Multivariate Correlations: Applying a Dynamic Constraint and Variable Localization in an Ensemble Context

Cathy Thomas[1,2,3], Kayo Ide[1]

Additional thanks to Daryl Kleist, Eugenia Kalnay, Takemasa Miyoshi, Brian Hunt, Jim Carton, Steven Greybush, Fred Kucharski

6th EnKF Workshop
20 May 2014

1 – University of Maryland
2 – NOAA/NWS/NCEP/EMC
3 – IMSG

Background Error Covariance Matrix

- Diagonal of $\mathbf{B}$ and its weight relative to the diagonal of $\mathbf{R}$ determine the magnitude of the analysis increment
- Off diagonal of $\mathbf{B}$ determines the spatial structure of the analysis increment

- 3DVar
  - Constant in time
  - Usually isotropic and homogeneous
  - Estimated prior to the experiment
  - Intervariable correlations represented with dynamic constraint
  - Full rank

- Ensemble
  - Contains flow dependent errors
  - Variable in time and anisotropic
  - Estimated using the ensemble
  - Contains sampling error
  - Rank deficient

Hybrid 4DEnVar

f – fixed
e – ensemble

Lorenc 2003, Buehner 2005, Wang 2008a,b

Extended control variable :

$$\delta\tilde{\mathbf{x}} = \begin{pmatrix} \delta\mathbf{x}^f \\ \boldsymbol{\alpha} \end{pmatrix}$$

Analysis increment $\delta\mathbf{x}$:

$$\delta\mathbf{x}_t = \beta^f \delta\mathbf{x}^f + \beta^e \sum_{m=1}^{M} (\alpha_m \circ (\mathbf{X}_m^e)_t)$$

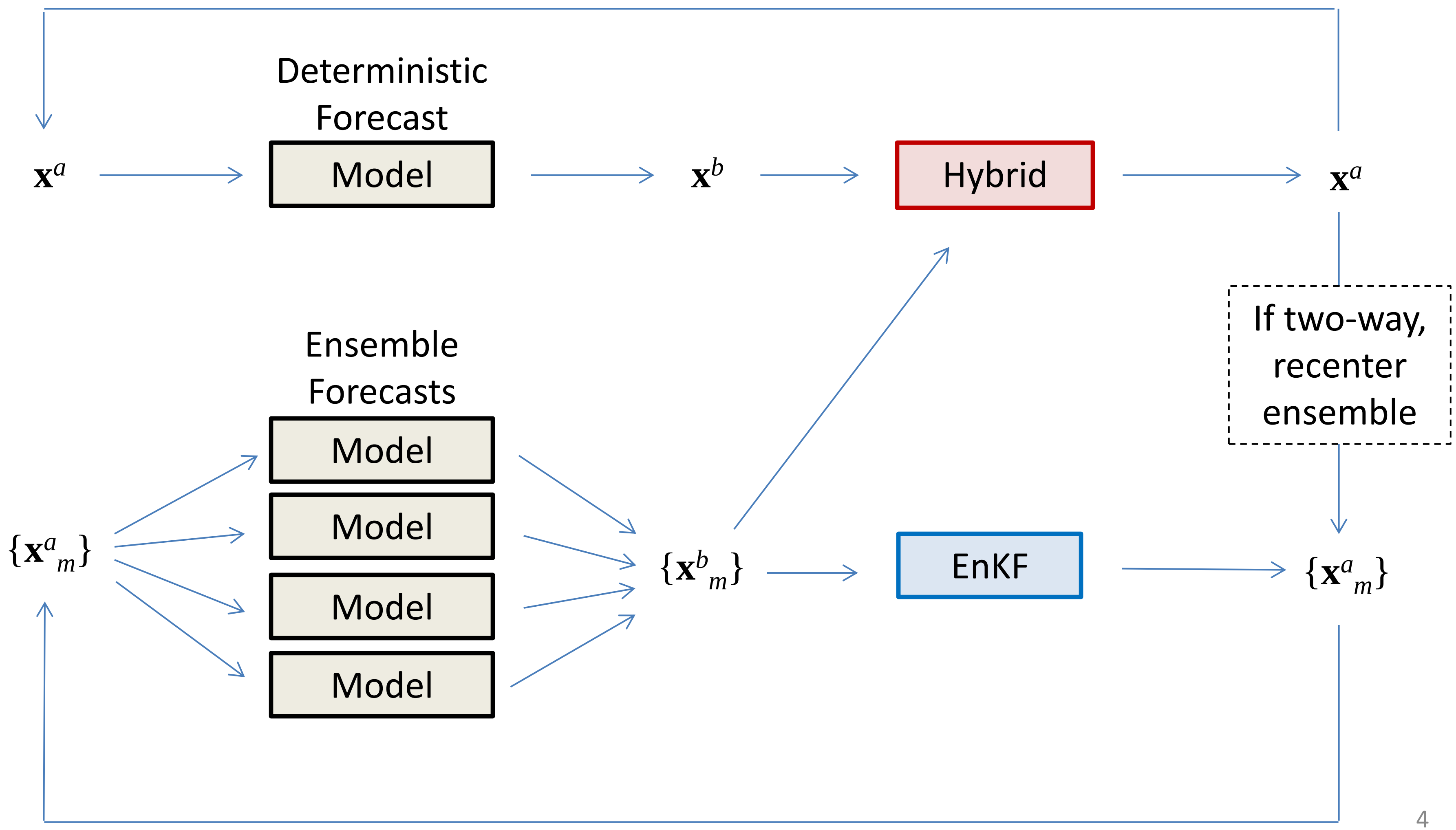
Minimize the cost function:

$$J(\delta\mathbf{x}^f, \boldsymbol{\alpha}) = \frac{1}{2}(\delta\mathbf{x}^f)^T \mathbf{B}^{-1}\delta\mathbf{x}^f + \frac{1}{2}\boldsymbol{\alpha}^T\mathbf{L}^{-1}\boldsymbol{\alpha} + \frac{1}{2}\sum_{t=1}^{\tau}(\mathbf{d}_t - \mathbf{H}_t\delta\mathbf{x}_t)^T\mathbf{R}^{-1}(\mathbf{d}_t - \mathbf{H}_t\delta\mathbf{x}_t)$$

