JCSDA 2012 Summer Colloquium on Data Assimilation,
Santa Fe, NM, July 24 – August 3, 2012

Ensemble data assimilation of all-sky-radiances in TC inner core

Milija Zupanski and Man Zhang

(contributed by Karina Apodaca, John Knaff, Mark DeMaria, Ming-Jeong Kim)

Cooperative Institute for Research in the Atmosphere
Colorado State University
Fort Collins, Colorado, U. S. A.

Acknowledgements:

- JCSDA

- NOAA NESDIS/GOES-R

- NSF Collaboration in Mathematical Geosciences

- JCSDA S4 computer, NCAR computing (bluefire)

Overview

- All-sky satellite radiance assimilation in TC inner core
- Ensemble data assimilation for hurricanes with NOAA HWRF
- Shannon information measures – observation impact

Motivation for data assimilation in TC inner core

- Develop a robust and efficient data assimilation for high impact weather events

  - tropical cyclones

  - severe weather

- Focus on assimilation of cloud and precipitation affected satellite measurements to gain information about high-resolution TC structure

- Assimilate cloudy radiance from various sources:

  - microwave, infrared, lightning

  - combine information from different sources to find most beneficial combinations

- Utilize operational codes as much as possible, focus on realistic issues

  - WRF NMM, HWRF

  - GSI

  - CRTM

Nonlinear DA uncertainty subspace formulation

1) Initial state and uncertainty: Define an initial state and a subspace (span-vectors)

$$\left[\boldsymbol{x}^0, span\{\boldsymbol{u}_i^0\}\right] \qquad \boldsymbol{x}_i^0 = \boldsymbol{x}^0 + \boldsymbol{u}_i^0; \qquad (i = 1,\ldots,N_E)$$

2) Prediction: Transport the uncertainty span-vectors in time by a prediction model

$$\left.\begin{array}{l} \boldsymbol{x}^t = \mathcal{M}(\boldsymbol{x}^{t-1}) \\ \boldsymbol{x}^t + \boldsymbol{u}_i^t = \mathcal{M}(\boldsymbol{x}^{t-1} + \boldsymbol{u}_i^{t-1}) \end{array}\right\} \left[\boldsymbol{x}^t, span\{\boldsymbol{u}_i^t\}\right] \qquad \boldsymbol{u}_i^t = \mathcal{M}(\boldsymbol{x}^{t-1} + \boldsymbol{u}_i^{t-1}) - \mathcal{M}(\boldsymbol{x}^{t-1})$$

3) Analysis: Maximum a posteriori estimate

$$\max P(X \mid Y) = \min\left[-\ln P(X \mid Y)\right]$$

$$\left.\begin{array}{l} \boldsymbol{y} = \mathcal{K}(\boldsymbol{x}^t) \\ \boldsymbol{y} + \boldsymbol{v}_i = \mathcal{K}(\boldsymbol{x}^t + \boldsymbol{u}_i) \end{array}\right\} \left[\mathcal{K}(\boldsymbol{x}^t), span\{\boldsymbol{v}_i\}\right] \qquad \boldsymbol{v}_i = \mathcal{K}(\boldsymbol{x}^t + \boldsymbol{u}_i) - \mathcal{K}(\boldsymbol{x}^t)$$

Uncertainty

Error covariance localization by model dynamics

- **Nonlinear dynamical systems have a natural capability to localize and scale perturbations**

  - important to maintain that capability in DA

- **Due to severe restriction of ensemble size in high-dimensional application, localization is still needed.**

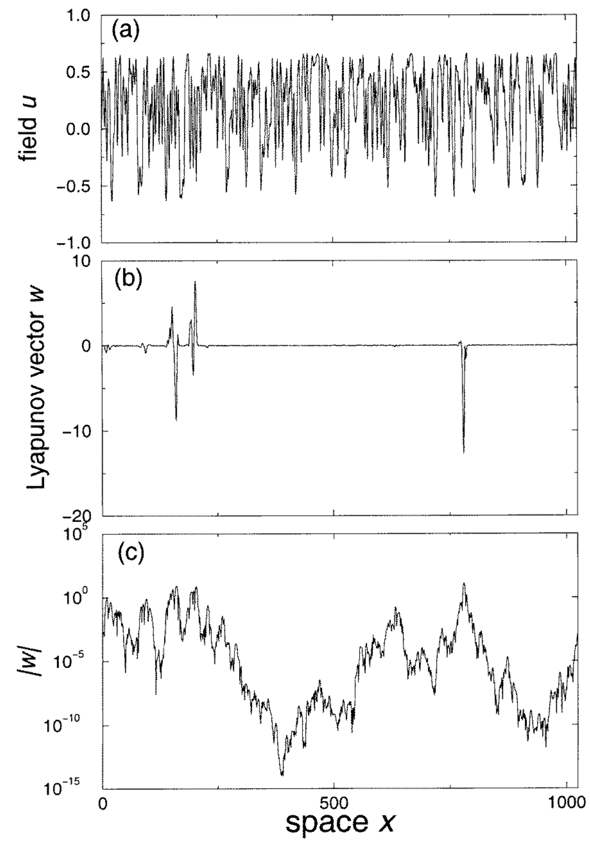


Figure 1. (a) The field u in coupled logistic maps (L = 1024); (b) the instantaneous profile of the corresponding Lyapunov vector at the same time as u; (c) the profile of the Lyapunov vector in the logarithmic scale. (from Pikovsky and Politi 1998)

$$u(t+1,x) = (1-2e)f(u(t,x)) + \varepsilon(f(u(t,x-1)) + f(u(t,x+1)))$$
$$x = 1,\ldots,L$$

Recursive ensemble DA example (KF)

➢ **Assume there is an initial error covariance at** $t=0$ **with columns** a_i (N_E =*ensemble dimension*)

$$A_0 = \begin{bmatrix} a_1 & a_2 & \cdots & a_{N_E} \end{bmatrix}$$

➢ **Denote**

forecast	*analysis*	*scaling*
$F = P_f^{1/2}$	$A = P_a^{1/2}$	$S = \left[I + \left(R^{-1/2} H P_f^{1/2} \right)^T \left(R^{-1/2} H P_f^{1/2} \right) \right]^{-1/2}$

➢ **Recursive formula for reduced-rank DA**

For the *k*-th data assimilation cycle (*k*=1,2,…):

Analysis step (dynamically preferred):

	Analysis approximations	Dynamically preferred
Analysis step:	$A_k^{comp} = F_k^{comp} S_k^{comp}$	$A_k = F_k S_k$
Forecast step :	$F_k = M_k A_{k-1}^{comp}$	$F_k = M_k A_{k-1}$

Recursive ensemble DA example (KF)

Given an initial uncertainty matrix A_0

Cycle 1 $\Rightarrow$

$$F_1 = M_1 A_0$$
$$A_1 = F_1 S_1 = M_1 A_0 S_1$$

. . . . . .

