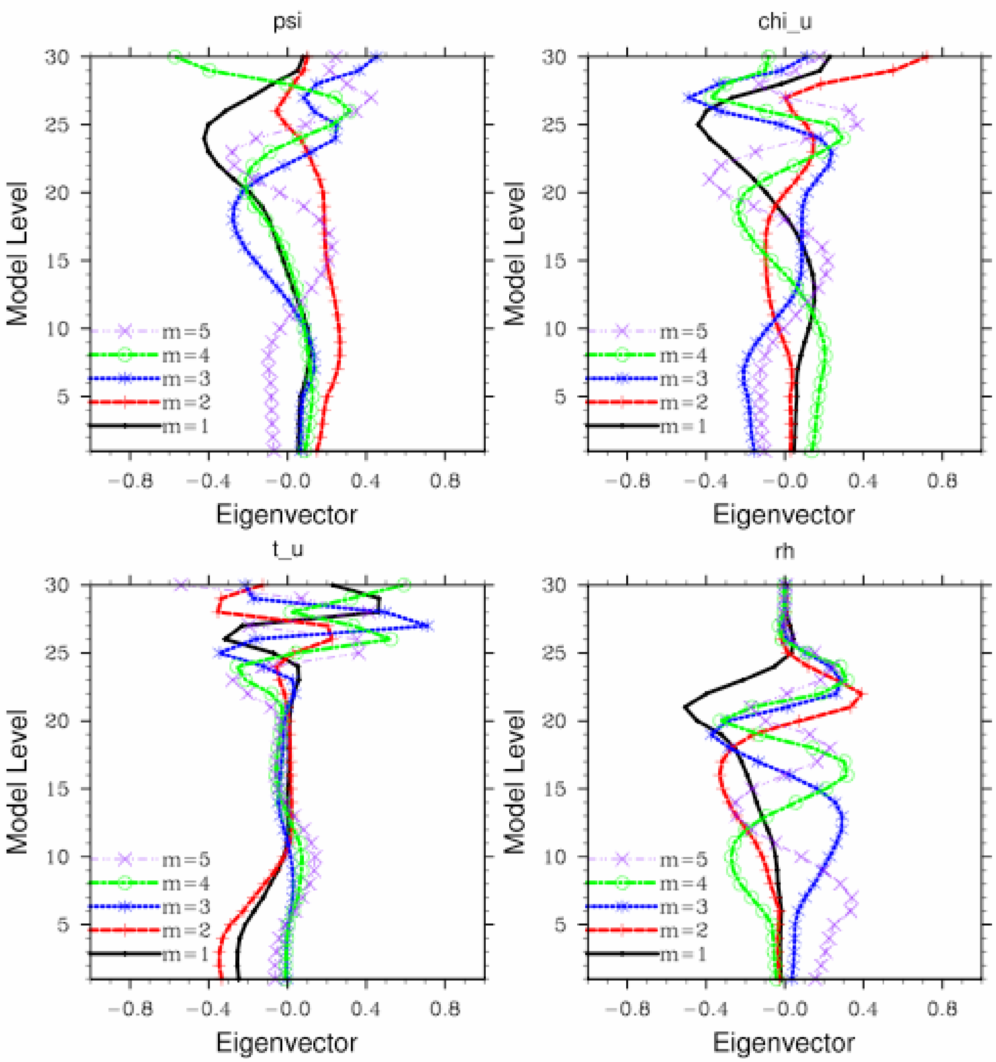
| cv_options | | 2 (original MM5) | 3(GSI) | 4 (Global) | 5(regional) |
|---|---|---|---|---|---|
| Analysis increments | $x'$ | $u',v',T',q',p_s'(i,j,k)$ | | | |
| Change of Variable | $U_p$ | $\psi',\chi',p_u',q'(i,j,k)$ | $\psi',\chi_u',T_u',\tilde{r}',p_{su}'(i,j,k)$ | | |
| Vertical Covariances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