Weights β^f and β^e satisfy: $(\beta^f)^2 + (\beta^e)^2 = 1$

$\beta^f = 1$, $\beta^e = 0$ ← More fixed $\mathbf{B}$ —— More ensemble $\mathbf{P}^b$ → $\beta^f = 0$, $\beta^e = 1$

Hybrid 4DEnVar



Dynamic Constraint, Fixed

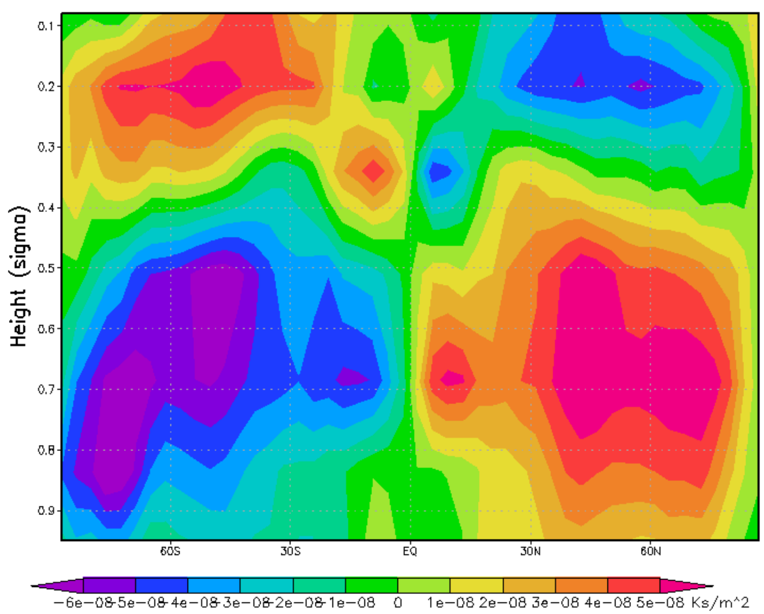
- Wu et al 2002
- Conventionally present in 3DVar
- Represents the physical relationship between the variables
- Magnitude is derived from climatological information

$$\delta\chi = \delta\chi^u + c\delta\psi$$

$$\delta T = \delta T^u + \mathbf{G}\delta\psi$$

$$\delta P = \delta P^u + \mathbf{\Omega}\delta\psi$$

Mid-level $\mathbf{G}$



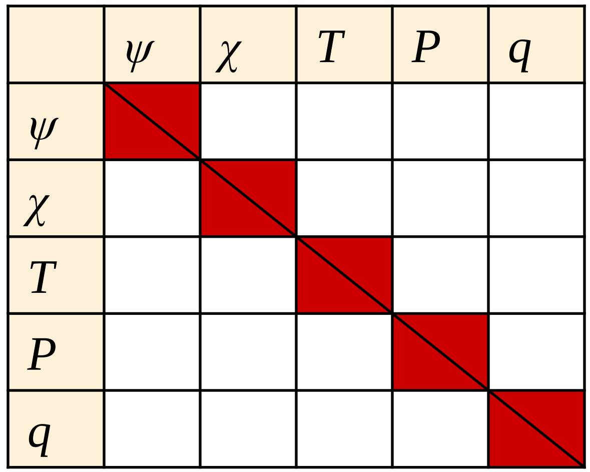
$$\mathbf{x} = \mathbf{\Gamma z}$$

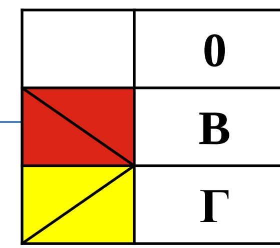
$$\mathbf{x} = \begin{pmatrix} \psi \\ \chi \\ T \\ P \\ q \end{pmatrix} \qquad \mathbf{z} = \begin{pmatrix} \psi \\ \chi^u \\ T^u \\ P^u \\ q \end{pmatrix}$$

$$\mathbf{\Gamma} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ c & 1 & 0 & 0 & 0 \\ \mathbf{G} & 0 & 1 & 0 & 0 \\ \mathbf{\Omega} & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

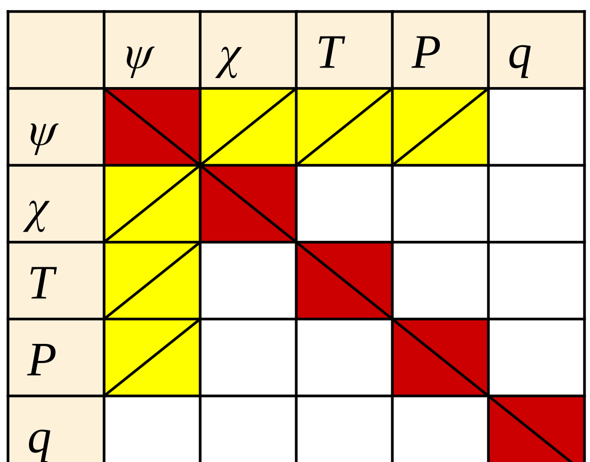
Covariance Matrix, Fixed

No Constraint





Constraint

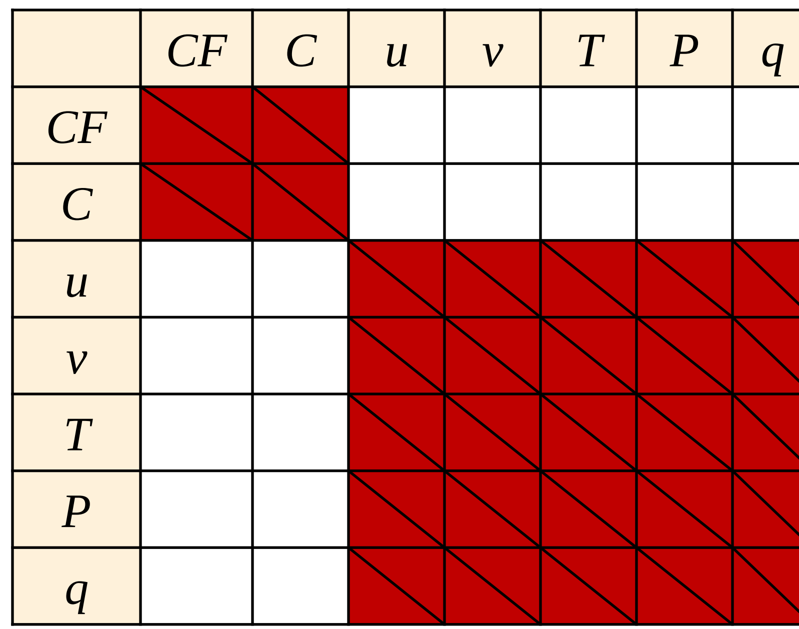


Dynamic Constraint, Ensemble

- Multivariate Correlations
  - Ensembles provide these covariances
    - If the ensemble is large and properly represents reality, these covariances are suitable
    - If the ensemble is too small, the sampling error is large
  - Dynamic constraint can also provides these covariances within the ensemble
    - If ensemble is poor, the dynamic constraint could provide more useful, balanced information
    - Can provide covariances outside of the traditional physical localization radius

