WRF-Var Background Error Generation

Dr. Dale M. Barker

MMM Division, NCAR,

PO Box 3000, Boulder, Colorado,

80307-3000, USA

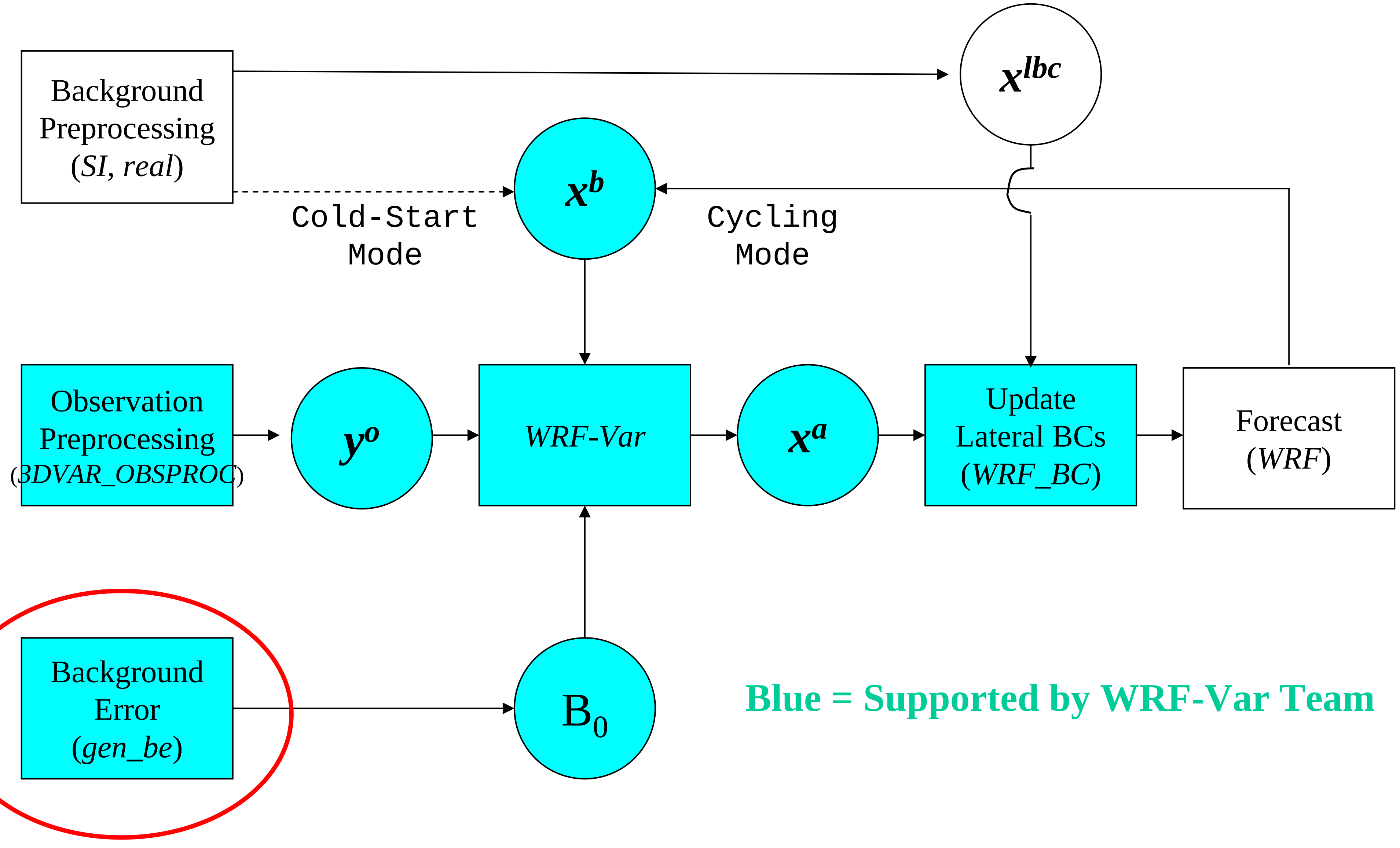
http://www.mmm.ucar.edu/people/barker/

Outline of Talk

- Use of Background Errors in WRF-Var.
- WRF-Var Background Error Covariance Generation (gen_be).

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.

Model-Based Estimation of Climatological Background Errors

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

$$\mathbf{B}_0 = \overline{(\boldsymbol{x}^b - \boldsymbol{x}^t)(\boldsymbol{x}^b - \boldsymbol{x}^t)^T} \approx \overline{\boldsymbol{x}'\boldsymbol{x}'^T}$$

Two ways of defining ***x’***:

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

$$\mathbf{B}_0 = \overline{\boldsymbol{x}'\boldsymbol{x}'^T} \approx A\overline{(\boldsymbol{x}^{t2} - \boldsymbol{x}^{t1})(\boldsymbol{x}^{t2} - \boldsymbol{x}^{t1})^T}$$