Cycle k $\Rightarrow$

$$F_k = M_k A_{k-1} = M_k M_{k-1} \ldots M_1 A_0 S_1 S_2 \ldots S_{k-1}$$
$$A_k = F_k S_k = M_k M_{k-1} \ldots M_1 A_0 S_1 S_2 \ldots S_k$$

$$M_k M_{k-1} \ldots M_1 A_0 = M\left(t_{k-1}, t_k\right) M\left(t_{k-2}, t_{k-1}\right) \ldots M\left(t_0, t_1\right) A_0 = M\left(t_0, t_k\right) A_0$$

Long forecast

Scaling

Forecast uncertainty after *k* data assimilation cycles:

$$F_k = [M(t_0, t_k) A_0][S_1 S_2 \ldots S_{k-1}] = [MA_0][S]$$

$$MA_0 = M\begin{bmatrix} a_1 & a_2 & \cdots & a_{N_E} \end{bmatrix} = \begin{bmatrix} Ma_1 & Ma_2 & \cdots & Ma_{N_E} \end{bmatrix}$$

The forecast uncertainty is a scaled, long “ensemble” forecast

Ensemble DA relation to Lyapunov exponents

- Lyapunov exponent (λ) measures the separation between two trajectories in phase space

$$|\delta Z_t| \approx e^{\lambda t} |\delta Z_0|$$

$$\lambda_{\max} = \lim_{t \to \infty} \lim_{|\delta Z_0| \to 0} \frac{1}{t} \ln \frac{|\delta Z_t|}{|\delta Z_0|}$$

In our notation: $\delta Z_0 = A_0 = \begin{pmatrix} a_1 & a_2 & \dots & a_K \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$

$$(\delta Z_t)_i = M_t(x_0 + a_i) - M_t(x_0) \qquad \Rightarrow \qquad \delta Z_t \approx M_t A_0$$

- Perturbations directions $(\delta Z_t)_i$ are Lyapunov vectors when $t \to \infty$

- Lyapunov exponents form a so-called Lyapunov spectrum $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$

which is relevant for information dimension of the system (e.g., Kaplan-Yorke dimension)

Ensemble DA that produces uncertainties spanned by Lyapunov vectors can:

- optimally control dominant instabilities in a dynamical system
- have uncertainties localized by model dynamics

(Carrassi et al. 2009)

Ensemble DA algorithm used for assimilation in TC inner core

- Maximum Likelihood Ensemble Filter (MLEF)

  - well-suited for nonlinear analysis problems

  - cost function minimization

  - implicit Hessian preconditioning

  - flow-dependent ensemble error covariance

  - error covariance localization of Yang et al. (2009)

- NOAA Hurricane WRF (HWRF) (27:9 km)
  - Ferrier microphysics (total cloud condensate)

- GSI forward component
  - no adjoint, background, or minimization

- CRTM as forward operator for assimilation of all-sky radiances

- Calculations on JCSDA S4 computer

MLEF IS AN ENSEMBLE DATA ASSIMILATION SYSTEM BASED ON CONTROL THEORY

Forecast:

$$x^f = \mathcal{M}(x^a)$$

$$P_f^{1/2} = \left[p_1^f \quad p_2^f \quad \ldots \quad p_n^f \right] \qquad p_i^f = \mathcal{M}(x^a + p_i^a) - \mathcal{M}(x^a)$$

Change of variable (Hessian preconditioning):

$$x - x^f = P_f^{1/2} \left(I + Z(x^f)^T Z(x^f) \right)^{-1/2} \zeta$$

$$Z(x) = \left[z_1(x) \quad z_2(x) \quad \ldots \quad z_n(x) \right] \qquad z_i(x) = R^{-1/2} \left[\mathcal{K}(x + \delta x) - \mathcal{K}(x) \right]$$

$$\left(I + Z(x^f)^T Z(x^f) \right)^{-1/2} = U \left(I + \Lambda \right)^{-1/2} U^T$$

Analysis (iterative minimization):

$$\zeta_{k+1} = \zeta_k + \alpha_k d_k$$

$$x^a = x^f + P_f^{1/2} \left[I + Z(x^a)^T Z(x^a) \right]^{-1/2} \zeta_{opt}$$

$$P_a^{1/2} = P_f^{1/2} \left[I + Z(x^a)^T Z(x^a) \right]^{-1/2}$$

Hessian preconditioning and cost function

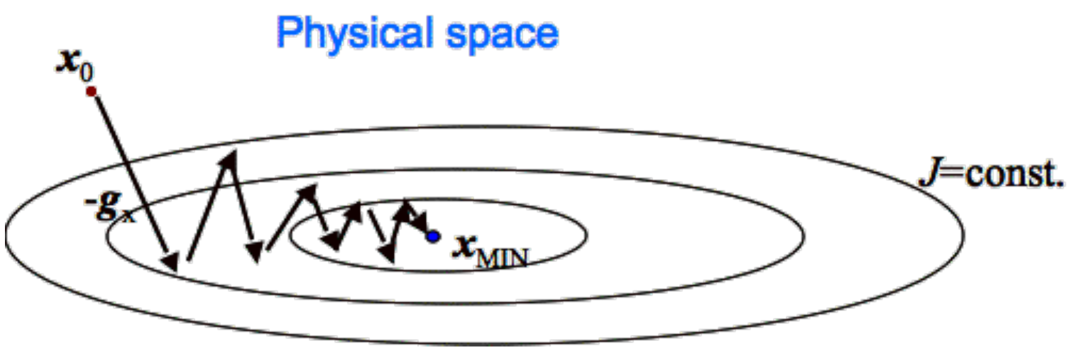
Cost function: $J(x) = \frac{1}{2}\left(x - x^f\right)^T \left[P_f^{-1}\right]_{ens} \left(x - x^f\right) + \frac{1}{2}\left[y - \mathcal{K}(x)\right]^T R^{-1}\left[y - \mathcal{K}(x)\right]$

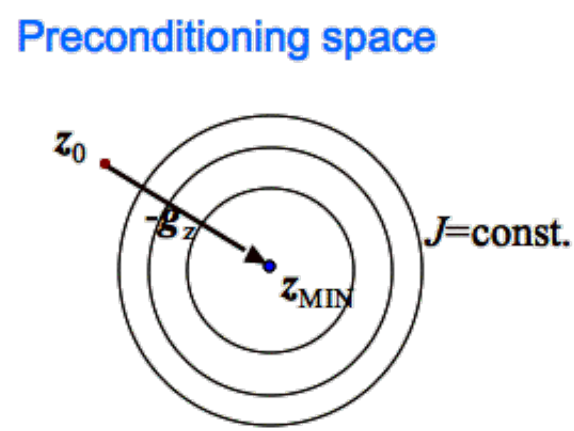
Hessian: $H = \frac{\partial^2 J}{\partial x^2} = P_f^{-1} + K^T R^{-1} K = P_f^{-T/2}(I + Z^T Z) P_f^{-1/2}$ $\qquad H = EE^T$