Variable Localization (Kang et al 2011)

- If the ensemble is too small, sampling error is large
  - Covariances may exist between variables that should not be correlated
- Implemented in LETKF through observation selection
  - For example, do not use T observations in the calculation of u analysis
- What if one type of observation impacts two variables that we want to be uncorrelated?



Reproduced from
Figure 1 of Kang et al 2011
for the application of
carbon data assimilation

Objectives

- Apply variable localization to the ensemble covariances using a cost function formulation
- Apply the dynamic constraint to the ensemble covariances
- When we combine the dynamic constraint and localization, we get two effects:
  - Localization eliminates spurious correlations
  - Then the constraint propagates the balanced information

Variable Localization

- Specify multiple sets of weights for variable types we wish to be uncorrelated. For this case, variables are split into (ψ) and (χ, T, P, q).

Control Vector:

$$\delta\tilde{\mathbf{x}}_\alpha = \begin{pmatrix} \delta\mathbf{x}^f \\ \boldsymbol{\alpha}_\psi \\ \boldsymbol{\alpha}_\chi \end{pmatrix}$$

Increment:

$$\delta\mathbf{x}_t = \beta^f \delta\mathbf{x}^f + \beta^e \sum_{m=1}^{M} \left(\boldsymbol{\alpha}_{\psi,m} \circ \left(\mathbf{X}^e_{\psi,m} \right)_t \right) + \beta^e \sum_{m=1}^{M} \left(\boldsymbol{\alpha}_{\chi,m} \circ \left(\mathbf{X}^e_{\chi,m} \right)_t \right)$$

Cost Function:

$$J(\delta\tilde{\mathbf{x}}_\alpha) = \frac{1}{2}\left(\delta\mathbf{x}^f\right)^T \mathbf{B}^{-1}\delta\mathbf{x}^f + \frac{1}{2}\left(\boldsymbol{\alpha}_\psi\right)^T \mathbf{L}^{-1}\boldsymbol{\alpha}_\psi + \frac{1}{2}\left(\boldsymbol{\alpha}_\chi\right)^T \mathbf{L}^{-1}\boldsymbol{\alpha}_\chi + \frac{1}{2}\sum_{t=1}^{\tau}(\mathbf{d}_t - \mathbf{H}_t\delta\mathbf{x}_t)^T \mathbf{R}^{-1}(\mathbf{d}_t - \mathbf{H}_t\delta\mathbf{x}_t)$$

Dynamic Constraint, Ensemble

- Apply the dynamic constraint to the ensemble perturbations rather than the extended control variable.

Same Control Variable:

$$\mathbf{z} = \begin{pmatrix} \mathbf{z}^f \\ \mathbf{z}^e \end{pmatrix}$$

Transform the ensemble perturbations:

$$(\mathbf{Z}_m^e)_t = \mathbf{\Gamma}^{-1}(\mathbf{X}_m^e)_t$$

$\mathbf{\Gamma}$ transforms between the full and unbalanced variables:

$$\mathbf{\Gamma} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ c & 1 & 0 & 0 & 0 \\ \mathbf{G} & 0 & 1 & 0 & 0 \\ \mathbf{\Omega} & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

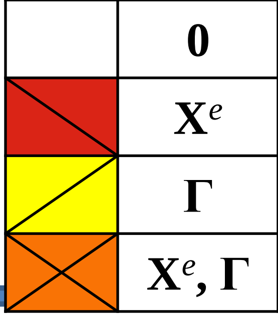
Apply the dynamic constraint to the **whole increment**:

$$\delta\mathbf{x} = \beta^f \mathbf{\Gamma}\mathbf{z}^f + \beta^e \sum_{m=1}^{M} (\alpha_m \circ (\mathbf{X}_m^e)_t) \quad \rightarrow \quad \delta\mathbf{x} = \mathbf{\Gamma}\left(\beta^f \mathbf{z}^f + \beta^e \sum_{m=1}^{M} (\mathbf{z}_m^e \circ (\mathbf{Z}_m^e)_t)\right)$$

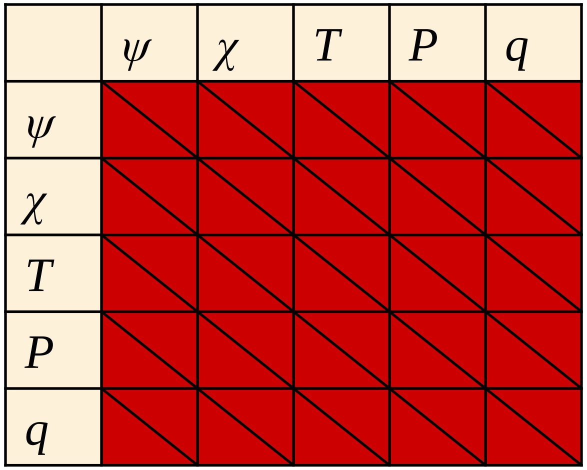
Constraint and Localization

	No Constraint	Constraint
No Localization	- $\mathbf{X}$ perturbations - Keeps $\mathbf{X}_\psi/\mathbf{X}_\chi$ ensemble covariances	- $\mathbf{Z}$ perturbations - Keeps $\mathbf{Z}_\psi/\mathbf{Z}_\chi$ ensemble covariances - Adds $\mathbf{X}_\psi/\mathbf{X}_\chi$ statistical covariances
Localization	- $\mathbf{X}$ perturbations - Removes $\mathbf{X}_\psi/\mathbf{X}_\chi$ ensemble covariances	- $\mathbf{Z}$ perturbations - Removes $\mathbf{Z}_\psi/\mathbf{Z}_\chi$ ensemble covariances - Adds $\mathbf{X}_\psi/\mathbf{X}_\chi$ statistical covariances