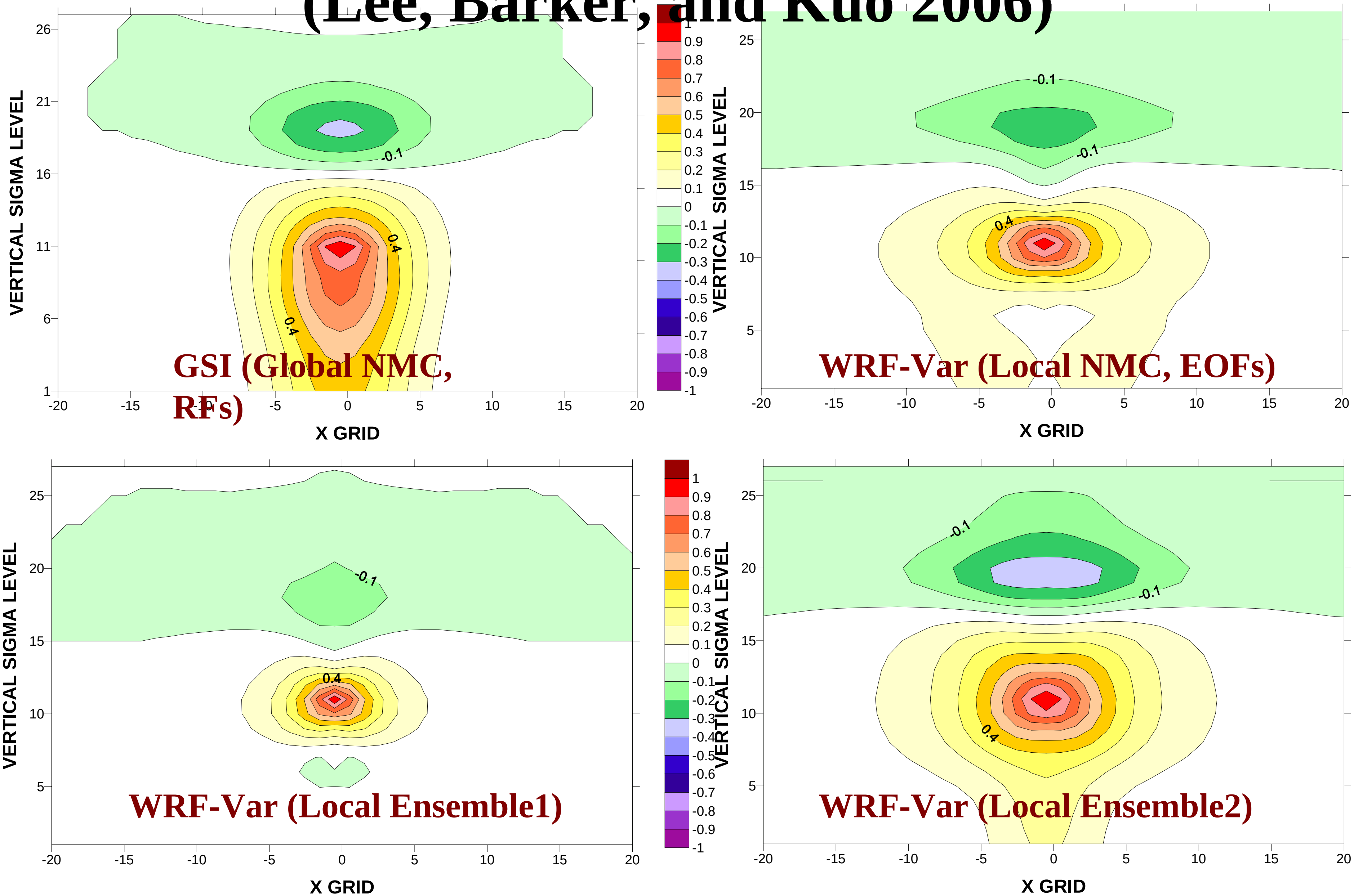
where e.g. t2=24hr, t1=12hr forecasts…

- …or ensemble perturbations (Fisher 2003):

$$\mathbf{B}_0 = \overline{\boldsymbol{x}'\boldsymbol{x}'^T} \approx C\overline{(\boldsymbol{x}^{k} - \langle \boldsymbol{x} \rangle)(\boldsymbol{x}^{k} - \langle \boldsymbol{x} \rangle)^T}$$

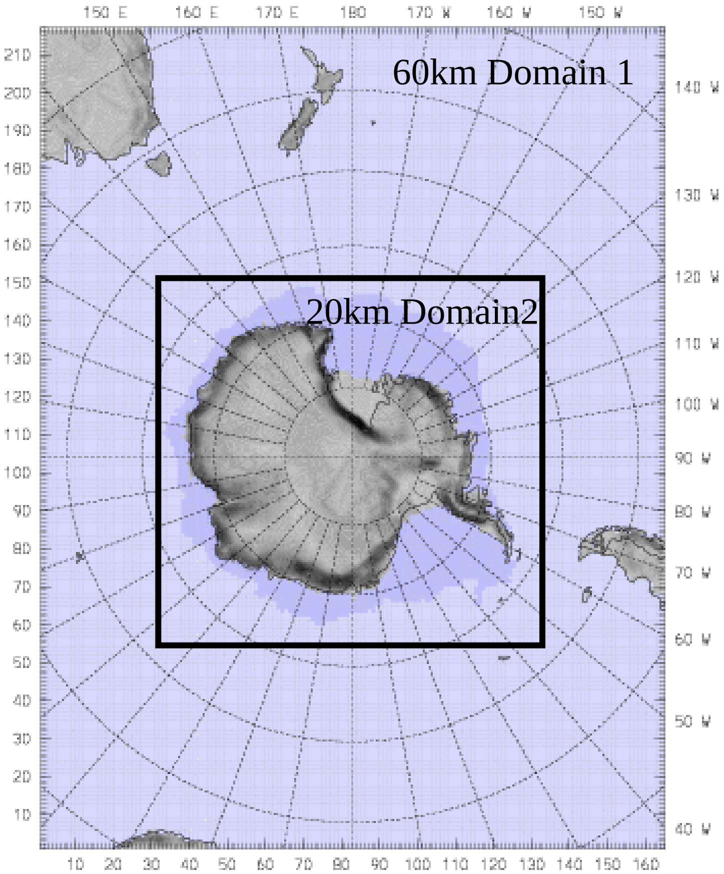
- Tuning via innovation vector statistics and/or variational methods.

Analysis Sensitivity to Background Error Model (Lee, Barker, and Kuo 2006)

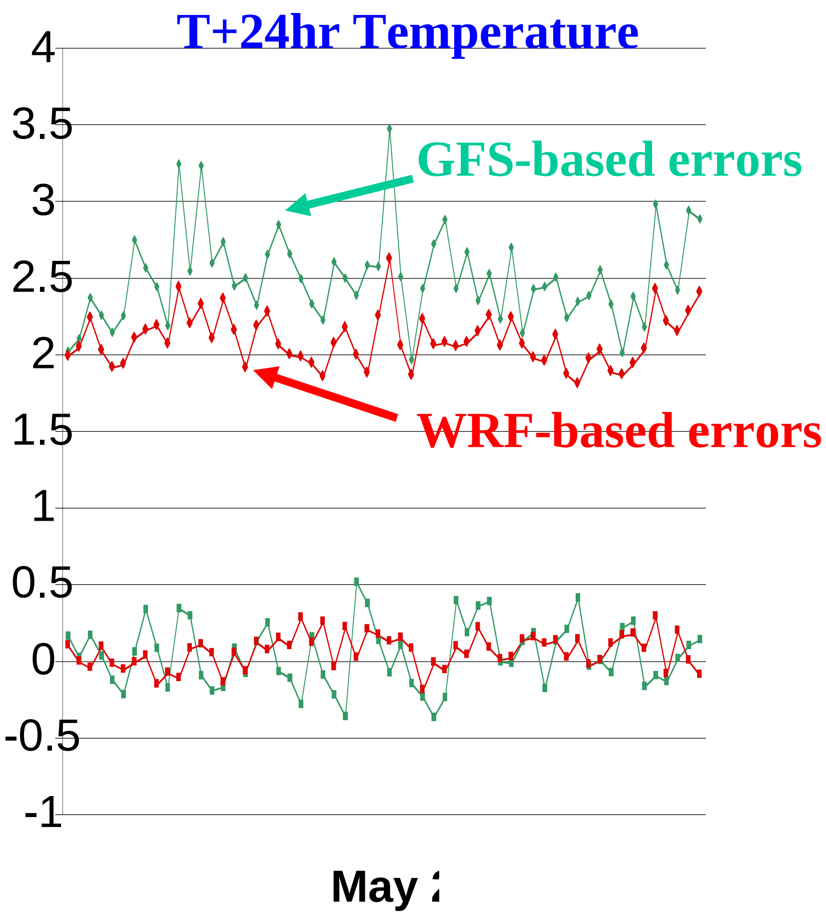


T increments : T Observation (1 Deg , 0.001 error around 850 hPa)

