The Abdus Salam
International Centre for Theoretical Physics

United Nations Educational, Scientific and Cultural Organization

International Atomic Energy Agency

Spring Colloquium on
'Regional Weather Predictability and Modeling'
April 11 - 22, 2005

1) Workshop **on Design and Use of Regional Weather Prediction Models,** *April 11 - 19*

2) Conference on **Current Efforts Toward Advancing the Skill of Regional Weather Prediction. Challenges and Outlook,** *April 20 - 22*

301/1652-7

The Maximum Likelihood Ensemble Filter (MLEF): An ensemble assimilation/prediction system based on control theory

M. Zupanski
Cooperative Institute for Research in the Atmosphere, CSU, Ft. Collins
USA

Spring Colloquium on
Regional Weather Predictability and Modeling
The Abdus Salam International Centre for Theoretical Physics
11-22 April, 2005, Trieste, Italy

The Maximum Likelihood Ensemble Filter (MLEF): An ensemble assimilation/prediction system based on Control Theory

Milija Zupanski

Cooperative Institute for Research in the Atmosphere
Colorado State University
Fort Collins, CO 80523-1375
ZupanskiM@CIRA.colostate.edu

In collaboration with:

Colorado State University: D. Zupanski, S. Fletcher, D. Randall, R. Heikes
Florida State University: I.M. Navon, B. Uzunoglu
Portland State University: D. Daescu

Outline

- MLEF algorithm
- Error covariance
- Parameter and model error estimation
- Non-Gaussian MLEF framework
- Future research directions

Maximum Likelihood Ensemble Filter (MLEF)

- Estimate of the *conditional mode* of the posterior PDF

- Ensembles used to estimate the *uncertainty* of the conditional mode

- *Non-derivative* minimization with Hessian preconditioning:
  (Generalized conjugate-gradient and BFGS quasi-Newton algorithms)

- *Non-Gaussian* errors allowed (minimization of non-Gaussian cost-function)

- Augmented control variable: *initial conditions*, *model bias*, *empirical parameters*, *boundary conditions*

- Related to: (*i*) variational data assimilation, (*ii*) Iterative Kalman filters, and (*iii*) Ensemble Transform Kalman Filter – ETKF

MLEF Mathematical Framework

Forecast (prior) error covariance

$$\boldsymbol{P}_f^{1/2} = [\boldsymbol{p}_1^f \quad \boldsymbol{p}_2^f \quad \cdots \quad \boldsymbol{p}_{N_E}^f]$$

$$\boldsymbol{p}_i^f = \mathcal{M}(\boldsymbol{x} + \boldsymbol{p}_i^a) - \mathcal{M}(\boldsymbol{x})$$

Minimize cost function in subspace spanned by ensemble perturbations $\boldsymbol{p}_i^f$

$$J = \frac{1}{2}[\boldsymbol{x} - \boldsymbol{x}^b]^T \boldsymbol{P}_f^{-1} [\boldsymbol{x} - \boldsymbol{x}^b] + \frac{1}{2}[\boldsymbol{y}_{obs} - \mathcal{H}(\boldsymbol{x})]^T \boldsymbol{R}^{-1} [\boldsymbol{y}_{obs} - \mathcal{H}(\boldsymbol{x})]$$

Non-derivative iterative minimization with preconditioning

- No differentiability assumption: works for all bounded operators
- Use with Conjugate gradient, LBFGS quasi-Newton algorithms
- Define generalized gradient, generalized Hessian

Employ Taylor formula for variations:

First variation = generalized gradient

Second variation = generalized Hessian

$$\Delta J = J(\boldsymbol{x} + \delta\boldsymbol{x}) - J(\boldsymbol{x}) = \delta J + \frac{1}{2}\delta^2 J + \frac{1}{3!}\delta^3 J + \cdots$$

MLEF Mathematical Framework

Generalized **inverse Hessian**

$$\left(\nabla_G^2 J\right)^{-1} = \boldsymbol{P}_f^{1/2}\left[\boldsymbol{I} + \left(\boldsymbol{Z}^b\right)^T \boldsymbol{Z}^b\right]^{-1} \boldsymbol{P}_f^{T/2}$$