Optimal change of variable: $\boxed{x - x_f = E^{-T}\zeta = P_f^{1/2}(I + Z^T Z)^{-T/2}\zeta}$

since $\boxed{K_\zeta = E^{-1} K E^{-T} = E^{-1} E E^T E^{-T} = I}$

Geometric interpretation of Hessian preconditioning in MLEF:





Non-differentiable RT observation operators

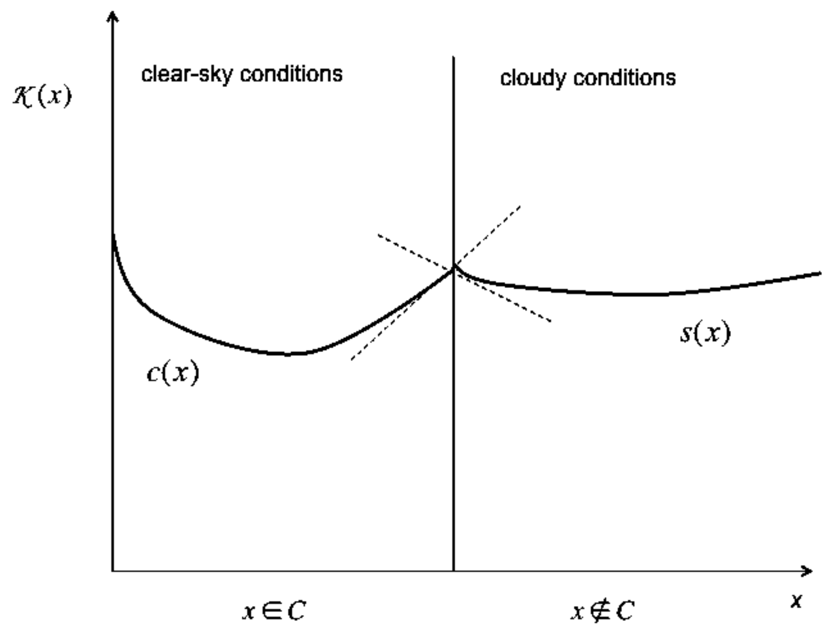
➢ All-sky radiative transfer calculation has two computational branches:

- clear-sky
- cloudy/precipitation-affected

➢ Decision about required calculation depends on model variables, thus creates a discontinuity in gradient and/or cost function

➢ Since commonly used iterative minimization is gradient-based, non-differentiability could have a large impact on the analysis (Steward et al. 2011)

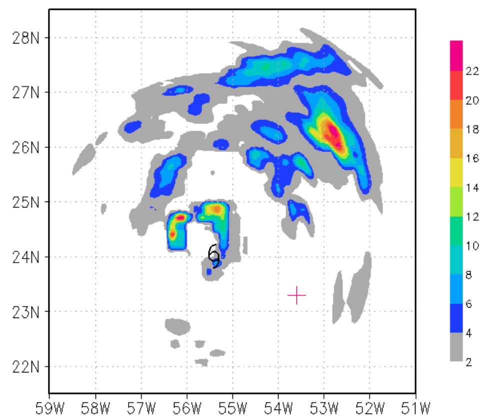
$$\mathcal{K}(x) = \begin{cases} c(x) & x \in C \\ s(x) & x \notin C \end{cases}$$



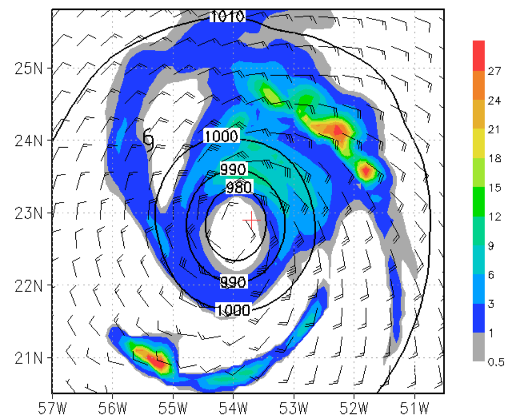
Assimilation of all-sky radiances may benefit from non-differentiable minimization, or other means of addressing discontinuities

Impact of minimization in all-sky MW radiance DA: Hurricane Danielle (2010)

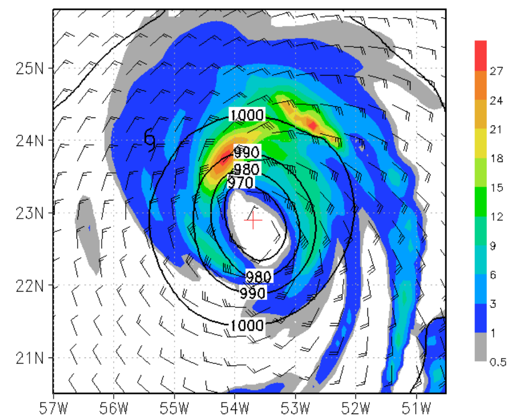
TRMM precipitation radar observations



One minimization iteration



Two minimization iterations



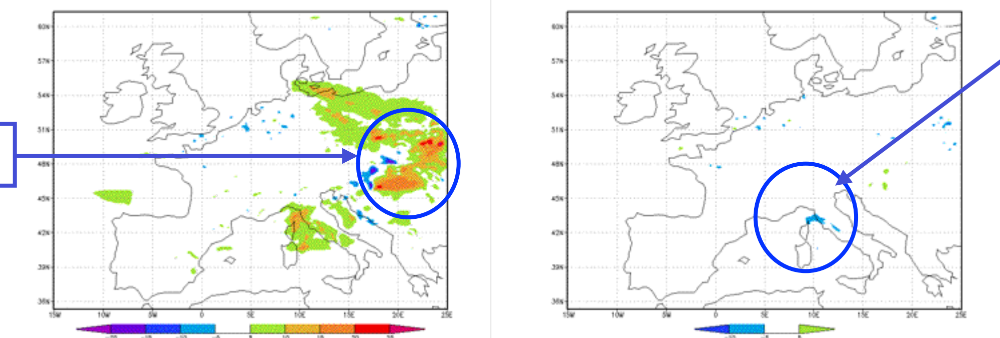
- Assimilation of AMSU-A all-sky radiances with HWRF-MLEF
- TC circulation represented by total cloud condensate (g/kg)
- Solid lines represent the MSLP (hPa)
- The plots are for DA cycle 8 valid 1200 UTC 26 August 2010

Relevance of microphysical control variables

Adjustment of microphysical control variables:

- provides a more complete control of initial conditions
- allows most direct impact of cloud observations on the analysis
- critical for high impact weather (e.g., TC and severe weather)

Microphysics control variables: impact on DA
(wind-storm Kyrill (2007) in Europel)



Temperature analysis increment at 850 hPa

Physically unrealistic analysis adjustment without hydrometeor control variable (cloud ice in this example)