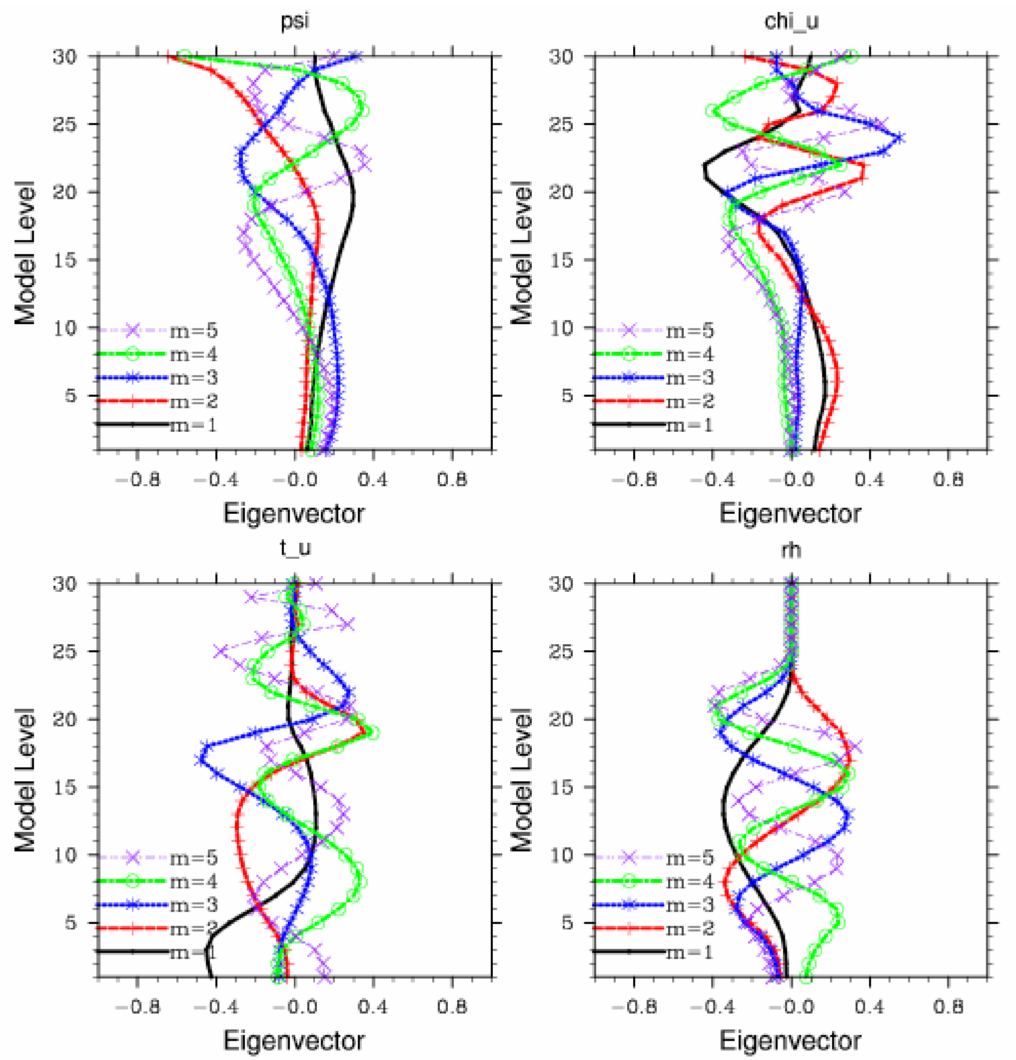



# Comparison of Domain-Averaged Eigenvectors $E_v^g$

CWB 45km WRF:



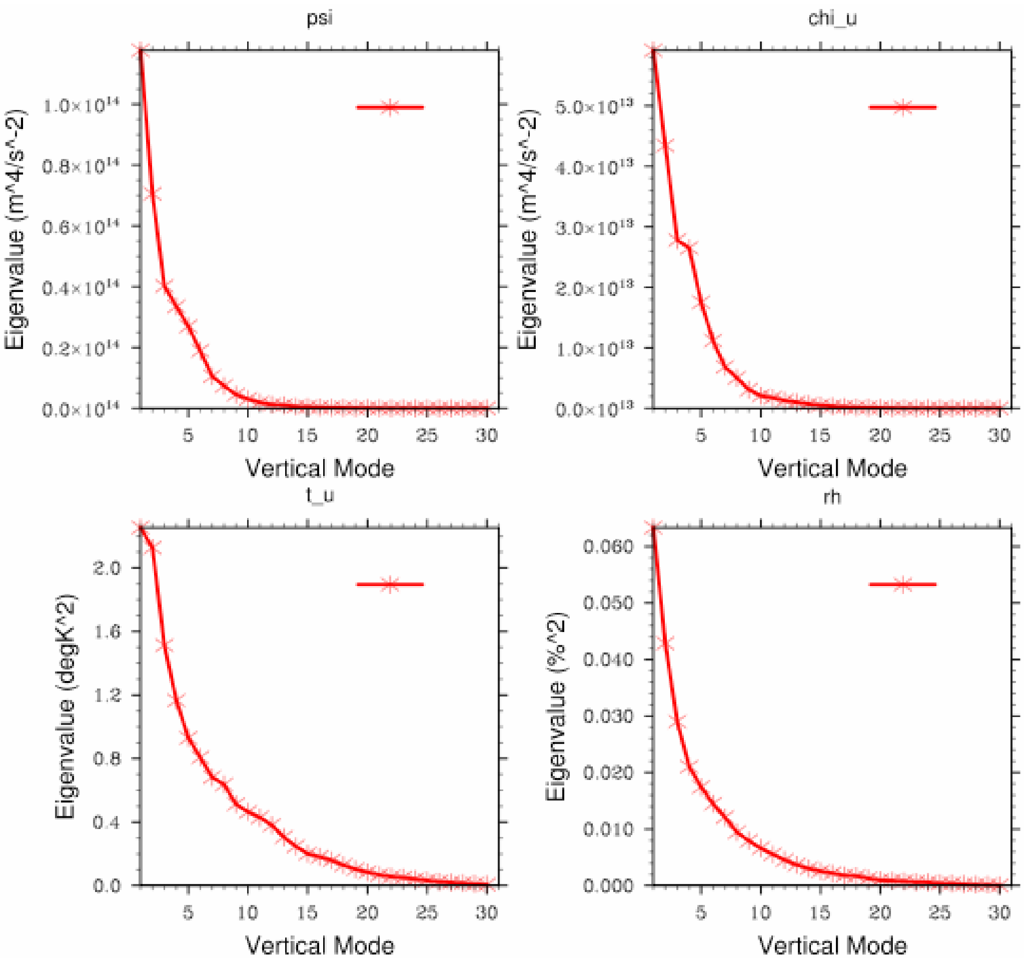

Korean T213 Global:



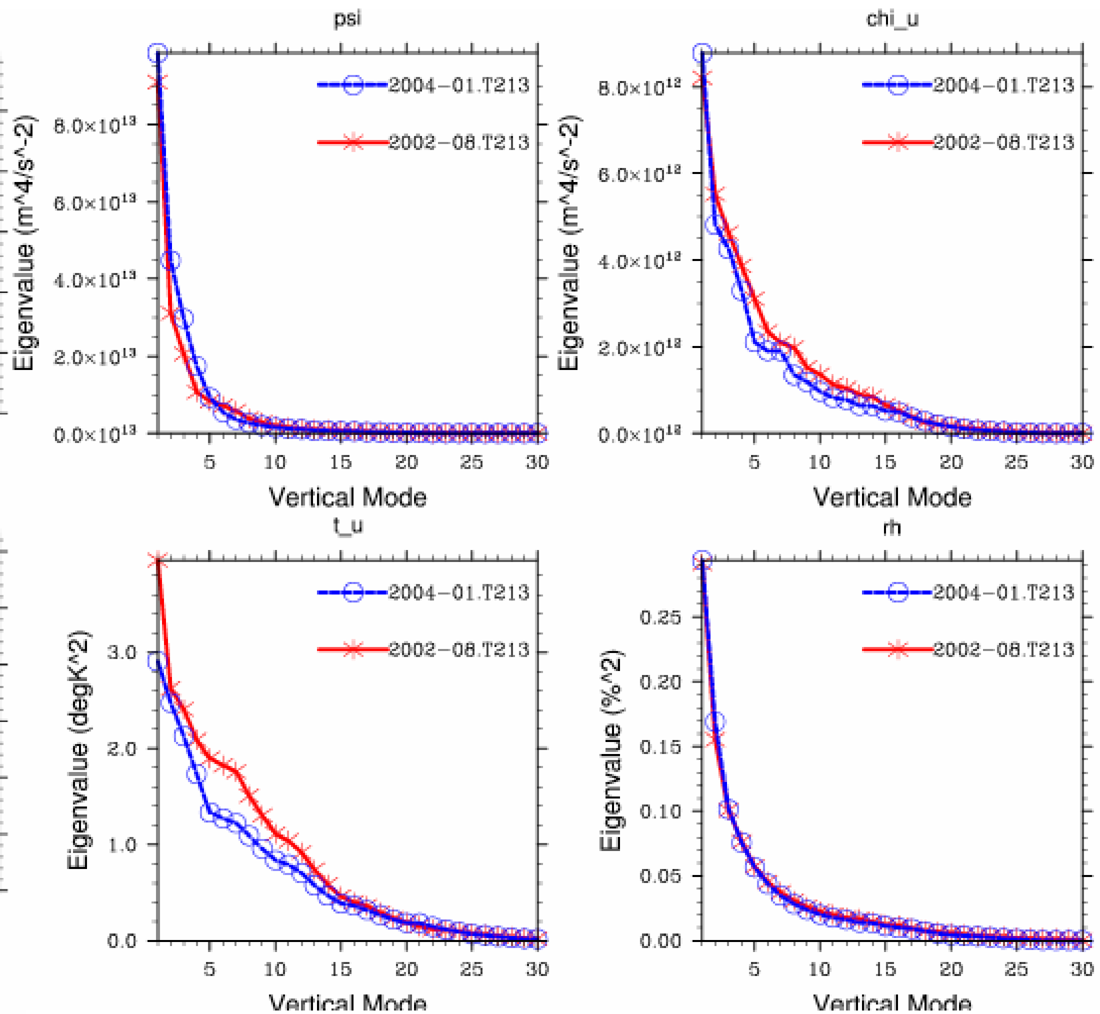

Conclusion: Significant differences in regional/global models (as one might expect), so it is is important to calculate domain-specific background error correlations.

# Comparison of Domain-Averaged Eigenvalues $\Lambda_v^g$

CWB 45km WRF:




Korean T213 Global:




Conclusion: Significant differences in regional/global models (as one might expect), so it is is important to calculate domain-specific background error variances.

# Background Error (BE) Estimation in WRF-Var

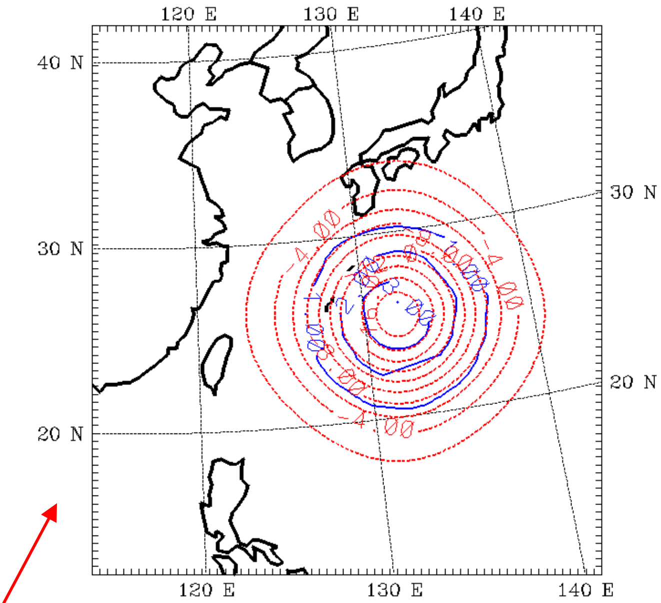
A new utility *gen_be* has been developed at NCAR to calculate BEs. The new *gen_be* code is split into a number of stages:

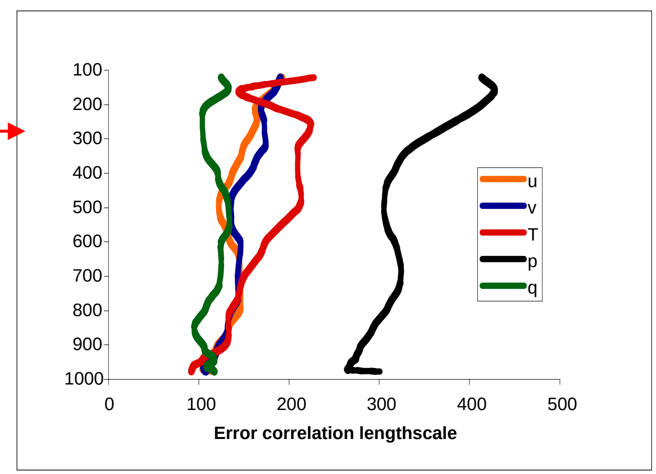
- Stage0: Convert model-specific data to "standard fields".
- Stage1: Remove time-domain mean from fields.
- Stage2: Calculate regression coefficients, and use them to define "unbalanced" control variables.
- Stage3: Calculate vertical error covariances (eigenvectors and eigenvalues) for control variables.
- **Stage 4: Calculate horizontal error correlations: lengthscales (in regional domains),** and "power spectra" in global domains.

# WRF-Var Control Variable Transform

$$x' = U\mathbf{v} = U_p U_v U_h \mathbf{v}$$

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



NCAR

Mesoscale & Microscale Meteorology @ NCAR

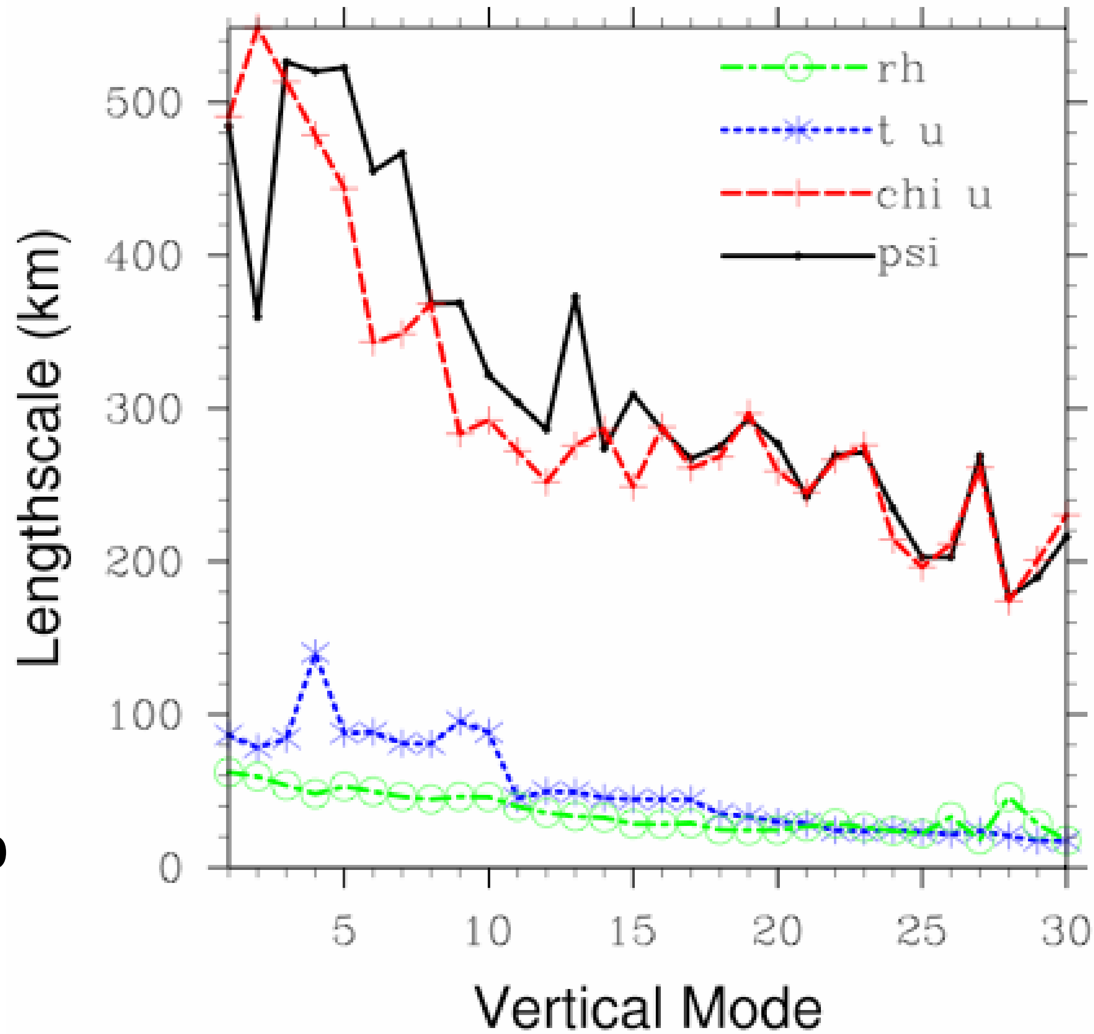
# Stage4: Regional model (CWB Sept. 2003) data