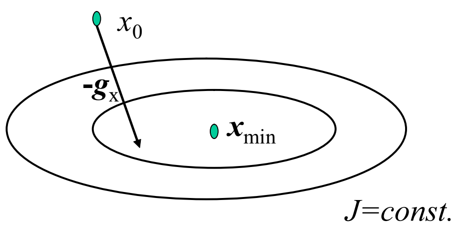
$$\boldsymbol{Z}^b = \left[\boldsymbol{z}_1^b \quad \boldsymbol{z}_2^b \quad \cdot \quad \cdot \quad \boldsymbol{z}_{N_E}^b\right] \qquad \boldsymbol{z}_i^b = \boldsymbol{R}^{-1/2}\left[\mathcal{H}(\boldsymbol{x}^b + \boldsymbol{p}_i^f) - \mathcal{H}(\boldsymbol{x}^b)\right]$$

Change of variable (Hessian preconditioning)

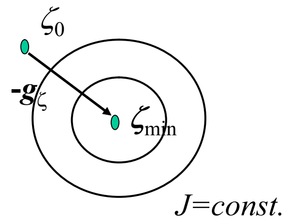
$$\boldsymbol{x}_k = \boldsymbol{x}^b + \boldsymbol{P}_f^{1/2}\left[\boldsymbol{I} + \left(\boldsymbol{Z}^b\right)^T \boldsymbol{Z}^b\right]^{-1/2} \zeta_k$$

$$\zeta_{k+1} = \zeta_k + \alpha_k \boldsymbol{d}_k$$

Physical space



Preconditioning space



Generalized **gradient**

$$\left(\nabla_G J\right)_k = \left[\boldsymbol{I} + \left(\boldsymbol{Z}^b\right)^T \boldsymbol{Z}^b\right]^{-1/2} \zeta_k - \left(\boldsymbol{Z}^k\right)^T \boldsymbol{R}^{-1}\left(\boldsymbol{y} - \mathcal{H}(\boldsymbol{x}_k)\right)$$

$$\boldsymbol{Z}^k = \left[\boldsymbol{z}_1^k \quad \boldsymbol{z}_2^k \quad \cdot \quad \cdot \quad \boldsymbol{z}_s^k\right] \qquad \boldsymbol{z}_i^k = \boldsymbol{R}^{-1/2}\left[\mathcal{H}(\boldsymbol{x}_k + \boldsymbol{p}_i^f) - \mathcal{H}(\boldsymbol{x}_k)\right]$$

k – iteration index

MLEF Mathematical Framework

Analysis (posterior) error covariance

$$\boldsymbol{P}_a^{1/2} = \boldsymbol{P}_f^{1/2}(\boldsymbol{I} + \boldsymbol{Z}^T\boldsymbol{Z})^{-1/2}$$

$$\boldsymbol{z}_i = \boldsymbol{R}^{-1/2}\mathcal{H}(\boldsymbol{x} + \boldsymbol{p}_i^f) - \boldsymbol{R}^{-1/2}\mathcal{H}(\boldsymbol{x})$$

- **Analysis error covariance estimated from minimization algorithm**
- **Inverse Hessian = Analyis error covariance (linear assumption)**

How to apply linear result to nonlinear problem?

- **If the Hessian is calculated at the minimum, it is close to the truth**
- $\|x^{true} - x^a\| < \varepsilon$ **: for small** ε **linear theory is valid**
- **Avoid sample error covariance calculations**
  - noisy for limited sample
  - inflation factors

• For good Hessian preconditioning, expect superior minimization convergence
• Valid for arbitrary nonlinear operators and cost functions

Prediction models

(1) Korteweg-de Vries-Burgers (KdVB)

- one-dimensional, one variable, two solitons
- nonlinear advection, dispersion, diffusion

(2) NASA GEOS-5 column precipitation model

- one-dimensional, two variables (tempertaure, specific humidity)
- forced by the global GEOS-5 model

(3) CSU global shallow-water model

- two-dimensional, three variables (height, stream function, velocity potential)
- finite-difference, on a twisted icosahedral grid

