Can we estimate and reduce major sources of forecast uncertainties employing a unified framework?

Dusanka Zupanski
CIRA/Colorado State University
Fort Collins, Colorado

International Workshop on Hydrometeorological and Hydrologic Ensemble Prediction
Boulder, CO, July 19-22, 2005

Collaborators
- ***M. Zupanski, L. Grasso (CIRA/CSU)***
- ***C.D. Kummerow (Atm. Sci. Dept./CSU)***
- ***A. Y. Hou, S. Zhang (NASA/GMAO)***
- ***M. DeMaria (NOAA/NESDIS)***

OUTLINE

- Major sources of forecast uncertainty
- General (probabilistic) framework
- Experimental results
- Conclusions and future work

Research supported by NOAA Grant NA17RJ1228, NASA Grants 621-15-45-78, NAG5-12105, and DoD Grant DAAD19-02-2-0005.

Dusanka Zupanski, CIRA/CSU
Zupanski@CIRA.colostate.edu

Major sources of forecast uncertainty

- Initial conditions
- Model errors (e.g., errors in dynamical equations, errors due to unresolved scales)
- Parametric errors (errors in empirical parameters)
- Forcing errors (e.g., errors in atmospheric forcing in hydrological models)
- Boundary conditions (e.g., lateral boundaries)

These uncertainties are not independent!