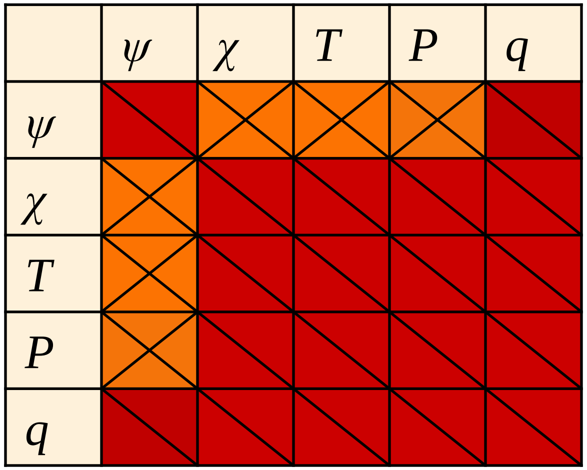
Covariance Matrix, Ensemble



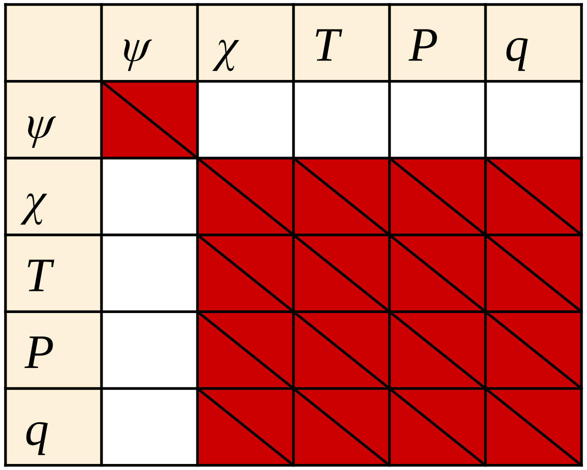
No Constraint, No Localization



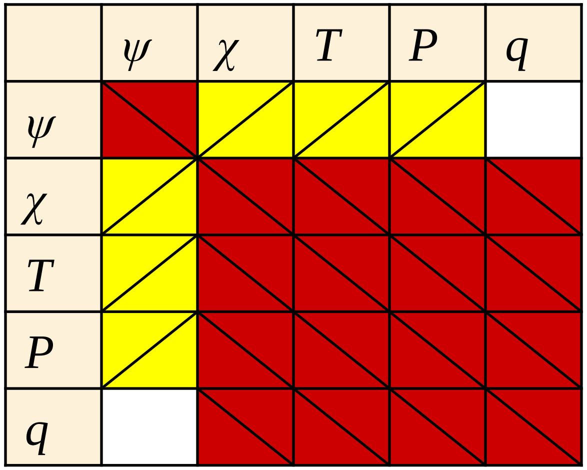
Constraint, No Localization



No Constraint, Localization



Constraint, Localization



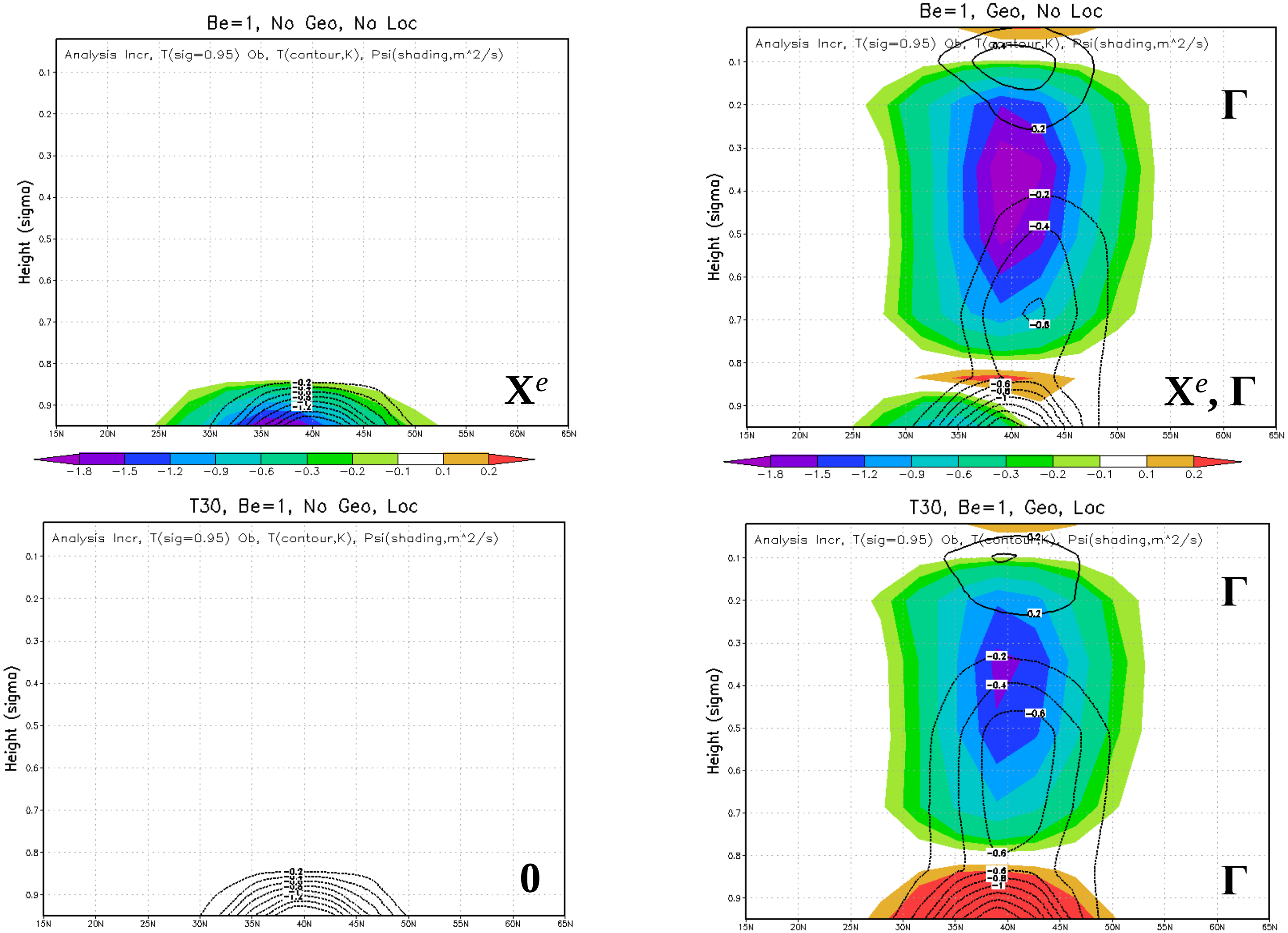
Single T surface Observation, β^e=1

Contours – T
Shading - ψ

No Constraint | Constraint

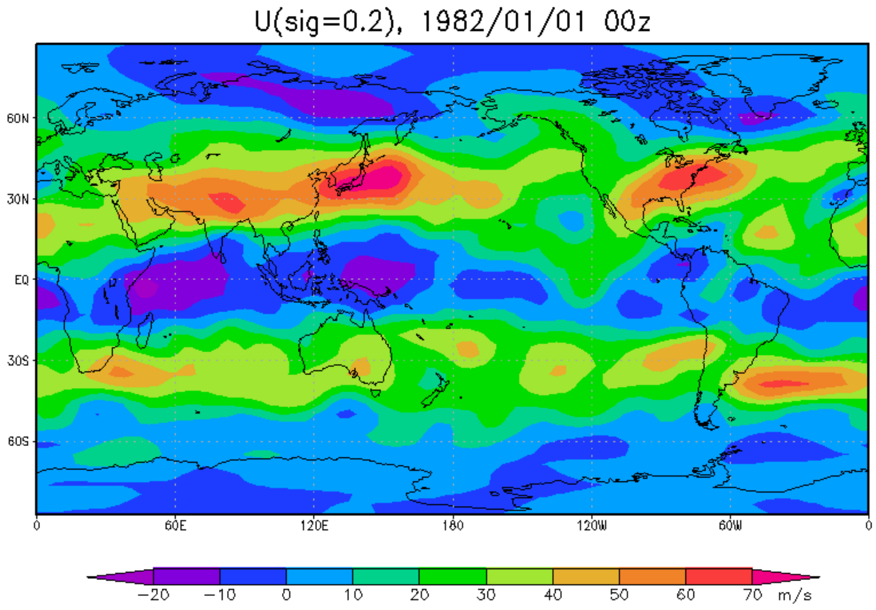
No Localization

Localization



Model Description – SPEEDY

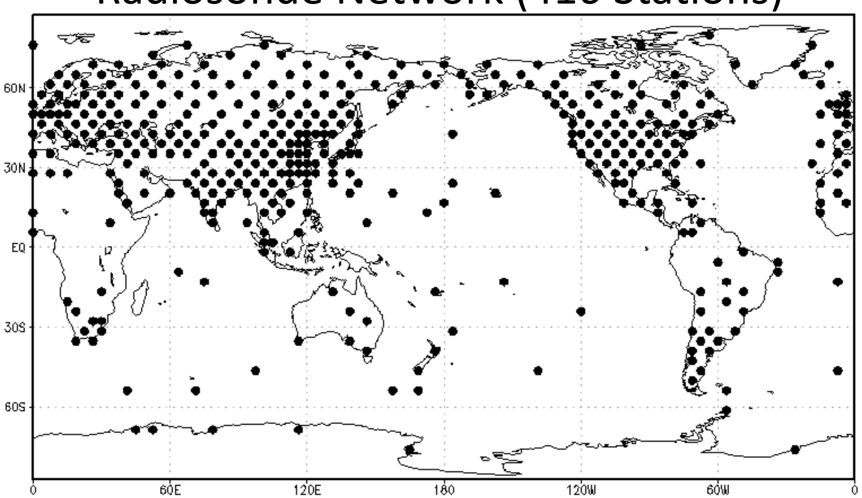
- Molteni 2003
- Model Description
  - **S**implified **P**arameterizations, primitiv**E**-**E**quation **DY**namics