Horizontal error correlation lengthscales are produced as follows:

a. Bin each 2D perturbation field as a function of gridpoint separation.

b. Calculate correlation matrix for all data.

c. Fit Gaussian curve to data to estimate horizontal error correlation lengthscales.



# Components of Global Background Error Statistics

a) **Regression Coefficients** (used to derive "balanced" components of fields) after Wu et al (2002) (output from stage2):

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

b) **Eigenvectors/eigenvalues** of vertical component of background error (output from stage3):

Domain/time averaged "global" values - $\mathbf{P}_{fv}^{g} = \mathbf{E}_v^{g}\Lambda_v^{g}\mathbf{E}_v^{g^T}$

Longitude/time averaged "local" values - $\mathbf{P}_{fv}^{l}(\phi) = \mathbf{E}_{fv}^{l}(\phi)\Lambda_{fv}^{l}(\phi)\mathbf{E}_{fv}^{l^T}(\phi)$

a) **Power spectra** (model horizontal error correlations for each vertical mode - output from stage4_global):

$$F^m(\phi) = \frac{1}{I}\sum_{i=1}^{I} F(\lambda_i,\phi)e^{-im\lambda_i} \longrightarrow F_n^m = \sum_{j=1}^{J} W_j F^m(\mu_j)P_n^m(\mu_j) \longrightarrow D_n = \sum_{m=-n}^{n}\left(F_n^m\right)^2$$

# Components of Regional Background Error Statistics

a) **Regression Coefficients** (used to derive "balanced" components of fields) after Wu et al (2002) (output from stage2):

$$\chi'_b = c\psi' \qquad T'_b(k) = \sum_{k1} G(k,k1)\psi'(k1) \qquad p'_{sb} = \sum_{k} W(k)\psi'(k)$$

b) **Eigenvectors/eigenvalues** of vertical component of background error (output from stage3):

Domain/time averaged "global" values - $\mathbf{P}^g_{fv} = \mathbf{E}^g_v \Lambda^g_v \mathbf{E}^{g^T}_v$

Longitude/time averaged "local" values - $\mathbf{P}^l_{fv}(\phi) = \mathbf{E}^l_{fv}(\phi)\Lambda^l_{fv}(\phi)\mathbf{E}^{l^T}_{fv}(\phi)$

a) **Lengthscales** (model horizontal error correlations *s* for each vertical mode - output from stage4_regional):

$$B(r) = B(0)\,\mathbf{exp}[-r**2/8s**2]$$

# Conclusions

1. Background Error Statistics are a crucial input to 3/4D-Var systems.

2. BE estimation methods are somewhat ad-hoc, and require significant efforts to accumulate the necessary forecast input data (e.g. 1 month 24hr forecasts)

3. A new utility *gen_be* has been written to calculate BEs for WRF-Var applications (for a variety of models, and techniques).

4. The output of *gen_be* must be further tuned to optimize WRF-Var performance. A number of additional utilities are available to perform this work.