Forecast error covariance: Algebra

Complex inter-variable correlations
(e.g., standard dynamical variables and microphysical variables)

$$P_f = \begin{bmatrix} P_{dd} & P_{dc} \\ P_{dc}^T & P_{cc} \end{bmatrix}$$

Correlations between **dynamical** variables

$$P_{dd} = \begin{bmatrix} P_{T,T} & P_{T,p} & P_{T,v} \\ P_{T,p} & P_{p,p} & P_{p,v} \\ P_{T,v} & P_{p,v} & P_{v,v} \end{bmatrix}$$

Correlations between **microphysical** variables

$$P_{cc} = \begin{bmatrix} P_{ice,ice} & P_{ice,snow} & P_{ice,rain} \\ P_{ice,snow} & P_{snow,snow} & P_{snow,rain} \\ P_{ice,rain} & P_{snow,rain} & P_{rain,rain} \end{bmatrix}$$

Cross-correlations between dynamical and microphysical variables

$$P_{dc} = \begin{bmatrix} P_{T,ice} & P_{T,snow} & P_{T,rain} \\ P_{p,ice} & P_{p,snow} & P_{p,rain} \\ P_{v,ice} & P_{v,snow} & P_{v,rain} \end{bmatrix}$$

Only P_{dd} is well known:

- Correlations among microphysical variables not well known
- Even less known correlations between dynamical and microphysical variables

SIGNIFICANCE OF FORECAST ERROR COVARIANCE

Singular value decomposition (SVD) of forecast error covariance

$$P_f^{1/2} = \sum_i \sigma_i u_i v_i^T$$

Kalman Filter and EnKF solution

$$x - x^f = P_f K^T \left[K P_f K^T + R \right]^{-1} (y - Kx^f) = P_f^{1/2} z_{KF}$$

$$z_{KF} = P_f^{T/2} K^T \left[K P_f K^T + R \right]^{-1} (y - Kx^f)$$

Similar is true for variational DA:

$$x - x^f = P_f^{1/2} z_{VAR}$$

Analysis correction of variational and ensemble DA can be represented as a linear combination of the forecast error covariance singular vectors u_i

$$x^a - x^f = P_f^{1/2} z = \sum_i \sigma_i u_i v_i^T z = \sum_i \gamma_i u_i$$

Structure of forecast error covariance defines analysis correction!

Fundamentally important to have adequate forecast error covariance.

Forecast error covariance structure

Use single observation experiment to assess the structure:

$$x - x^f = P_f K^T \left[K P_f K^T + R \right]^{-1} [y - \mathcal{K}(x^f)]$$

$$K = I\delta_{ij}$$

$$x^a - x^f = P_f \delta_{ij} \left\{ (P_f + R)^{-1} \left[y - \mathcal{K}(x) \right] \right\} = \begin{pmatrix} p_{11} & \cdots & p_{1k} & \cdots & p_{1n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ p_{j1} & \cdots & p_{jk} & \cdots & p_{jn} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nk} & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1_k \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} p_{1k} \\ \vdots \\ p_{jk} \\ \vdots \\ p_{nk} \end{pmatrix}$$

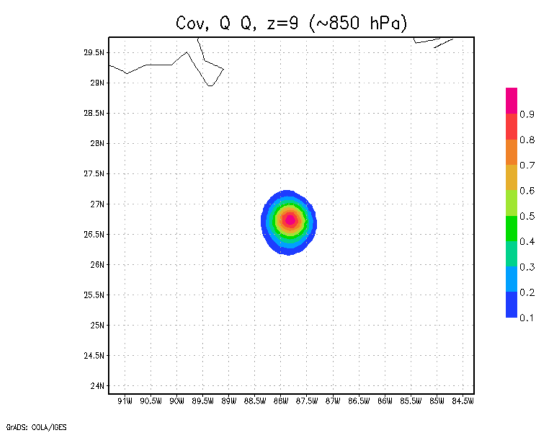
Analysis increment proportional to a column of forecast error covariance

Single observation of specific humidity

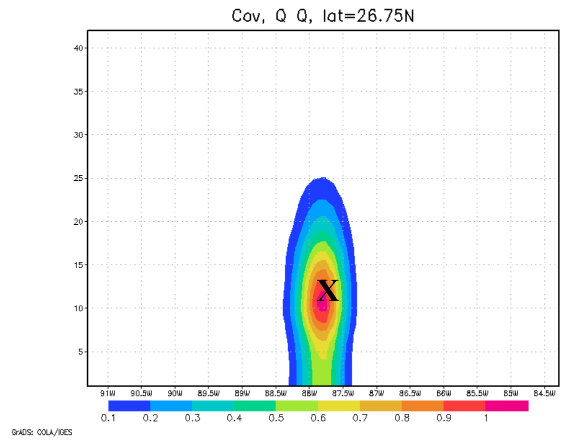
- Hurricane Gustav (2008)
- HWRF-MLEF data assimilation system

HWRF-MLEF analysis response to single specific humidity (Q) observation at 850 hPa

Q analysis at 850 hPa



Vertical cross-section of Q analysis at 26.75 N

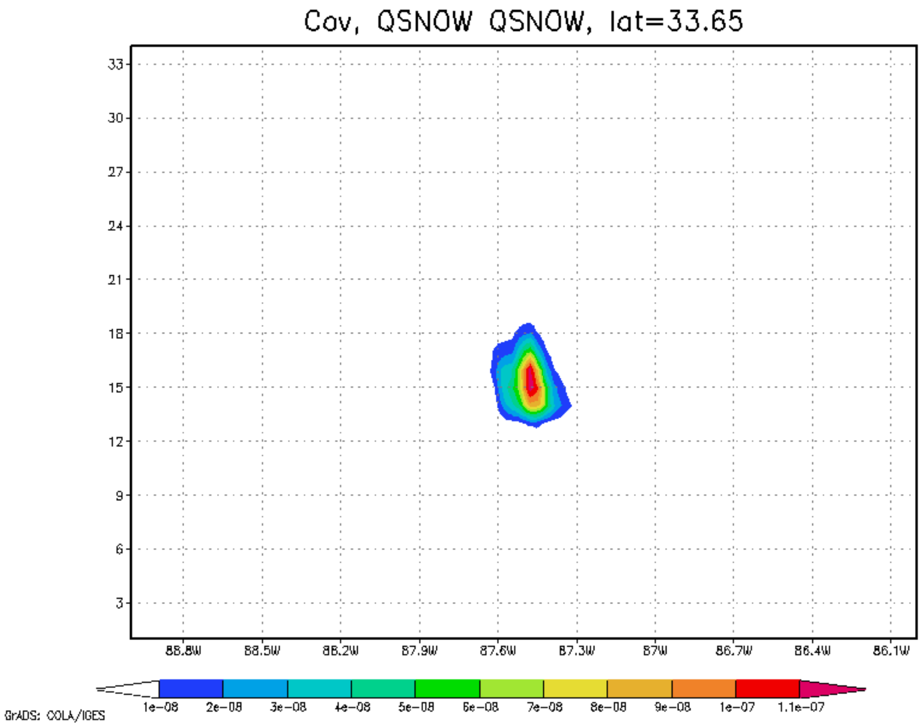


The results are valid for hurricane Gustav (2008) at 1200 UTC on August 31, 2008. The cross denotes the location of observation.