  - Global atmospheric general circulation model of intermediate complexity
- Version 41
  - Provided by Fred Kucharski (ICTP)
  - 3 horizontal resolution options: T30, T47, T63
  - 8 vertical levels
- Output every hour (addition by Miyoshi and Greybush)



Experiment Set-Up

- Resolution: T63 Truth with T30 forecasts and analyses
- Beta weighting: 25% Fixed, 75% Ensemble
- Ensemble Size: 20 members
- Inflation: Fixed at 6%
- Experiment length: 2 years (January 1982 – January 1984)
- Observations: simulated radiosonde network and satellite observations

Radiosonde Network (416 Stations)



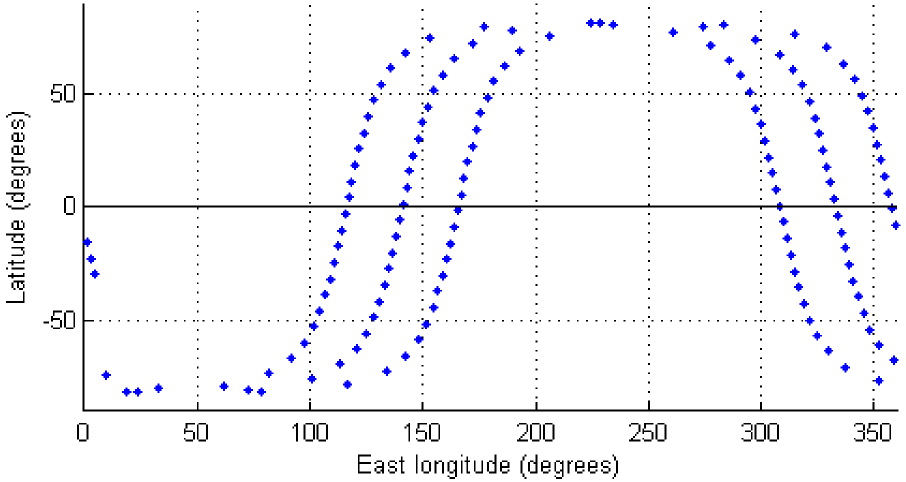
Observation Type	Observation Error
u	1 m/s
v	1 m/s
T	1 K
P	100 Pa
q	10^{-4} kg/kg