Sensitivity to Forecast Error Covariances in Antarctica (Rizvi et al 2006)



60km Verification (vs. radiosondes)



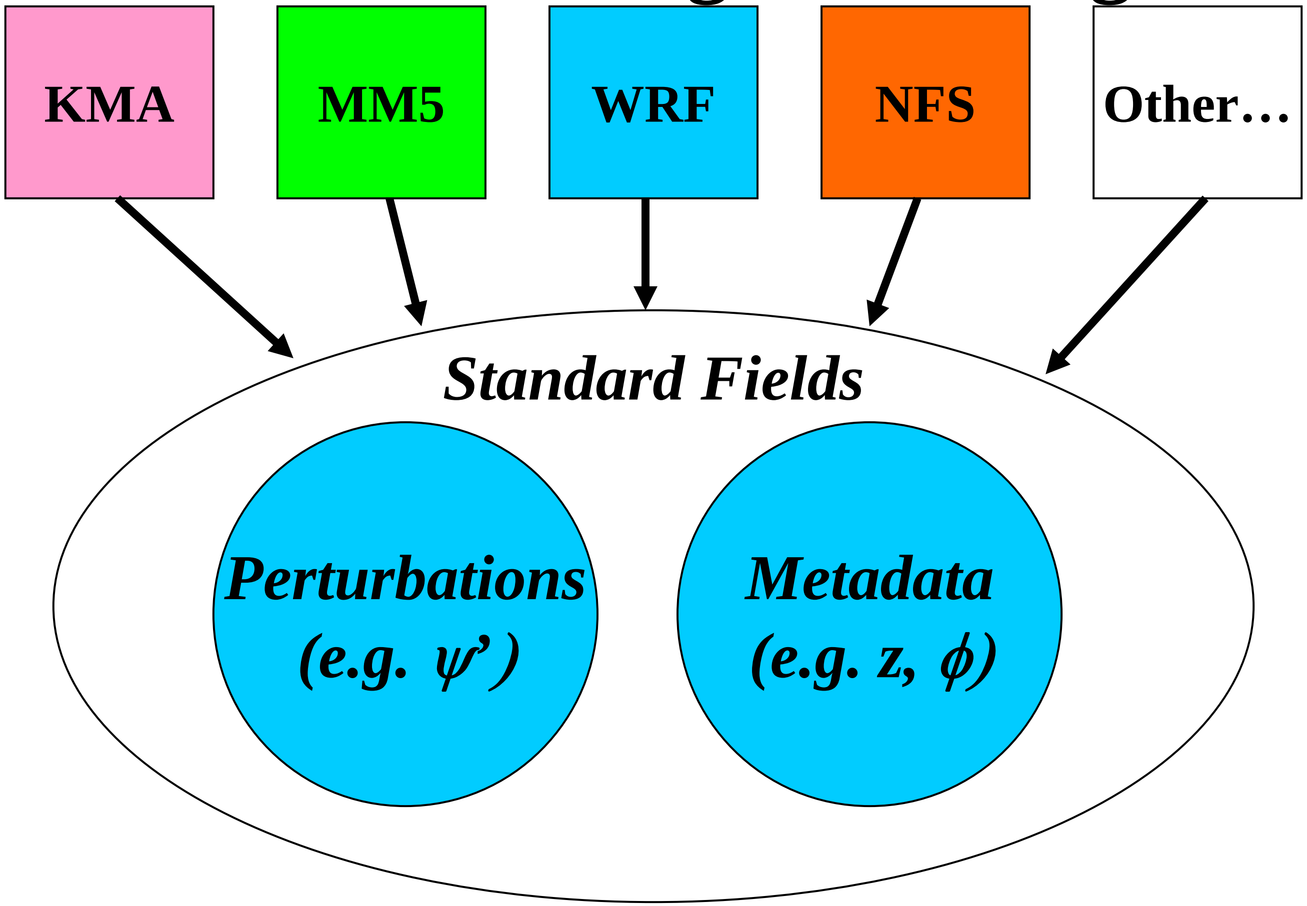
2. WRF-Var Background Error Covariance Generation (gen_be)

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

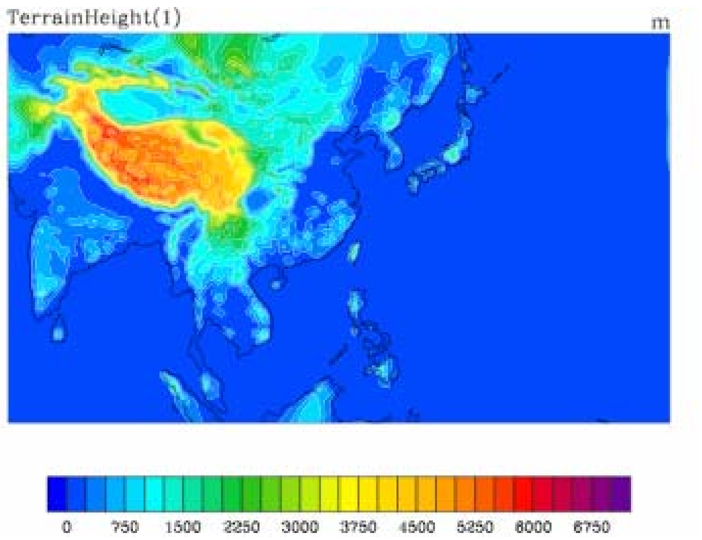


- Stage0 is the only stage that is model specific (e.g. gen_be_stage0_nfs).
- Perturbed fields are either forecast differences (NMC-Method) or ensemble perturbations (ensemble method).
- Standard fields are perturbations of streamfunction, velocity potential, temperature, relative humidity, and surface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