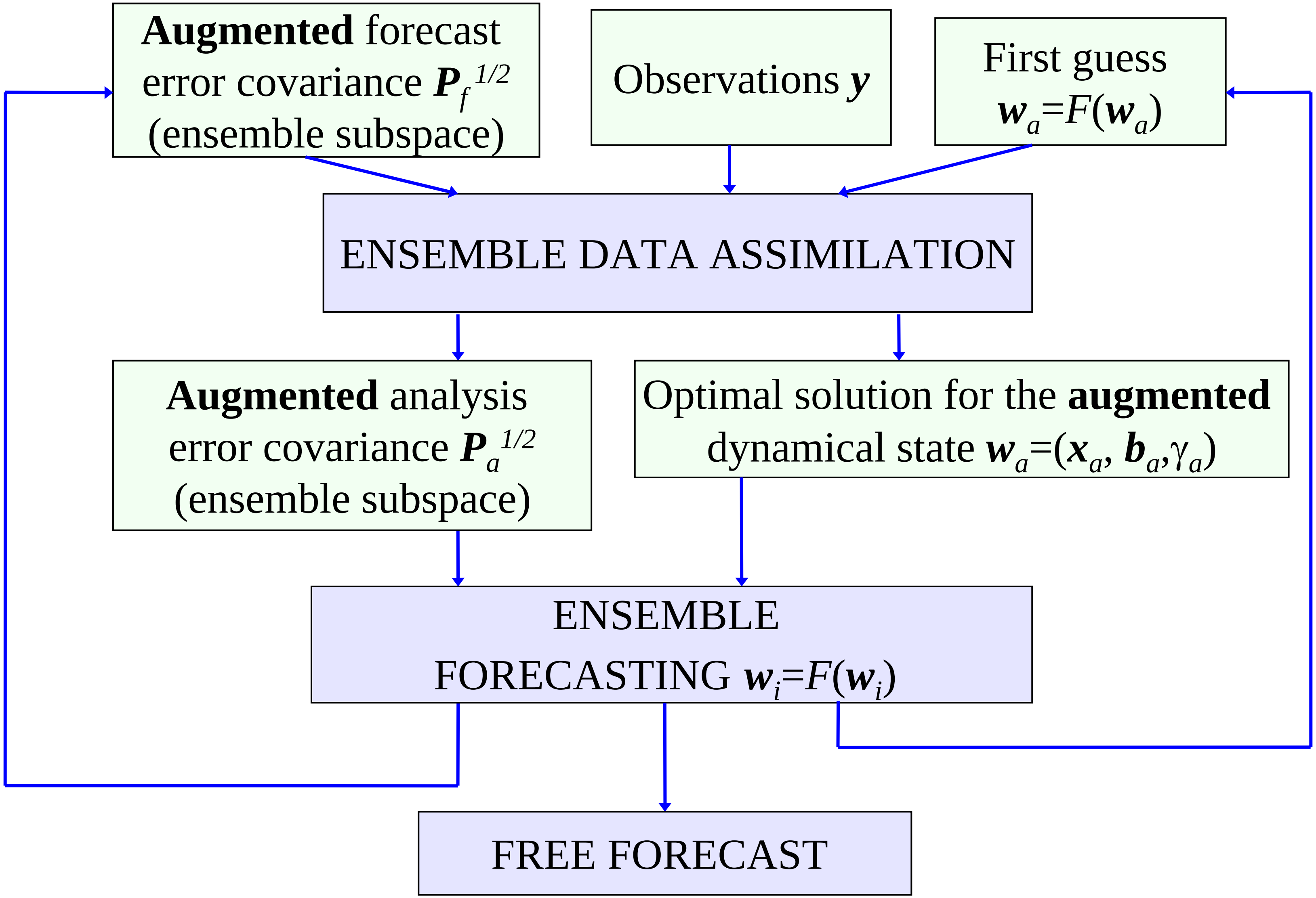
$\Downarrow$

All sources of uncertainty should be taken into account *simultaneously* within a unified mathematical approach.

$\Downarrow$

Verified analysis and forecast uncertainty

Can we do that?



Yes, it can be done employing ensemble data assimilation + state augmentation as a general probabilistic framework!

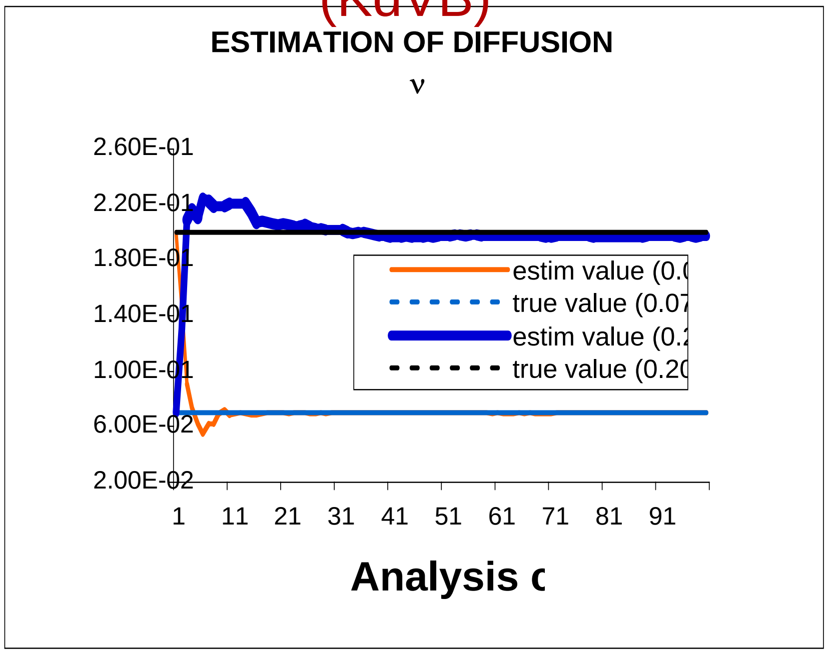
Research in this area

- Mitchell et al. 2000 (MWR)
- Anderson 2001 (MWR)
- Heemink et al. 2001 (MWR)
- Reichle et al. 2002 (MWR)
- Hansen 2002 (MWR)
- Houtekamer et al. 2005 (MWR)
- Hamill and Whitaker 2005 (MWR)
- Zupanski and Zupanski 2005 (MWR)