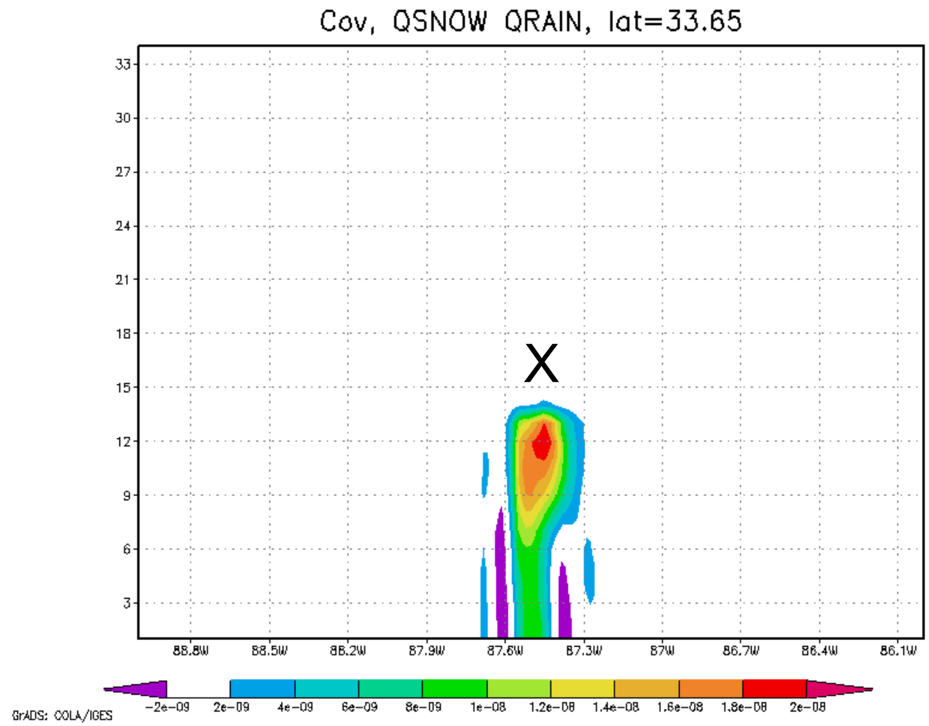
Well-defined, localized analysis response

Single observation of cloud snow at 650 hPa

(a) Snow at 34N ($P_{snow,snow}$)



(b) Rain at 34N ($P_{rain,snow}$)

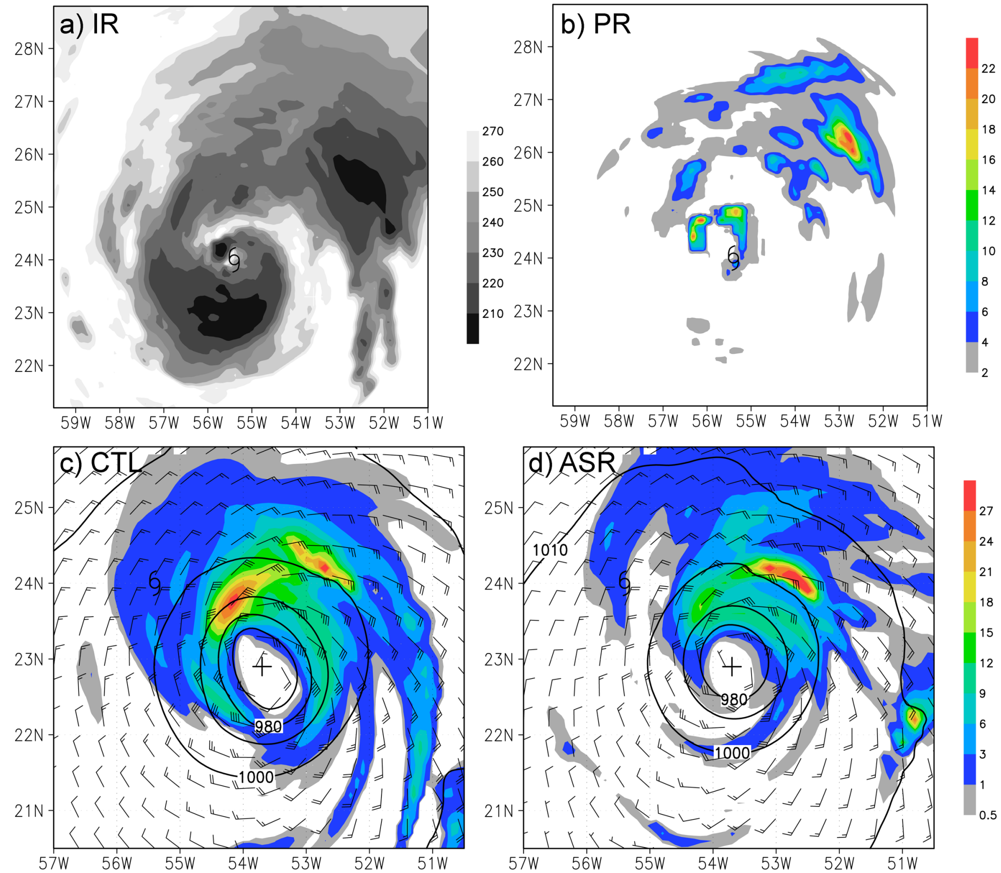


Difficult to model rain-snow correlation: non-centered response and time-dependence

Assimilation of all-sky AMSU-A radiances with HWRF-MLEF

CTL – control experiment

ASR – all-sky experiment



(a) Enhanced Infrared (IR) Imagery at 1145 UTC 26 Aug 2010 (Unit: K); (b) AMSU-retrieved precipitation rate map from MetOp-A at 1311 UTC 26 Aug 2010 (Unit: mm h-1). Distribution of the 6-h total column condensate (Colored; Unit: Kg m-2) forecasts start from cycle 7 analyses of (c) the CTL experiment, and (d) the ASR experiment, superposed with mean sea-level pressure and 10-m above ground wind barbs from, valid at 1200 UTC 26 Aug 2010. (Figure 7 from Zhang et al. 2012)

Shannon Information measures for all-sky radiances (and other observations)

Use information theory (e.g. entropy) as an objective, pdf-based quantification of information (Rodgers 2000; Zupanski et al. 2007):

- Change of entropy
- Relative entropy
- Mutual information

Entropy

$$H\{X\} = -\int p(x)\log(p(x))\,dx$$

Change of entropy due to observations

$$\Delta H = H\{X\} - H\{X \mid Y\}$$

Relative entropy

$$R\{X \mid Y, X\} = -\int p(x \mid y)\log\frac{p(x \mid y)}{p(x)}\,dx$$

Mutual information

$$I\{X;Y\} = -\int p(x,y)\log\frac{p(x,y)}{p(x)p(y)}\,dx$$

How applicable are these measures to realistic data assimilation?