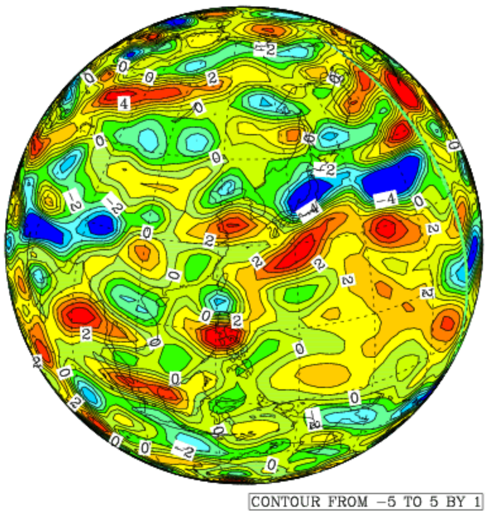
Background Error (BE) Estimation for different domains/models/times

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

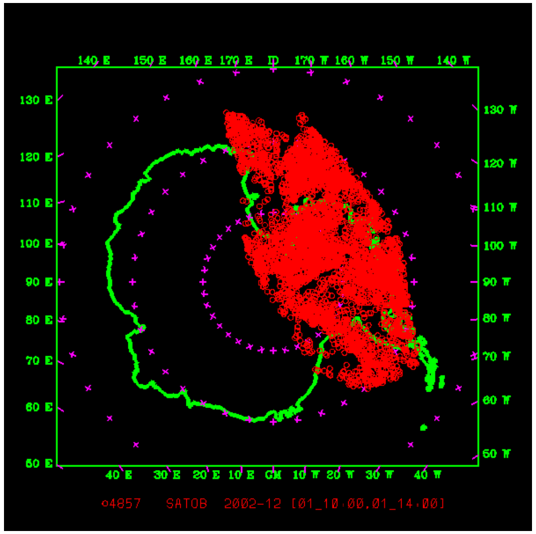
CWB 45km WRF



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 Background Error Modeling

cv_options		2 (original MM5)	3(GSI)	4 (Global)	5(regional)
Analysis increments	$\mathbf{x}'$	$u',v',T',q',p_s'(i,j,k)$			
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	$\mathbf{B}=\mathbf{E}\Lambda\mathbf{E}^{\mathrm{T}}$	RF	$\mathbf{B}=\mathbf{E}\Lambda\mathbf{E}^{\mathrm{T}}$	
Horizontal Correlations	U_h	RF		Spectral	RF
Control Variables	$\mathbf{v}$	$\mathbf{v}(i,j,m)$		$\mathbf{v}(l,n,m)$	$\mathbf{v}(i,j,m)$

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

Define control variables:

$$\psi'$$

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

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

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

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

3D-Var Statistical Balance Constraints

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

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

3D-Var Statistical Balance Constraints

Regression Coefficeints

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

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

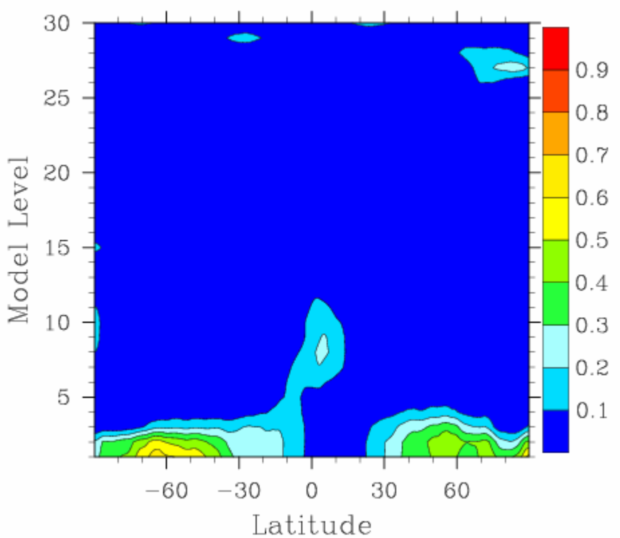
3D-Var Statistical Balance Constraints

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

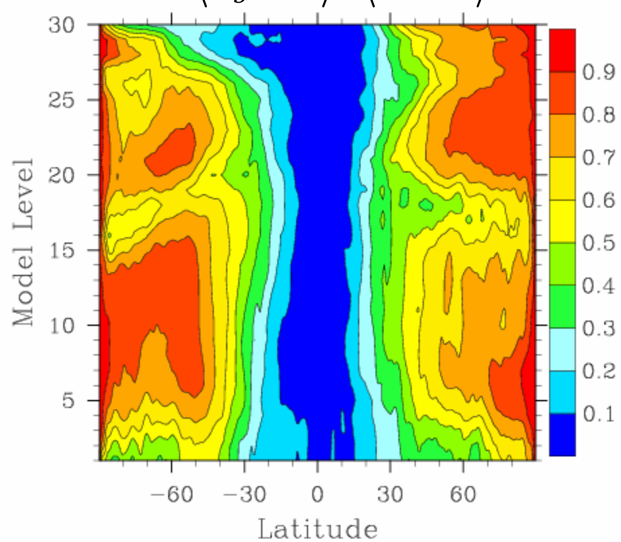
$$\chi_b' = c\psi' \quad T_b'(k) = \sum_{k1} G(k,k1)\psi'(k1) \quad p_{sb}' = \sum_k W(k)\psi'(k)$$

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

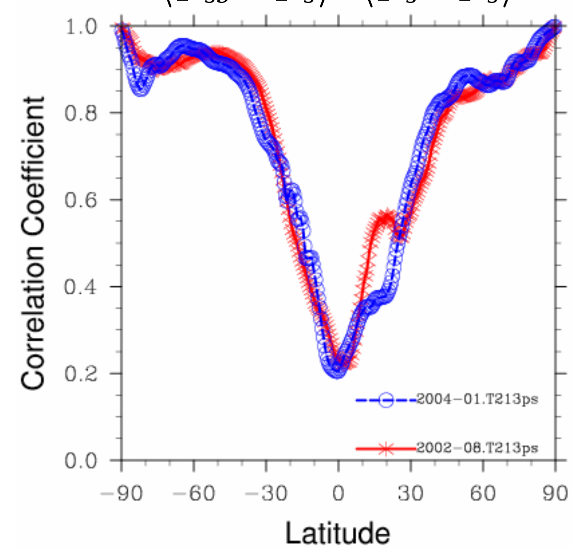
$$\langle \chi_b \bullet \chi \rangle / \langle \chi \bullet \chi \rangle$$