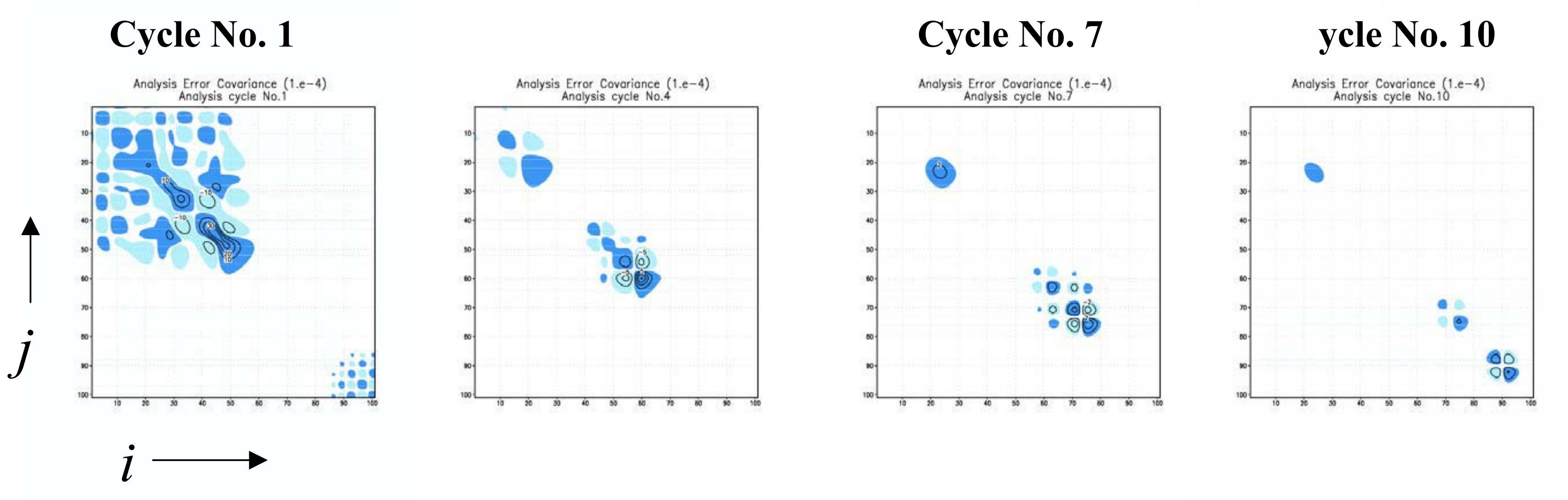
(4) CSU RAMS – non-hydrostatic model with explicit microphysics

- Exner function, potential temperature, winds, specific humidity, rain, graupel, hail, aggregates, pristine ice, snow
- finite-difference, Arakawa C-grid

KdVB model: covariance localization

• **KdVB model**

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + 6\frac{\partial^3 u}{\partial x^3} = \nu\frac{\partial^2 u}{\partial x^2}$$



• **Initial error covariance noisy, but quickly becomes spatially localized**

• **No need to artificially force error covariance localization**

Why does the localization happens by itself ?

Error covariance localization

$n = 1$

$$\left(P_f^{1/2}\right)_1 = M_1 P_0^{1/2}$$

$$\left(P_a^{1/2}\right)_1 = \left(P_f^{1/2}\right)_1 \left(I + Z_1^T Z_1\right)^{-1/2} = \left[M_1 P_0^{1/2}\right]\left(I + Z_1^T Z_1\right)^{-1/2}$$

$n = 2$

$$\left(P_f^{1/2}\right)_2 = M_2 \left(P_a^{1/2}\right)_1 = \left[M_2 M_1 P_0^{1/2}\right]\left(I + Z_1^T Z_1\right)^{-1/2}$$

$$\left(P_a^{1/2}\right)_2 = \left(P_f^{1/2}\right)_2 \left(I + Z_2^T Z_2\right)^{-1/2} = \left[M_2 M_1 P_0^{1/2}\right]\left(I + Z_1^T Z_1\right)^{-1/2}\left(I + Z_2^T Z_2\right)^{-1/2}$$

$n = N$

$$\left(P_f^{1/2}\right)_N = M_N \left(P_a^{1/2}\right)_{N-1} = \left[M_N \ldots M_2 M_1 P_0^{1/2}\right]\left(I + Z_1^T Z_1\right)^{-1/2}\left(I + Z_2^T Z_2\right)^{-1/2} \cdots \left(I + Z_{N-1}^T Z_{N-1}\right)^{-1/2}$$