Dusanka Zupanski, CIRA/CSU
Zupanski@CIRA.colostate.edu

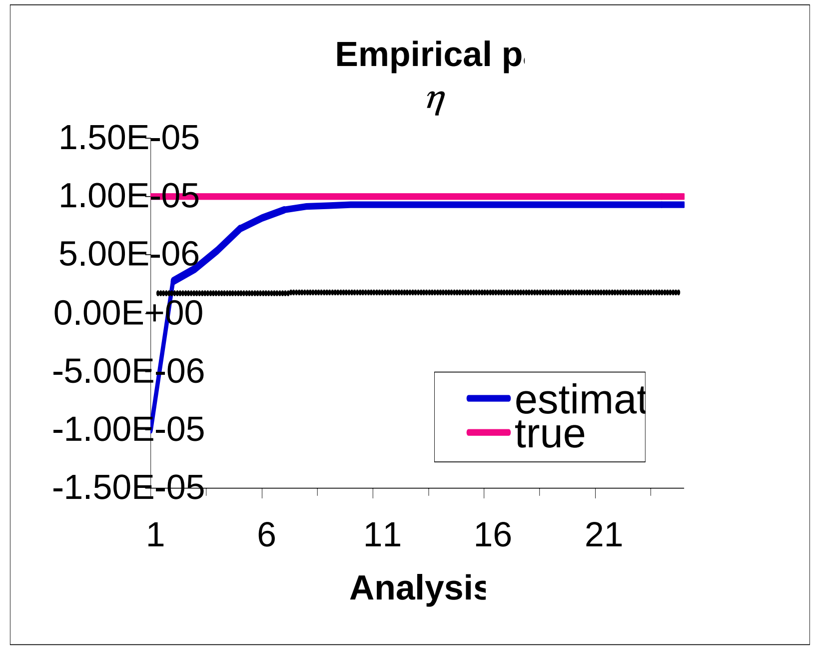
Parameter estimation

Korteweg-de Vries-Burgers model (KdVB)



N_{state} =101+1
N_{ens} =102
N_{obs} =101

RAMS (non-hydrostatic) model



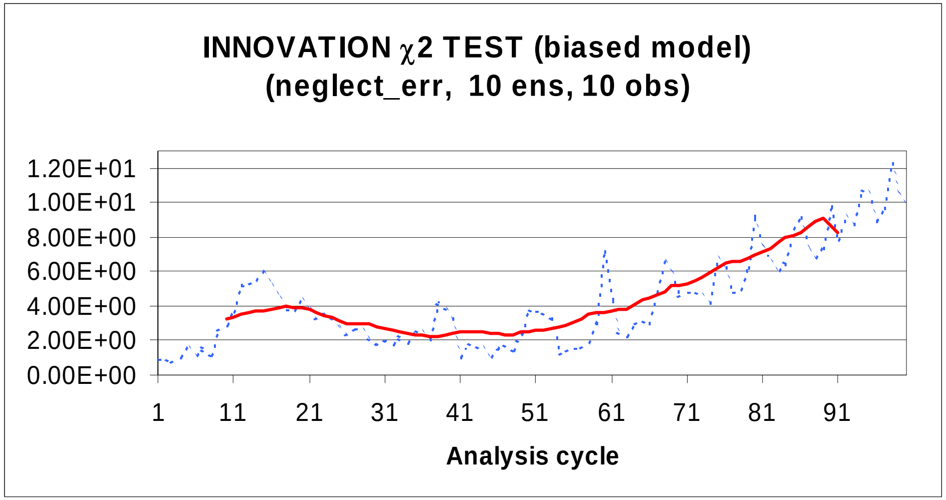
N_{state} =54000+1
N_{ens} =50
N_{obs} =7200

The method is applicable to simple and complex models.

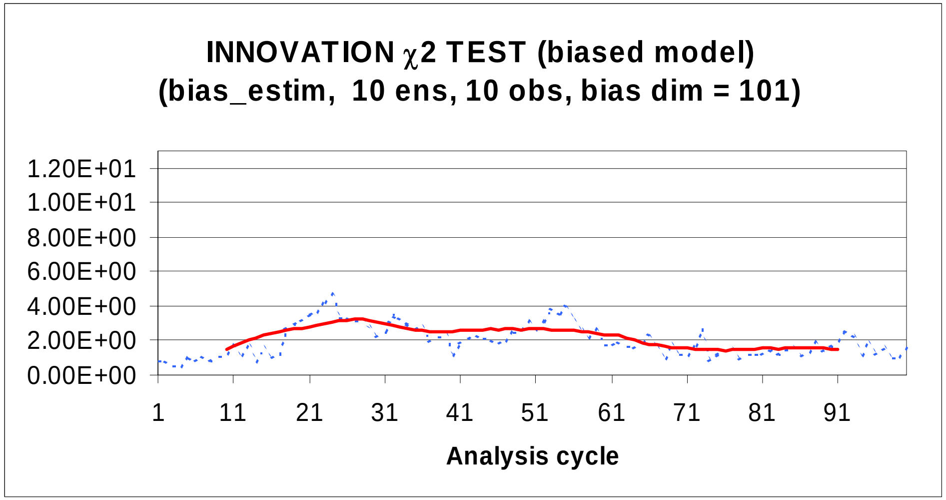
Dusanka Zupanski, CIRA/CSU
Zupanski@CIRA.colostate.edu