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



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



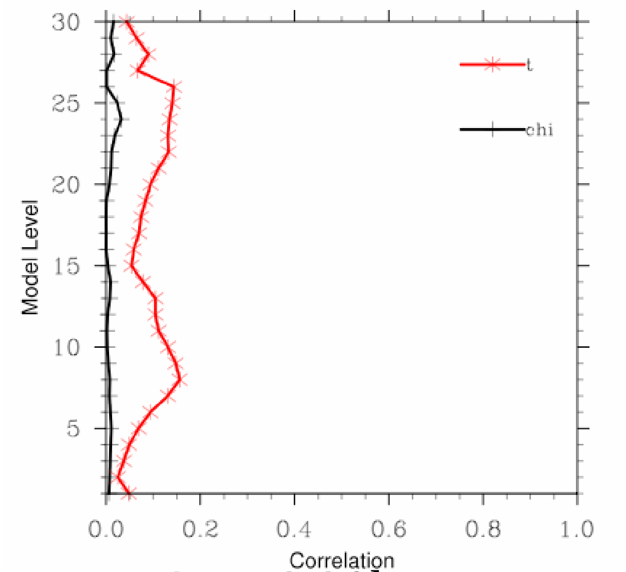
Stage2: CWB/KMA/AMPS Comparison

Calculate domain averaged correlations:

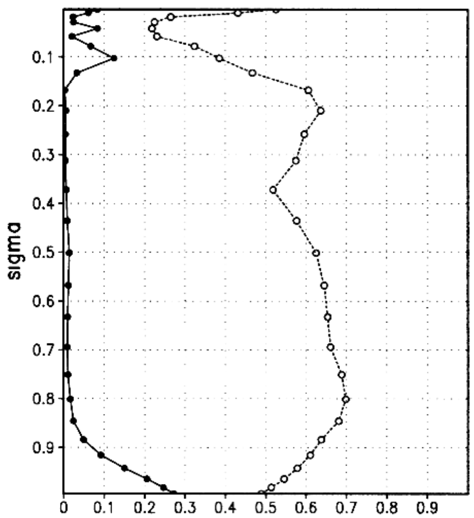
$$\langle T_b \bullet T \rangle / \langle T \bullet T \rangle \qquad \langle \chi_b \bullet \chi \rangle / \langle \chi \bullet \chi \rangle$$

For a variety of models using NMC-method.

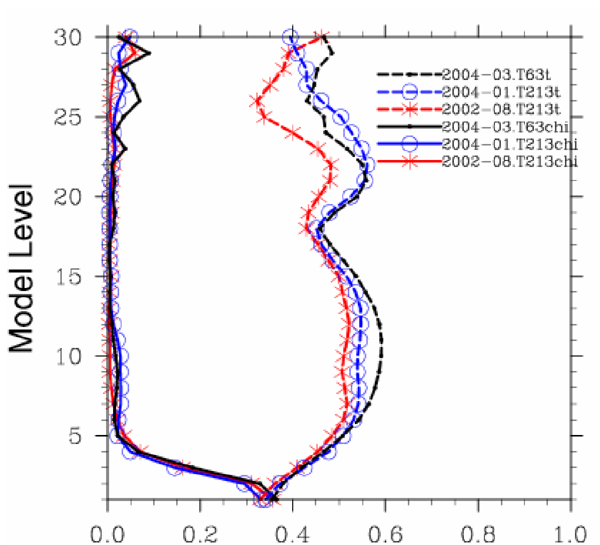
CWB 45km WRF:



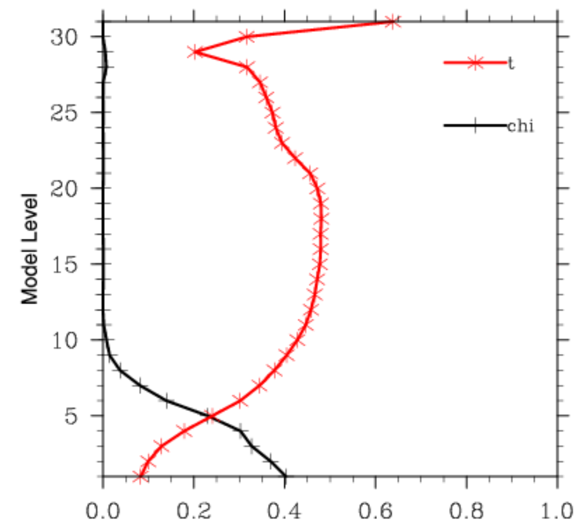
NCEP (Wu et al 2002):



Korean T213 Global:



AMPS 30km WRF:



Background Error (BE) Estimation in WRF-Var

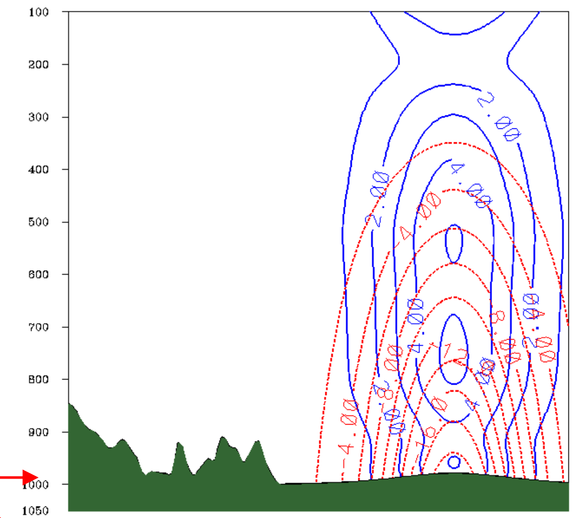
A new utility *gen_be* has been developed at NCAR to calculate BEs. The new *gen_be* code is split into a number of stages:

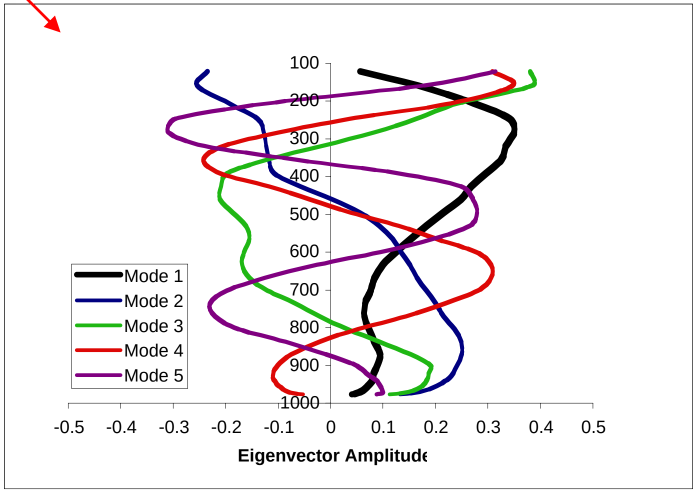
- Stage0: Convert model-specific data to “standard fields”.
- Stage1: Remove time-domain mean from fields.
- Stage2: Calculate regression coefficients, and use them to define “unbalanced” control variables.
- **Stage3: Calculate vertical error covariances (eigenvectors and eigenvalues) for control variables.**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