Experiment Set-Up

Simulated satellite observations

- AIRS on Aqua and SeaWinds on Quikscat
- 5 minute intervals with linear time interpolation to an hourly T63 truth
- AIRS:
  - T profile: 2 K error
  - q profile up to middle model level: 2×10^{-4} kg/kg error
- SeaWinds:
  - u and v at lowest model level: 1.5 m/s error

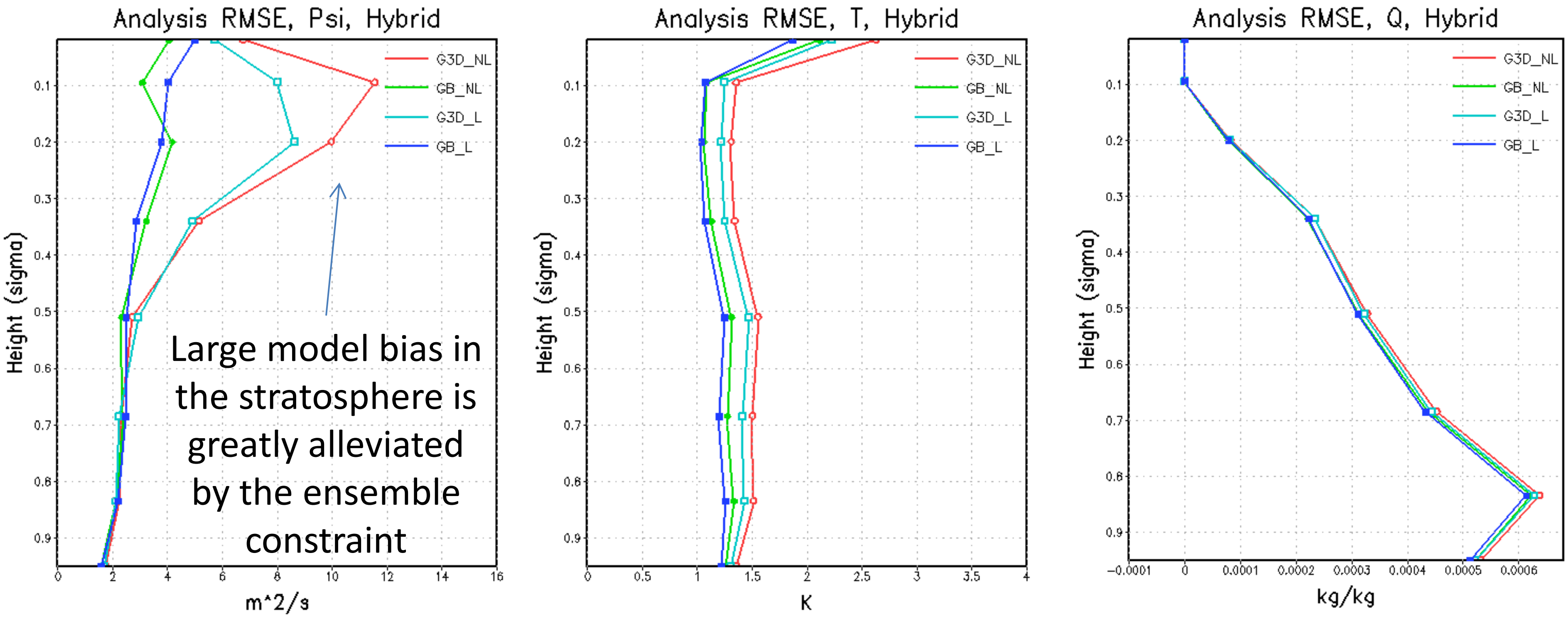
Simulated AIRS Observations for 6 Hours



RMSE – Analysis Skill

No Constr., No Loc.
Constr., No Loc.
No Constr., Loc.
Constr., Loc.

RMSE compared to the nature run

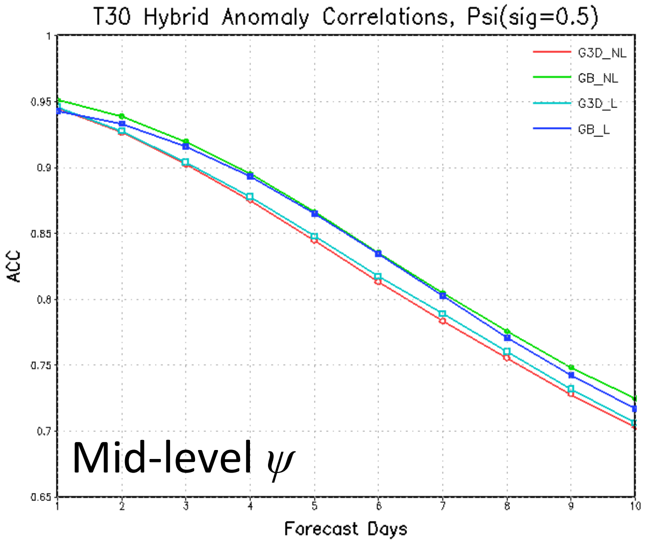


Ensemble constraint has largest impact.

Variable localization with the constraint provides additional value and produces the most accurate analysis in general.

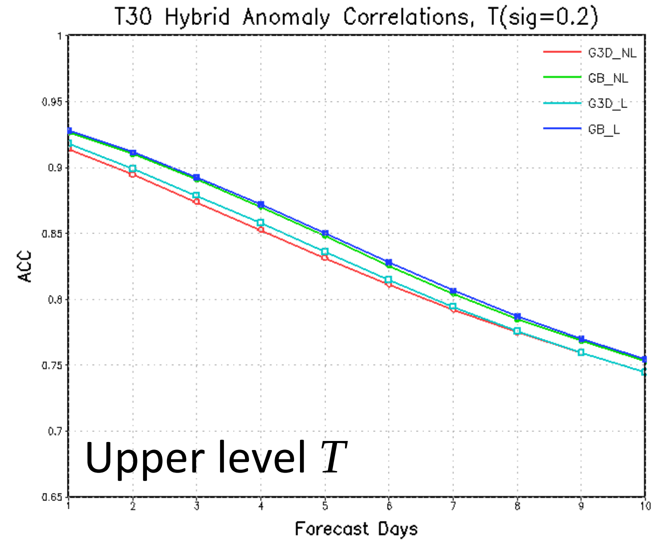
ACC – Forecast Skill

No Constr., No Loc.
Constr., No Loc.
No Constr., Loc.
Constr., Loc.