BIAS ESTIMATION, KdVB model

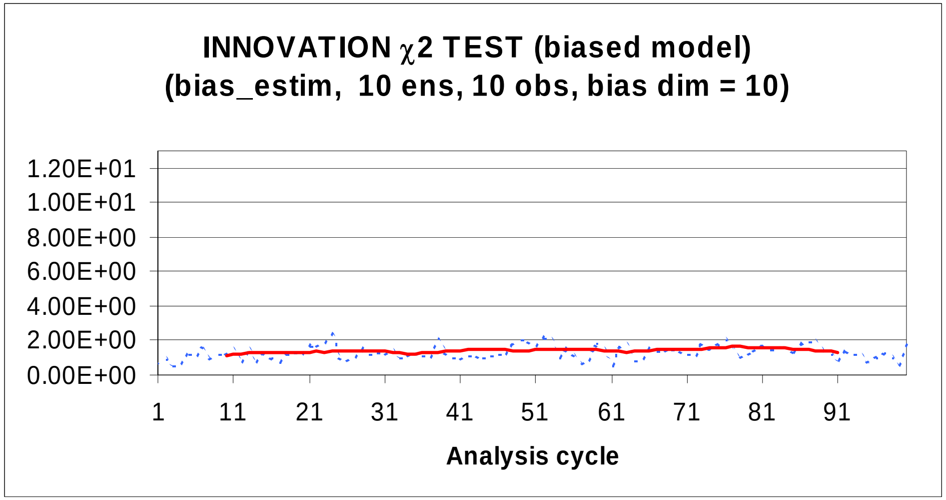
NEGLECT BIAS



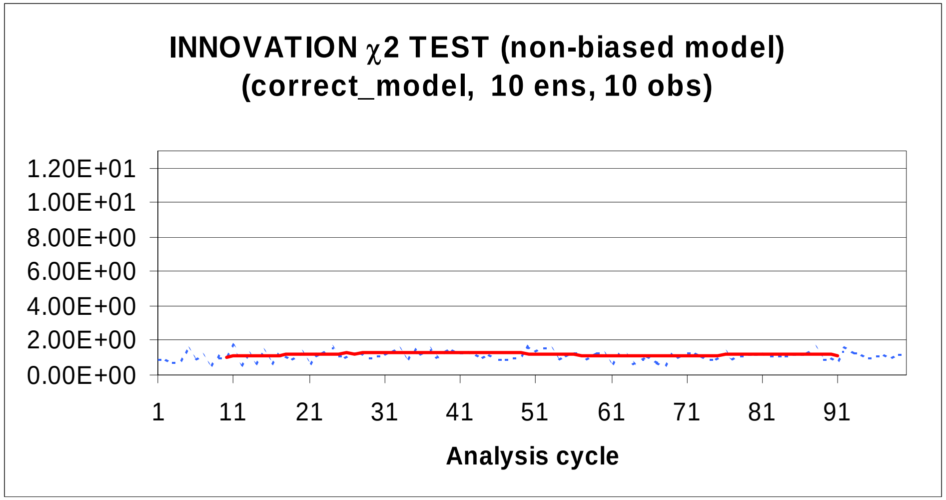
BIAS ESTIMATION (vector size=101)



BIAS ESTIMATION (vector size=10)



NON-BIASED MODEL



It is beneficial to reduce degrees of freedom of the model error.

How do we know the number of DOF?