WRF-Var Background Error Modeling

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	$\mathbf{B}=\mathbf{E}\Lambda\mathbf{E}^{\mathrm{T}}$	RF	$\mathbf{B}=\mathbf{E}\Lambda\mathbf{E}^{\mathrm{T}}$	
Horizontal Correlations	U_h	RF		Spectral	RF
Control Variables	v	$\mathbf{v}(i,j,m)$		$\mathbf{v}(l,n,m)$	$\mathbf{v}(i,j,m)$

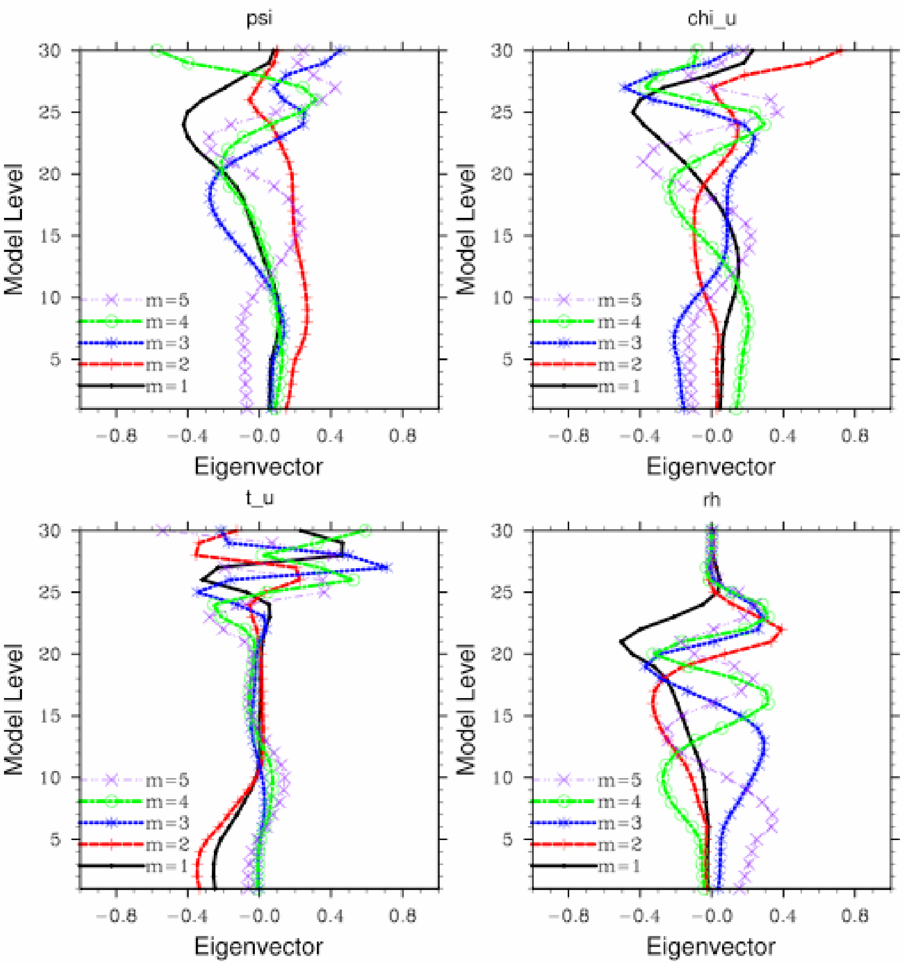
$$\delta\mathbf{x}(t_0) = \mathbf{U}\mathbf{v} = \mathbf{U}_p\mathbf{U}_v\mathbf{U}_h\mathbf{v}$$



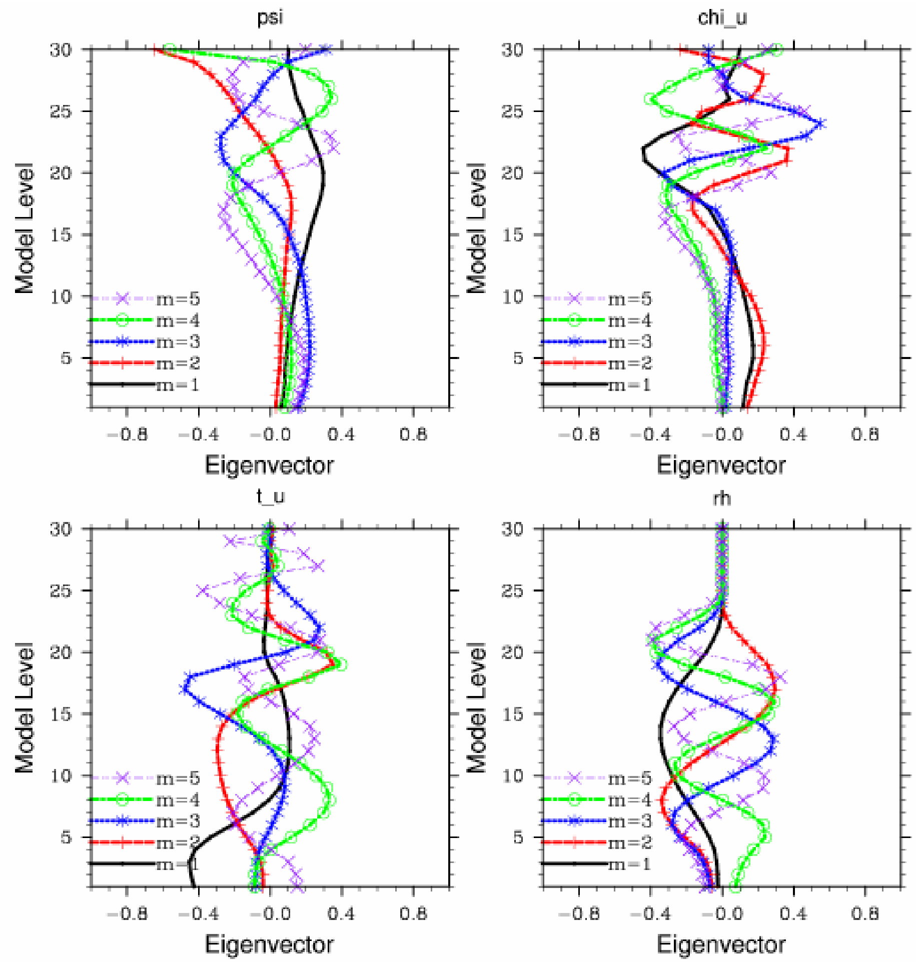


Comparison of Domain-Averaged Eigenvectors E_v^g

CWB 45km WRF:



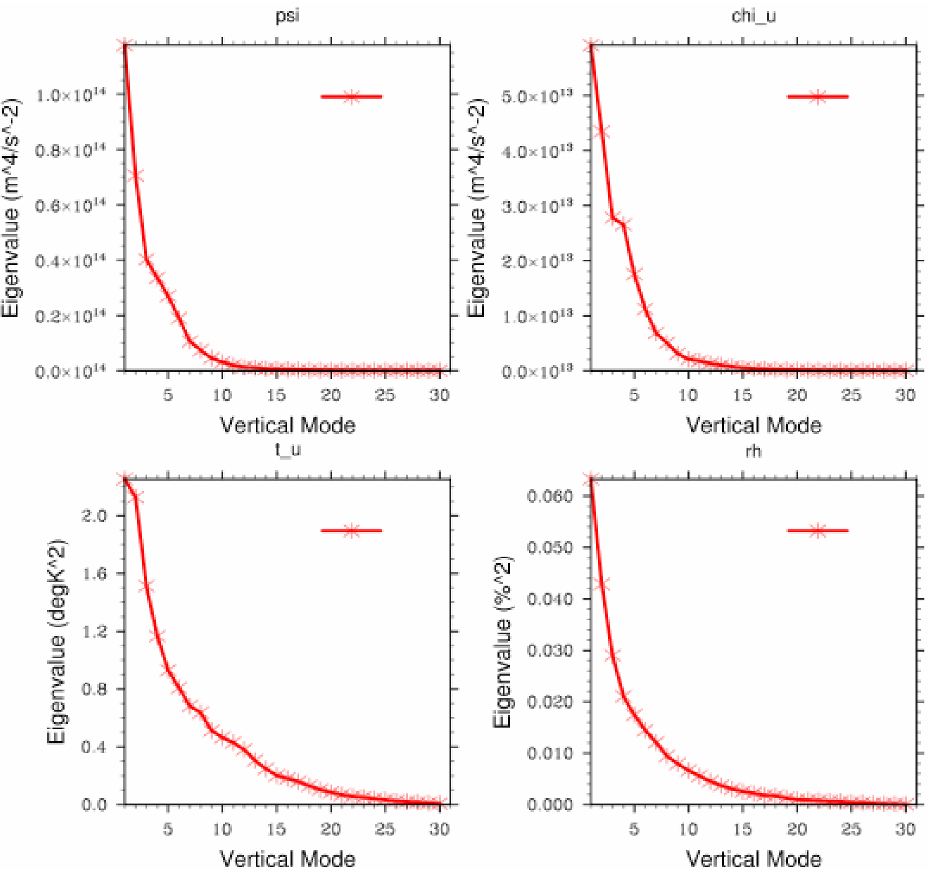
Korean T213 Global:



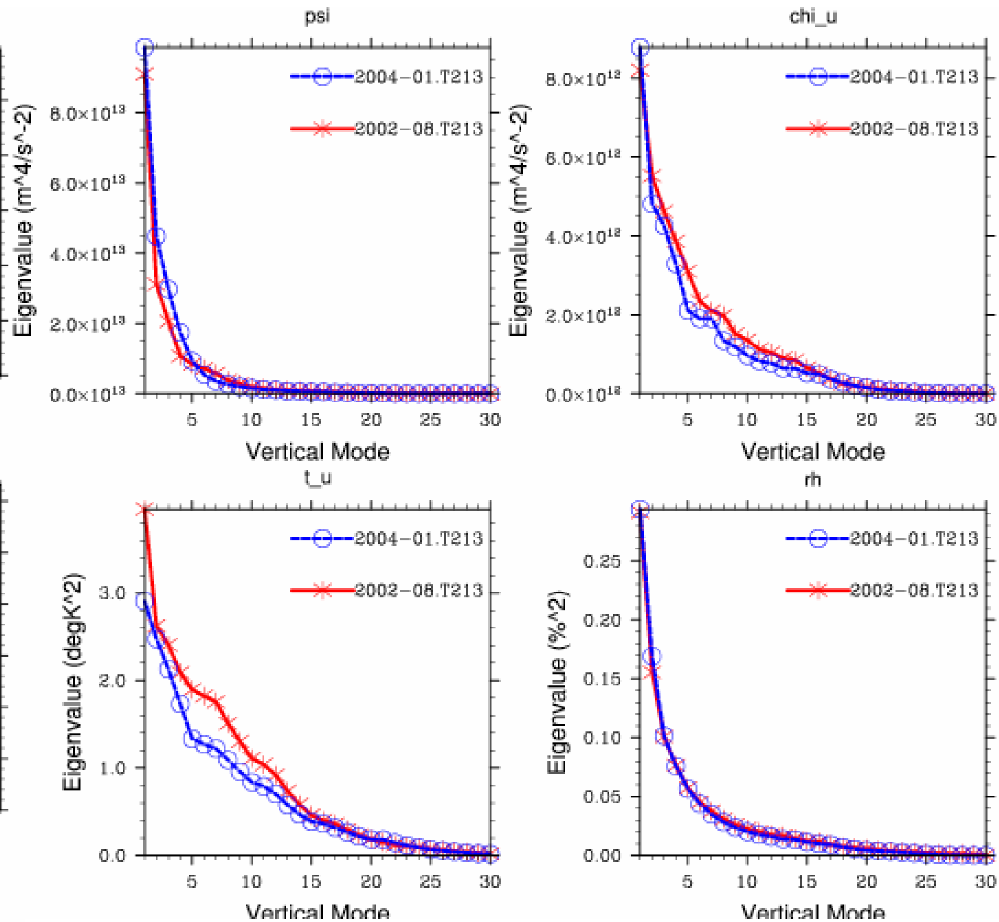
Conclusion: Significant differences in regional/global models (as one might expect), so it is is important to calculate domain-specific background error correlations.