Information measures for Gaussian pdfs

- Gaussian pdfs greatly reduce the complexity, since entropy is related to covariance
- Applicable to commonly used data assimilation algorithms

Change of entropy / degrees of freedom for signal (DFS)

$$\Delta H = d_s = trace\left[I - P_a P_f^{-1}\right]$$

In realistic ensemble methods d_s can be computed exactly in ensemble subspace

$$P_a = P_f^{1/2}\left(I + Z^T Z\right)^{-1} P_f^{T/2} \qquad Z = R^{-1/2} H P_f^{1/2} \qquad Z^T Z = U \Lambda U^T$$

$$d_s = trace\left[(I + Z^T Z)^{-1} Z^T Z\right] \qquad d_s = \sum_i \frac{\lambda_i^2}{1+\lambda_i^2}$$

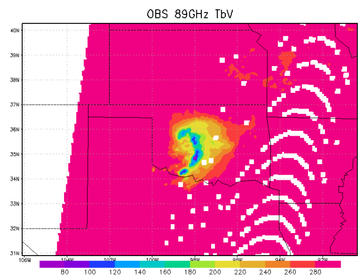
Since eigenvalues of the matrix $Z^T Z$ are known and the matrix inversion is defined in ensemble space, the flow-dependent d_s can be computed

ALL-SKY RADIANCE OBSERVATION INFORMATION CONTENT (DEGREES OF FREEDOM FOR SIGNAL – DFS)

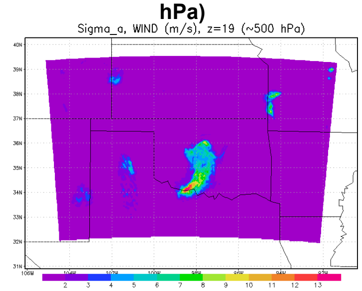
MW: AMSR-E all-sky radiance data assimilation (Erin, 2007)

(from Zupanski et al. 2011, *J. Hydrometeorology*)

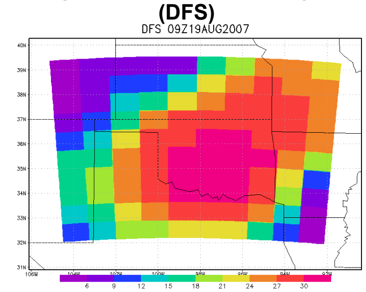
OBS 89v GHz Tb



Wind analysis uncertainty (500 hPa)



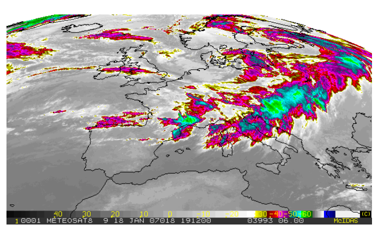
Degrees of Freedom for Signal (DFS)



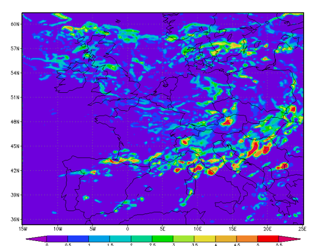
IR: Assimilation of synthetic GOES-R ABI (10.35 mm) all-sky radiances (Kyrill, 2007)

(from Zupanski et al. 2011, *Int. J. Remote Sensing*)

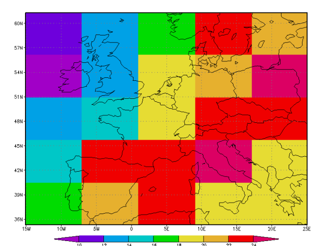
METEOSAT Imagery valid at 19:12 UTC 18 Jan 2007



Cloud ice analysis uncertainy

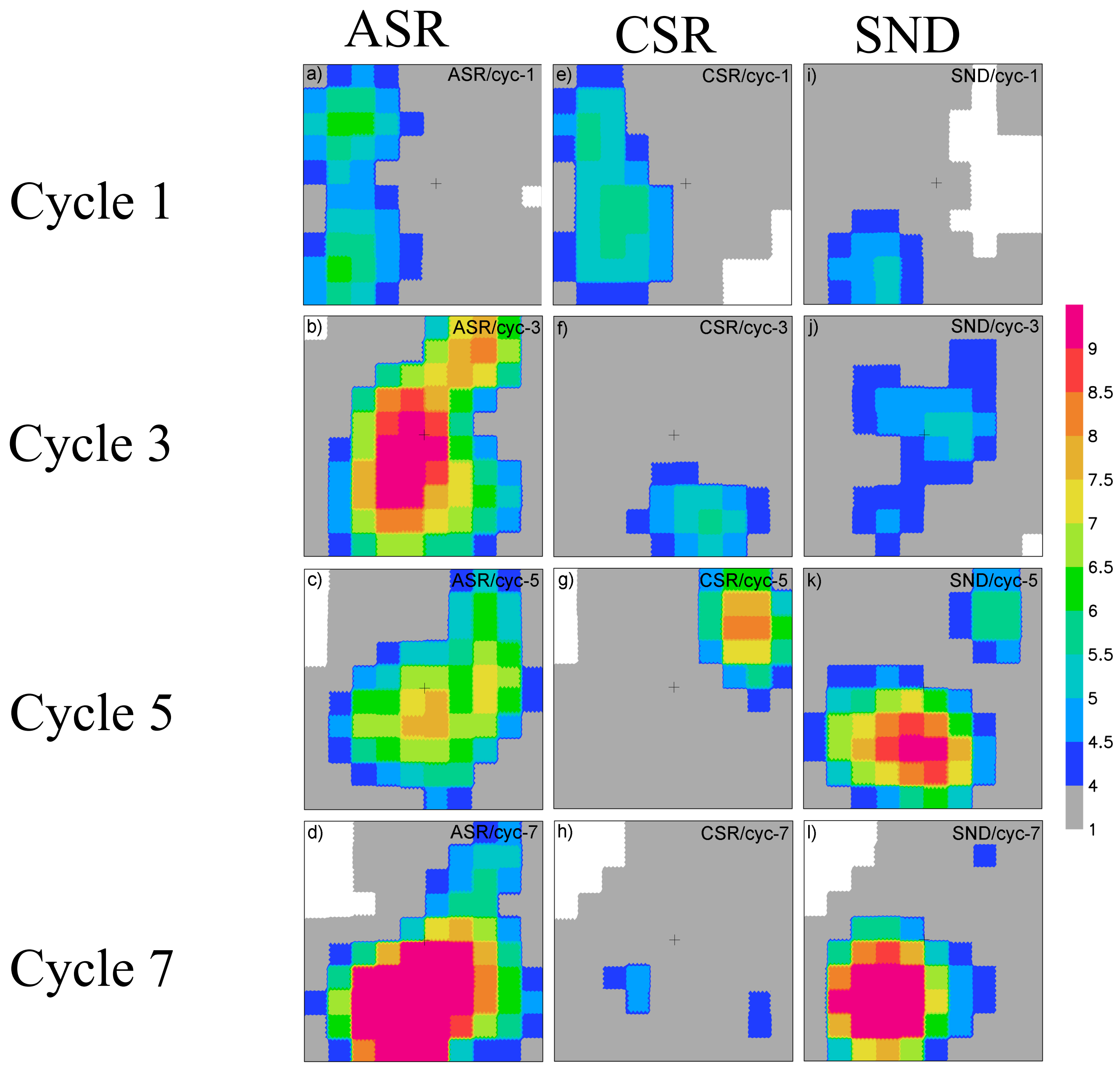


Degrees of Freedom for Signal (DFS)



Analysis uncertainty and DFS are flow-dependent, largest DFS in cloudy areas of the storm.