Answer can be obtained by using the following 3 components within a general framework:

Ensemble Data Assimilation +
State Augmentation +
Information theory

Degrees of freedom (DOF) for signal (Rodgers 2000; Zupanski et al. 2005)

$$d_s = tr[(\boldsymbol{I} + \boldsymbol{C})^{-1}\boldsymbol{C}] = \sum_i \frac{\lambda_i^2}{(1+\lambda_i^2)}$$

C **- information matrix in ensemble subspace**

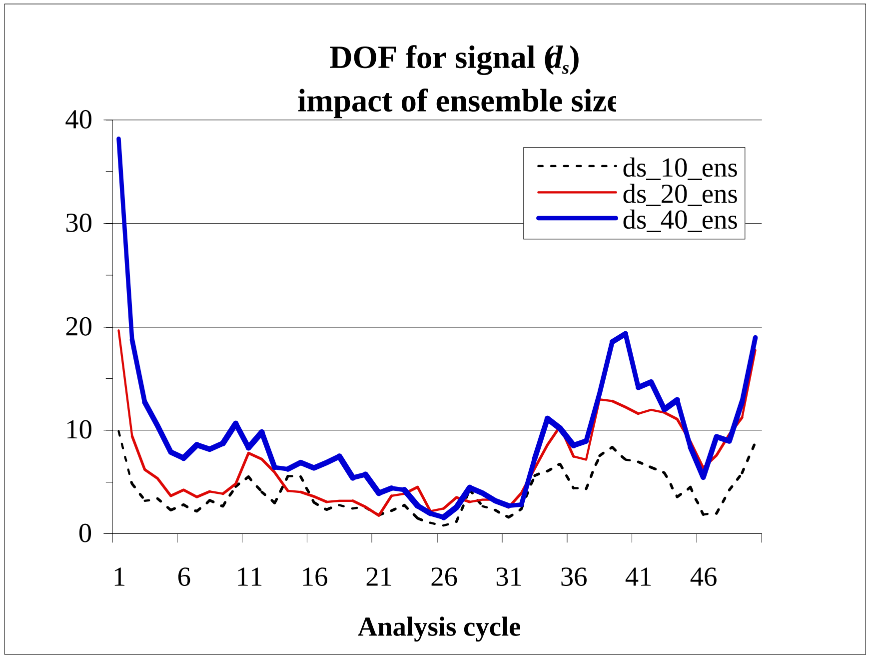
Shannon information content, or entropy reduction

$$h = \frac{1}{2}\sum_i \ln(1+\lambda_i^2)$$

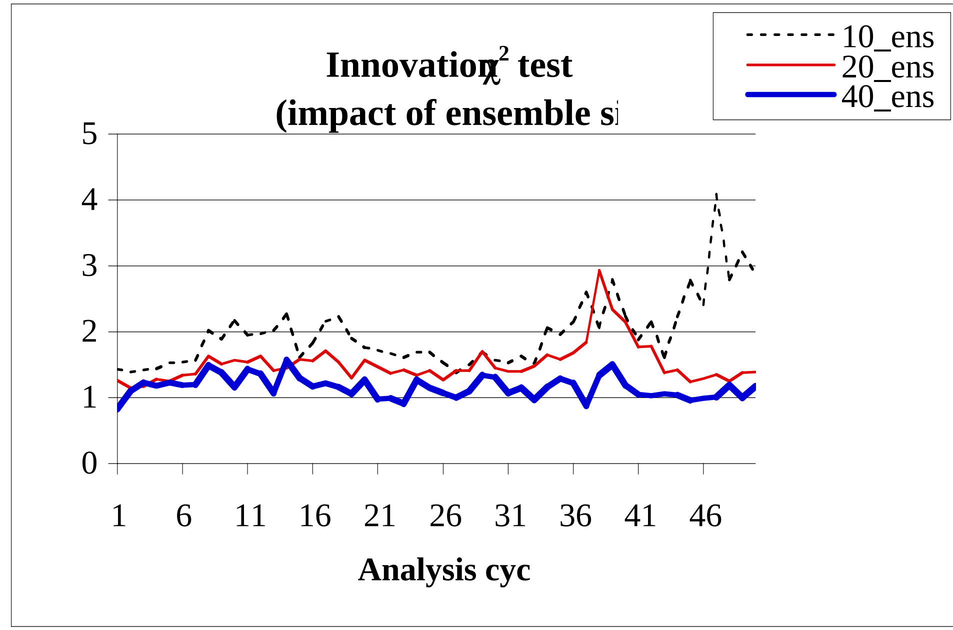
Dusanka Zupanski, CIRA/CSU
Zupanski@CIRA.colostate.edu

Information content analysis
GEOS-5 Single Column Model

DOF for Signal: 10, 20, and 20 Ens.