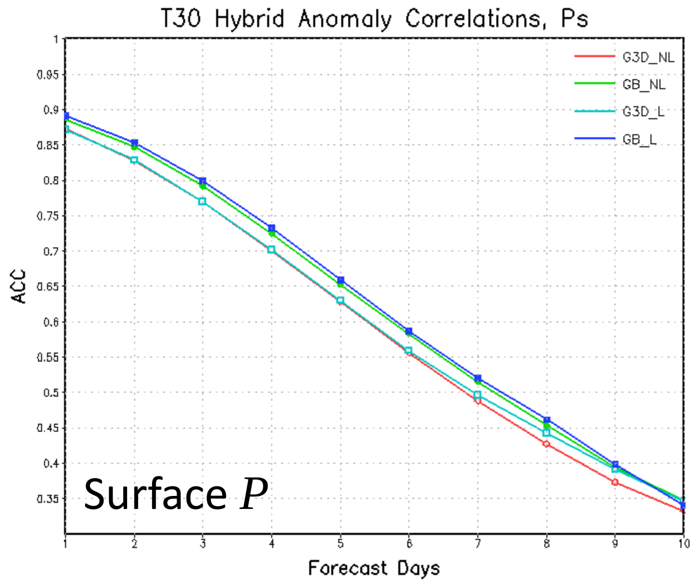
Mid-level ψ

Global anomaly correlation coefficients with respect to the truth.



Upper level T

Ensemble constraint produces largest increase in forecast skill.



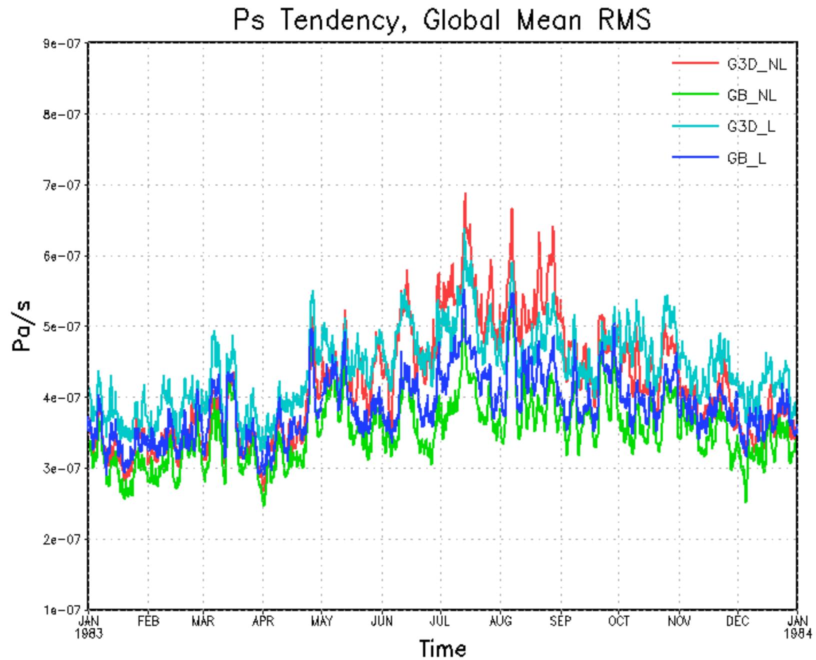
Surface P

Variable localization provides minor increase in forecast skill.

No Constr., No Loc.
Constr., No Loc.
No Constr., Loc.
Constr., Loc.

Measure of Balance

To determine the impact on balance, we examine the surface pressure tendency, which should be reduced with more balanced analyses.



According to this metric:

- Ensemble constraint without variable localization provides the most balanced state
- The case with the ensemble constraint and variable localization also appears well balanced

Summary

- Dynamic constraint
  - Provides additional information about the relationship between variables that the ensemble may overlook or was removed by spatial localization
  - Produces the largest impact in skill
- Variable localization
  - Removes ensemble-provided correlations that may not be correct due to limited ensemble size
  - Provides additional skill, although not as much as the dynamic constraint
- Using both methods simultaneously:
  - Remove unphysical, spurious correlation
  - Add physical, statistically-derived correlations to the ensemble
- Results are promising within the SPEEDY context. We recommend these methods be explored within a state-of-the-art system.

Thank You

References

Buehner, M., 2005: Ensemble-derived stationary and flow-dependent background error covariances: Evaluation in a quasi-operational NWP setting. Quart. J. Roy. Meteor. Soc., 131, 1013-1043.

Buehner, M, P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010a: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministicNWP. Part I: Description and single-observation experiments. Mon. Wea. Rev., 138, 1885–1901.

Buehner, M, P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He 2010b: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observations. Mon. Wea. Rev., 138, 1902–1921.

Etherton, B. J., and C. H. Bishop, 2004: Resilience of hybrid ensemble/3DVAR analysis schemes to model error and ensemble covariance error. Mon. Wea. Rev., 132, 1065–1080.

Hamill, T.M., and C. Snyder, 2000: A hybrid ensemble Kalman filter-3D variational analysis scheme. Monthly Weather Review, 128, 2905-2919.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. Mon. Wea. Rev., 126, 796–811.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman Filter. Physica D, 230, 112–126.

Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc, 1997: Unified notation for data assimilation: Operational, sequential, and variational. J. Meteor. Soc. Japan, 75 (1B), 181-189.

Kalnay et al., 1996: The NCEP/NCAR 40-year reanalysis project. Bull. Amer. Meteor. Soc., 77, 437-470.

Kalnay, E. and S.-C. Yang, 2010: Accelerating the spin-up of Ensemble Kalman Filtering. Q.J.R. Meteorol. Soc., 136: 1644–1651.