DFS for all-sky AMSU-A: hurricane Danielle (2010)



ASR – all-sky

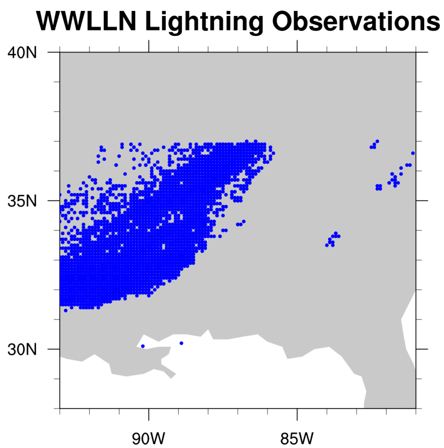
CSR – clear-sky

SND – all-sky, excluding AMSU-A channels 1-4 and 15

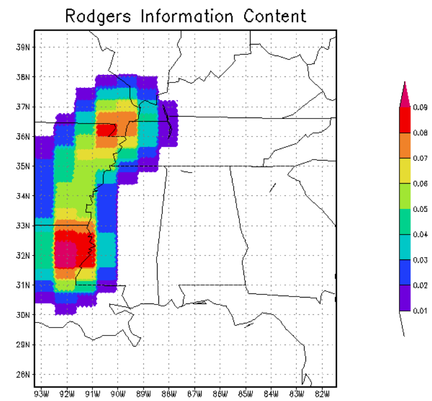
(From Zhang et al. 2012)

GOES-R GLM proxy: WWLLN lightning observation assimilation

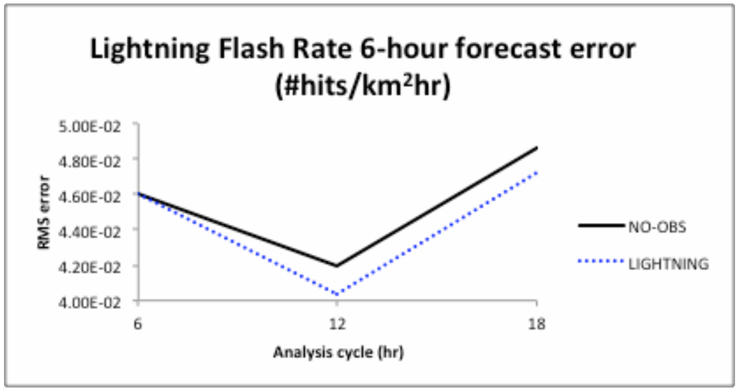
Preliminary results: **Assimilation of WWLLN lightning observations using WRF-NMM**



DFS



Forecast RMS error



- **DFS shows the utility of lightning observations in the analysis**
- **6-hour WRF-NMM forecast improved due to assimilating lightning observations**

Future

- Important to develop capability to extract maximum information from cloudy and precipitation-affected radiances

- Inner core all-sky satellite radiance observations are critical for improved analysis and prediction of hurricanes

- Cloudy radiance assimilation: computation aspects
  - RT computation can increase 2-3 times with scattering
  - number of observations can increase by an order of magnitude due to cloudy information
  - **20-30** times more expensive to compute – need parallelization, code optimization

- Shannon entropy measures are useful tool for assessing the impact of all-sky radiance assimilation

- Combine information from various sources: GOES-R, JPSS (MW, IR, Lightning)

References:

Non-differentiable minimization

Steward, J. L., I. M. Navon, M. Zupanski, and N. Karmitsa, 2011: Impact of Non-Smooth Observation Operators on Variational and Sequential Data Assimilation for a Limited-Area Shallow-Water Equation Model. *Quart. J. Roy. Meteorol. Soc.*, **138**, 323-339.

Chaotic system

Carrassi, A., S. Vannitsem, D. Zupanski, and M. Zupanski, 2009: The maximum likelihood ensemble filter performances in chaotic systems. *Tellus*, **61A**, 587-600.

All-sky infrared

Zupanski D., M. Zupanski, L. D. Grasso, R. Brummer, I. Jankov, D. Lindsey and M. Sengupta, and M. DeMaria, 2011: Assimilating synthetic GOES-R radiances in cloudy conditions using an ensemble-based method. *Int. J. Remote Sensing*, **32**, 9637-9659.

All-sky AMSR-E (microwave)

Zupanski, D., S. Q. Zhang, M. Zupanski, A. Y. Hou, and S. H. Cheung, 2011: A prototype WRF-based ensemble data assimilation system for downscaling satellite precipitation observations. *J. Hydromet.*, **12**, 118-134.

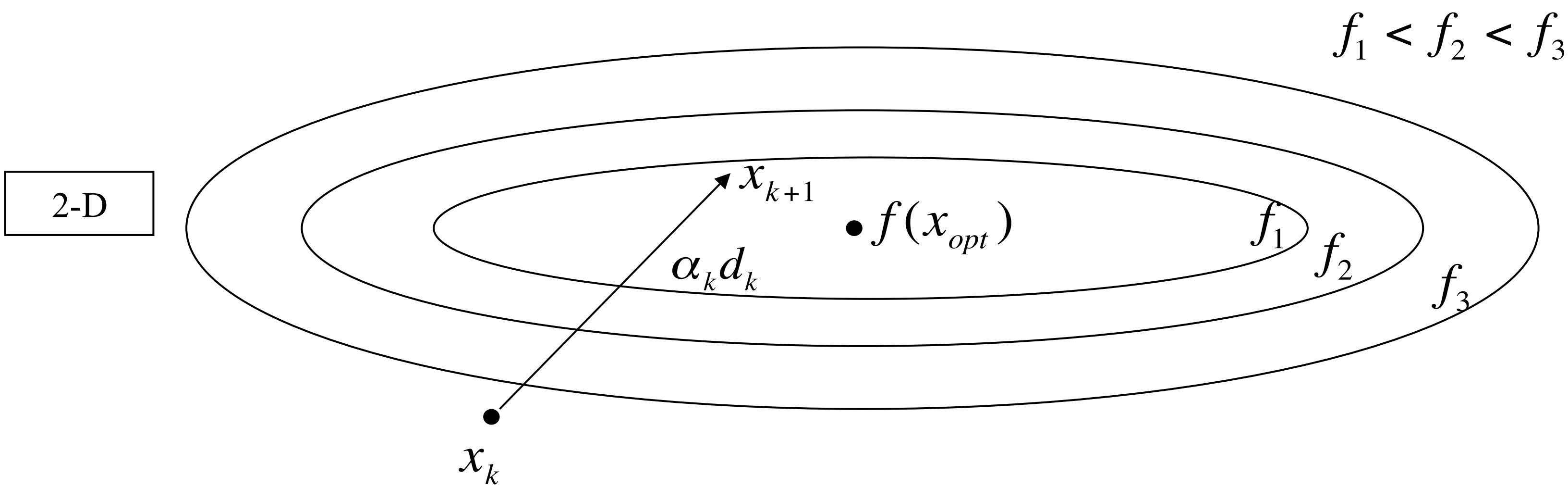
All-sky AMSU-A (microwave)