Chi-square statistics



N_{state} =80; N_{obs} =80

Chi-square test can be used to verify the results of the information content analysis.

Dusanka Zupanski, CIRA/CSU
Zupanski@CIRA.colostate.edu

CONCLUSIONS

➢Unified methodologies for estimating all major sources of analysis and forecast uncertainties and for quantifying DOF of a dynamical system are desirable.

➢Ensemble-based methods can address all these issues within the same mathematical (probabilistic) framework.

➢Hydrological models are good candidates to test the general approach (for estimating multiple empirical parameters, unknown atmospheric forcing, and model bias).

Dusanka Zupanski, CIRA/CSU
Zupanski@CIRA.colostate.edu

FUTURE RESEARCH DIRECTIONS

➢Evaluate this unified approach in application to complex dynamical models and real observations (e.g., atmospheric, hydrological, and carbon transport models).

➢Estimate multiple components of the augmented state variable (e.g., initial conditions, multiple parameters, and model bias).

➢Evaluate the unified approach employing non-Gaussian PDFs.

➢Learn about model errors and possibly improve the models.

Dusanka Zupanski, CIRA/CSU
Zupanski@CIRA.colostate.edu

HOW?

$$\boldsymbol{x}_n = M_{n,n-1}(\boldsymbol{x}_{n-1}, \boldsymbol{b}_{n-1}, \gamma_{n-1})$$

- Dynamical model for standard model state $\boldsymbol{x}$

$$\boldsymbol{b}_n = G_{n,n-1}(\boldsymbol{b}_{n-1})$$

- Dynamical model for model error (bias) $\boldsymbol{b}$

$$\gamma_n = S_{n,n-1}(\gamma_{n-1})$$

- Dynamical model for empirical parameters γ

Define augmented state vector w

$$\boldsymbol{w}_n = (\boldsymbol{x}_{n-1}, \boldsymbol{b}_{n-1}, \gamma_{n-1}) \quad ,$$

And augmented dynamical model F

$$\boldsymbol{w}_n = F_{n,n-1}(\boldsymbol{w}_{n-1}) \quad .$$

$$\Downarrow$$

Find optimal solution (augmented analysis) w_a by minimizing J (MLEF method):

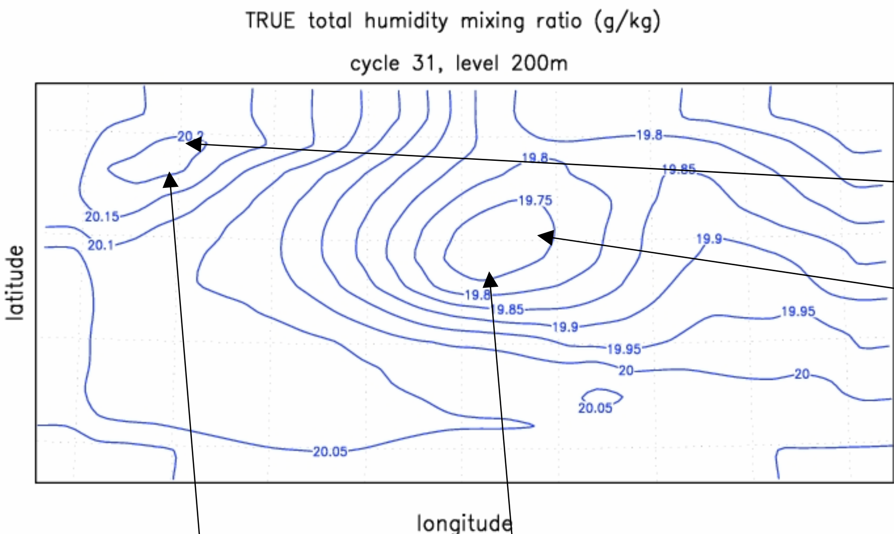
$$J = \frac{1}{2}[\mathbf{w} - \mathbf{w}_b]^T \boldsymbol{P}_f^{-1}[\mathbf{w} - \mathbf{w}_b] + \frac{1}{2}[H[F(\mathbf{w})] - \boldsymbol{y}_{obs}]^T \boldsymbol{R}^{-1}[H[F(\mathbf{w})] - \boldsymbol{y}_{obs}] = min$$