Comparison of Domain-Averaged Eigenvalues Λ_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

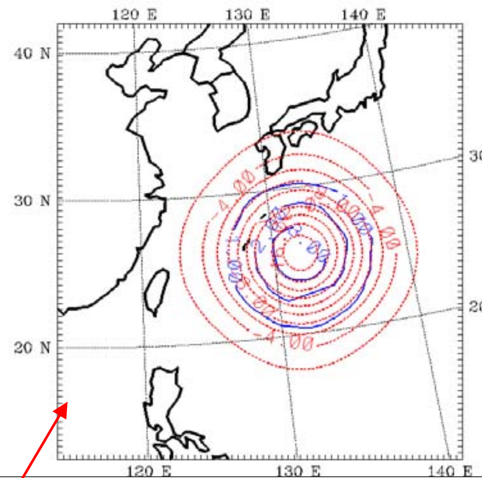
A new utility *gen_be* has been developed at NCAR to calculate BEs. The new *gen_be* code is split into a number of stages:

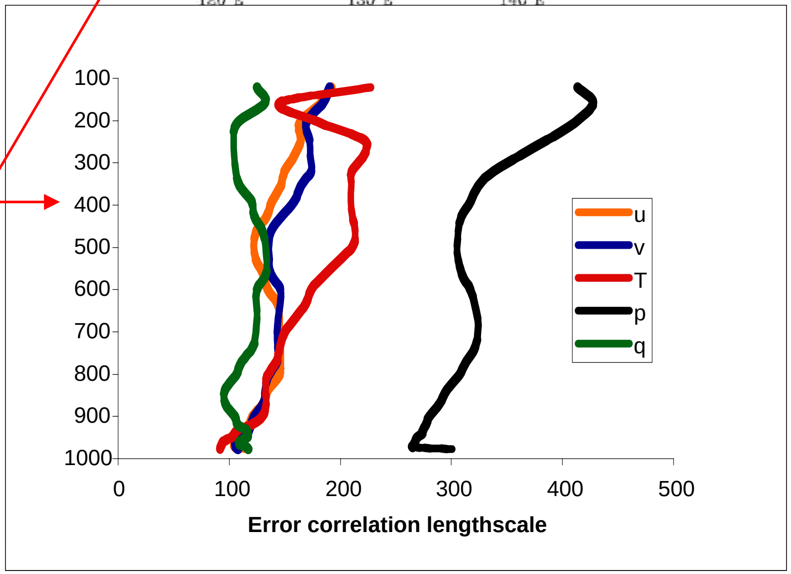
- Stage0: Convert model-specific data to “standard fields”.
- Stage1: Remove time-domain mean from fields.
- Stage2: Calculate regression coefficients, and use them to define “unbalanced” control variables.
- Stage3: Calculate vertical error covariances (eigenvectors and eigenvalues) for control variables.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

WRF-Var Background Error Modeling

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	$\mathbf{v}(i,j,m)$		$\mathbf{v}(l,n,m)$	$\mathbf{v}(i,j,m)$

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





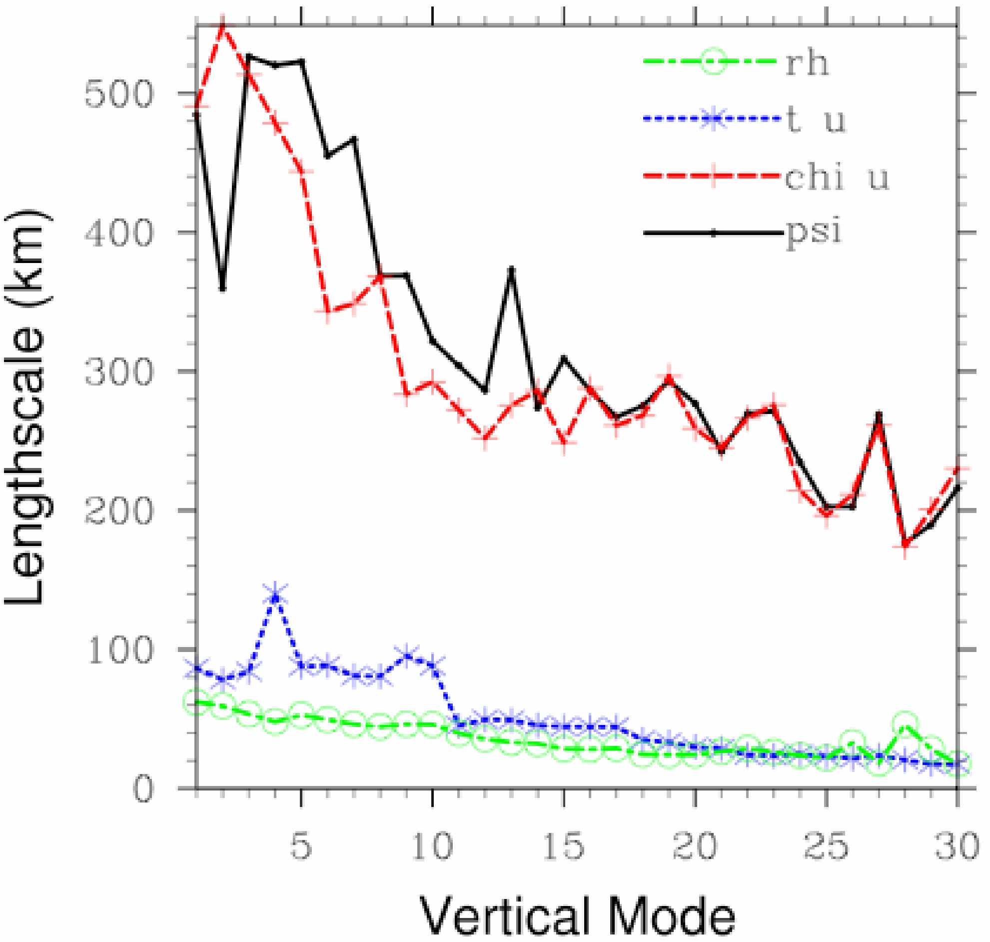
Stage4: Regional model (CWB Sept. 2003) data

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.



Conclusions

- Background Error Statistics are a crucial input to 3/4D-Var systems.
- BE estimation methods are somewhat ad-hoc, and require significant efforts to accumulate the necessary forecast input data (e.g. 1 month 24hr forecasts)
- A new utility *gen_be* has been written to calculate BEs for WRF-Var applications (for a variety of models, and techniques).
- 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.

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 - $\mathbf{P}_{fv}^{g} = \mathbf{E}_{v}^{g}\Lambda_{v}^{g}\mathbf{E}_{v}^{g\,T}$

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

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