Zhang, M., M. Zupanski, M.-J. Kim, and J. Knaff, 2012: Assimilating AMSU-A radiances in TC inner core with NOAA operational HWRF and a hybrid data assimilation system: Danielle (2010). Submitted to *Mon. Wea. Rev*.

Additional MLEF literature available at **http://www.cira.colostate.edu/projects/ensemble/**

Iterative minimization

Minimize cost-function $f(x)$



Iterative approach: $x_{k+1} = x_k + \alpha_k d_k$

x **control vector**

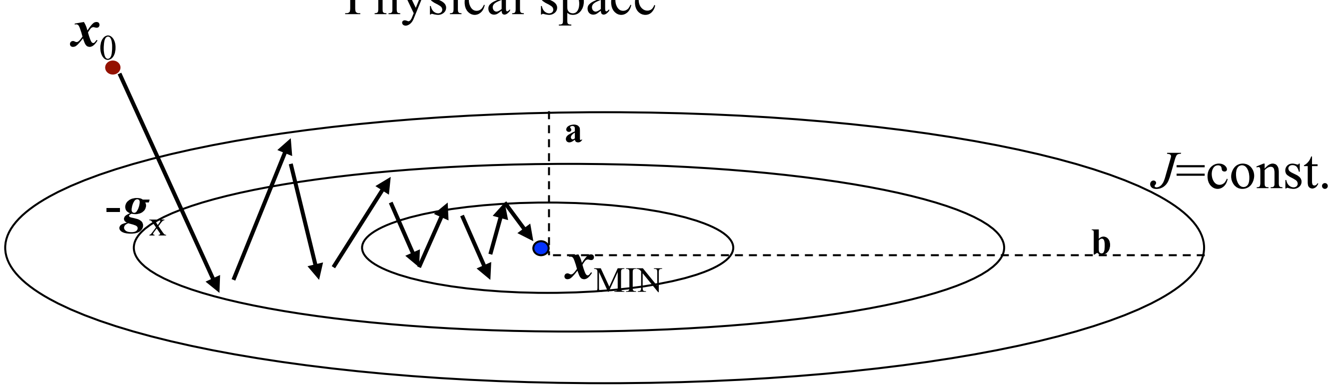
d **descent direction**

α **step length**

Iterative minimization can be used in both ensemble and variational DA

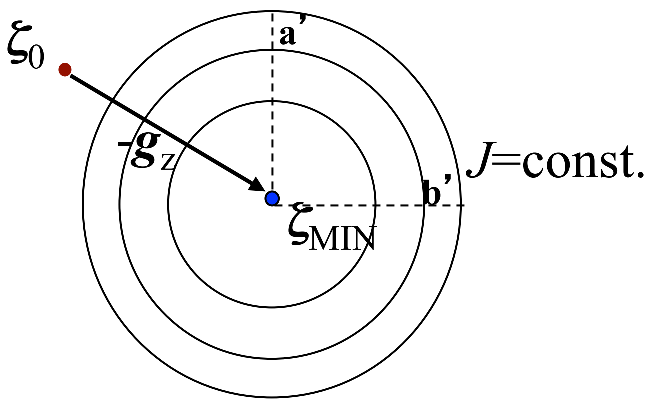
Hessian Preconditioning:
Geometric interpretation

Physical space



$$\kappa(Q) = \frac{b}{a} \approx 4$$

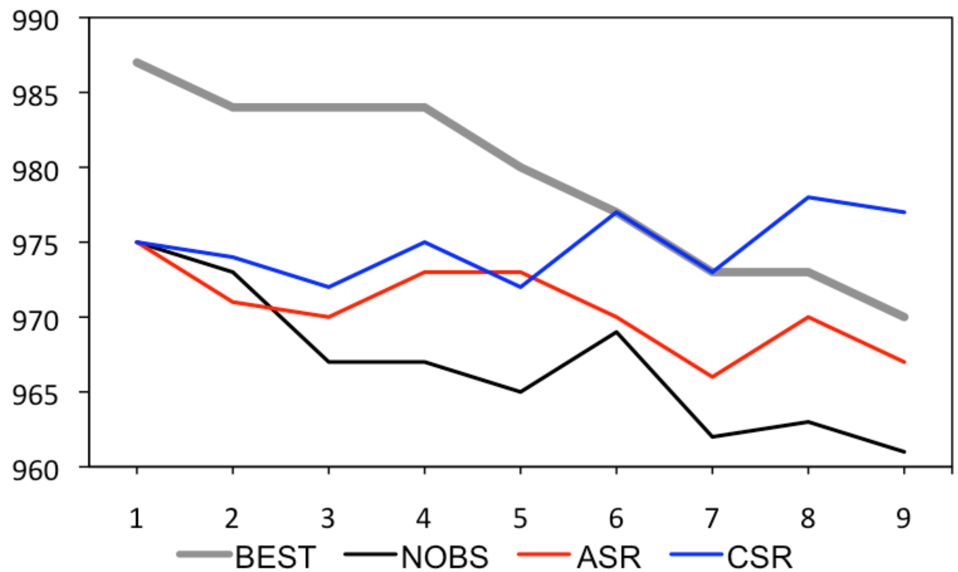
Preconditioning space



$$\kappa(Q_\zeta) = \frac{b'}{a'} = 1$$

In realistic applications $\kappa(Q)$~O(10^5-10^8). Hessian preconditioning is critical for efficient minimization.

Assimilation of all-sky AMSU-A radiances with HWRF-MLEF: Impact on hurricane intensity forecast



Hurricane Danielle (2010): Time series of the minimum sea level pressure (hPa) and NHC best track data (thick grey line), and MLEF-HWRF experiment outputs (colored lines) between 1800 UTC 24 Aug and 1800 UTC 26 Aug 2010 in 6 hour DA intervals.