Dusanka Zupanski, CIRA/CSU
Zupanski@CIRA.colostate.edu

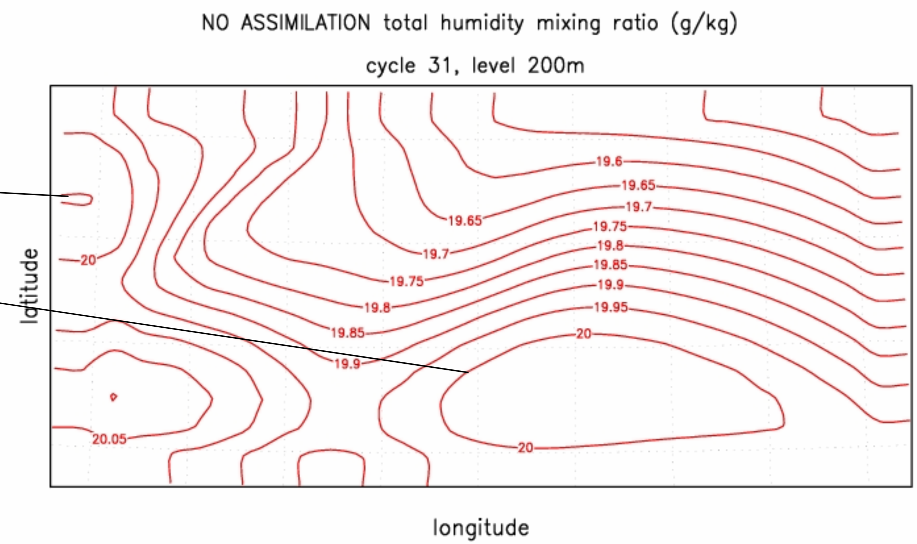
DATA ASSIMILATION application:

CSU-RAMS non-hydrostatic model: Total humidity mixing ratio (level=200m, *Nens*=50, *Nstate*=54000)

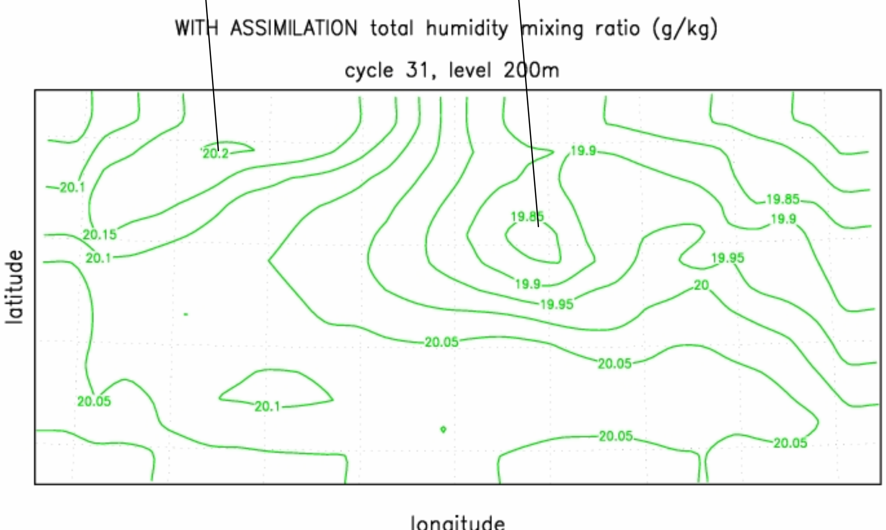
TRUTH



NO ASSIMILATION



ASSIMILATION



Locations of min and max centers are much improved in the experiment with assimilation.

Dusanka Zupanski, CIRA/CSU
Zupanski@CIRA.colostate.edu