Kleist, D.T., 2011: GSI Tutorial. Background and Observation Errors: Estimation and Tuning. 29-30 June 2011.

Kleist, D. T. and K. Ide, 2012: An OSSE-based evaluation of 4D-Ensemble-Var (and hybrid variants) for the NCEP GFS. CMOS 2012 Congress, AMS 21st NWP and 25th WAF Conferences, Montréal, Québec, Canada, < http://www.emc.ncep.noaa.gov/gmb/wd20dk/docs/presentations/WAFNWP-2012/Kleist_hybrid_wafnwp2012.ppt>.

References

Lorenc, A.C., 2003: The potential of the ensemble Kalman filter for NWP—A comparison with 4D-VAR. Quart. J. Roy. Meteor. Soc., 129, 3183-3203.

Lorenz, E., 1996: Predictability - A problem partly solved. Proc. Seminar on Predictability, Reading, United Kingdom, ECMWF, 1-18.

Miyoshi, T., 2005: Ensemble Kalman filter experiments with a primitive-equation global model. Ph.D. dissertation, University of Maryland, College Park, 197 pp.

Miyoshi, T., 2011: The Gaussian Approach to Adaptive Covariance Inflation and Its Implementation with the Local Ensemble Transform Kalman Filter. Mon. Wea. Rev., 139, 1519-1535.

Molteni, F., 2003: Atmospheric simulations using a GCM with simplified physical parameterizations. I: model climatology and variability in multi-decadal experiments. Clim. Dyn., 20, 175-191.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation system. Mon. Wea. Rev., 120, 1747–1763.

Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003a: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. Mon. Wea. Rev., 131, 1524–1535.

Wang, X., C. Snyder, and T. M. Hamill, 2007a: On the theoretical equivalence of differently proposed ensemble/3D-Var hybrid analysis schemes. Mon. Wea. Rev., 135, 222–227.

Wang, X., D. Barker, C. Snyder, and T. M. Hamill, 2008a: A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part I: Observing system simulation experiment. Mon.Wea.Rev., 136, 5116–5131.

Wang, X., D. Barker, C. Snyder, and T. M. Hamill, 2008b: A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part II: Real observation experiments. Mon. Wea. Rev., 136, 5132–5147.

Wang, X., 2010: Incorporating ensemble covariance in the Gridpoint Statistical Interpolation (GSI) variational minimization: a mathematical framework. Mon. Wea. Rev., 138, 2990-2995.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. Mon. Wea. Rev., 130, 1913–1924.

Wu, W.-S., D. F. Parrish, and R. J. Purser, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. Mon. Wea. Rev., 130, 2905–2916.

Back-up Slides

LETKF Variable Localization

Cost function:

$$J = \frac{1}{2}\Delta\mathbf{x}^T\mathbf{B}^{-1}\Delta\mathbf{x} + \frac{1}{2}(\mathbf{d} - \mathbf{H}\Delta\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{d} - \mathbf{H}\Delta\mathbf{x})$$

We split the increment into the stream function increment and the velocity potential increment, removing the correlation between the two.

$$\Delta\mathbf{x} = \begin{pmatrix}\Delta\mathbf{x}_\psi \\ \Delta\mathbf{x}_\chi\end{pmatrix} = \widetilde{\mathbf{X}}^b\mathbf{w} = \begin{pmatrix}\mathbf{X}_\psi & 0 \\ 0 & \mathbf{X}_\chi\end{pmatrix}\begin{pmatrix}\mathbf{w}_\psi \\ \mathbf{w}_\chi\end{pmatrix} = \begin{pmatrix}\mathbf{X}_\psi\mathbf{w}_\psi \\ \mathbf{X}_\chi\mathbf{w}_\chi\end{pmatrix} \qquad \overline{\mathbf{w}} \in \mathbb{R}^{2M}$$

We rewrite the cost function:

$$J = \frac{(M-1)}{2}\left(\mathbf{w}_\psi^T\mathbf{w}_\psi + \mathbf{w}_\chi^T\mathbf{w}_\chi\right) + \frac{1}{2}\left(\mathbf{d} - \mathbf{Y}_\psi\mathbf{w}_\psi - \mathbf{Y}_\chi\mathbf{w}_\chi\right)^T\mathbf{R}^{-1}\left(\mathbf{d} - \mathbf{Y}_\psi\mathbf{w}_\psi - \mathbf{Y}_\chi\mathbf{w}_\chi\right)$$

where

$$\mathbf{Y}_\psi = \mathbf{H}_\psi\widetilde{\mathbf{X}}_\psi \qquad \mathbf{Y}_\psi \in \mathbb{R}^{L\times M}$$

Y() represents the projection from the ensemble onto the observations using just the variables (). For example, for $Y\psi$, the u observation type only contains information from the streamfunction. $Y\chi$ contains the information on that same observation from the velocity potential.

LETKF Variable Localization

The cost function is equivalent to the previous LETKF formulation:

$$J = \frac{(M-1)}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}(\mathbf{d} - \mathbf{Y}\mathbf{w})^T\mathbf{R}^{-1}(\mathbf{d} - \mathbf{Y}\mathbf{w})$$

where:

$$\mathbf{Y} = \begin{pmatrix}\mathbf{Y}_\psi & \mathbf{Y}_\chi\end{pmatrix} \qquad \mathbf{Y} \in \mathbb{R}^{L\times 2M}$$

This cost function is written the same as the original LETKF formulation, except with different definitions for $\mathbf{w}$ and $\mathbf{Y}$. We can similarly solve for $\mathbf{w}$ by setting the cost function gradient to zero.

$$\nabla J = (M-1)\mathbf{w} - \mathbf{Y}^T\mathbf{R}^{-1}(\mathbf{d} - \mathbf{Y}\mathbf{w}) = \mathbf{0}$$

$$\bar{\mathbf{w}} = \big((M-1)\mathbf{I} + \mathbf{Y}^T\mathbf{R}^{-1}\mathbf{Y}\big)^{-1}\mathbf{Y}^T\mathbf{R}^{-1}\mathbf{d}$$

	ψ	χ^u	T^u	q	Ps^u
ψ					
χ^u					
T^u					
q					
Ps^u					