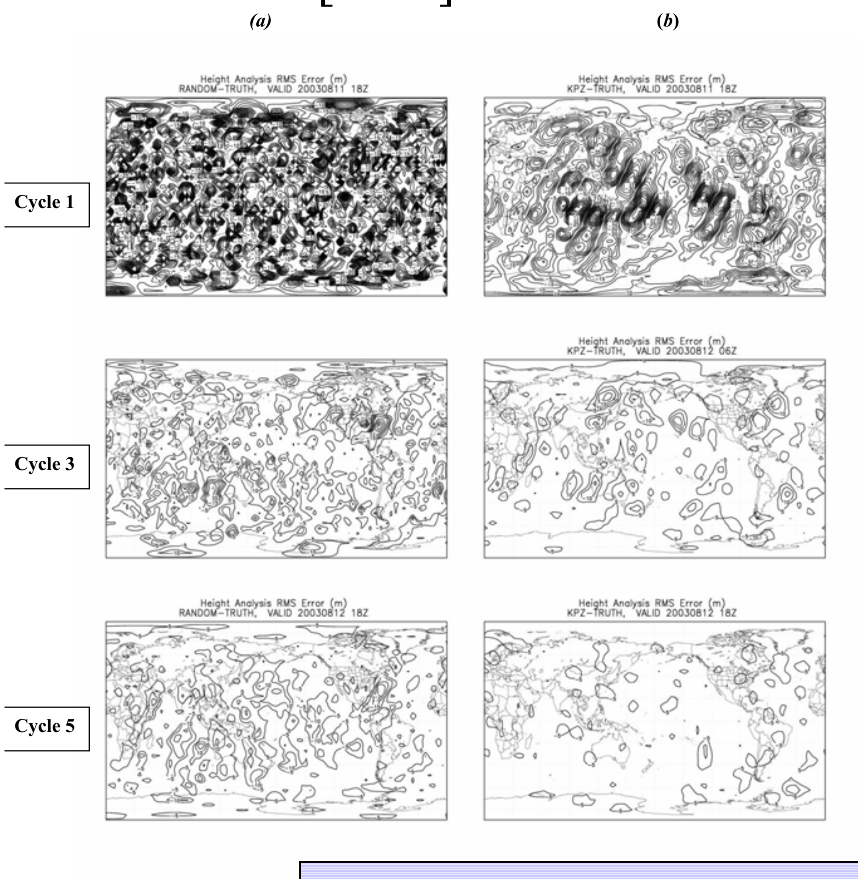
$$\left(P_a^{1/2}\right)_N = \underbrace{\left[M_N \ldots M_2 M_1 P_0^{1/2}\right]}_{\text{Lyapunov vector}}\left(I + Z_1^T Z_1\right)^{-1/2}\left(I + Z_2^T Z_2\right)^{-1/2} \cdots \left(I + Z_N^T Z_N\right)^{-1/2}$$

Possible explanation for error covariance localization:

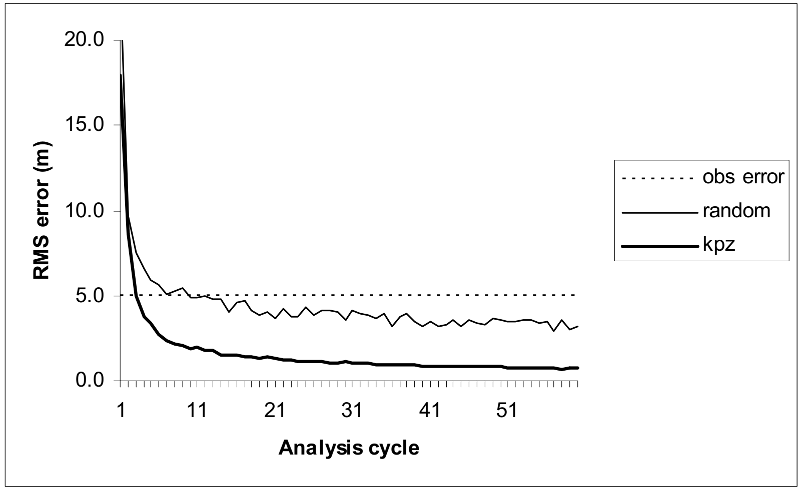
Dynamic localization of Lyapunov vectors

MLEF with CSU global shallow-water model

Height analysis increment
$[x^a - x^b]$



Height RMS error $[x^a - x^t]$



Impact of the initial choice for ensemble perturbations:
Correlated initial ensemble perturbations can significantly improve algorithm performance (Kardar-Parisi-Zhang equation + correlations)

Analysis increment

• **Analysis increment is a linear combination of the square-root forecast error covariance column-vectors:**

$$\boldsymbol{x}^a - \boldsymbol{x}^b = \boldsymbol{P}_f^{1/2}\boldsymbol{w} = w_1\boldsymbol{p}_1^f + w_2\boldsymbol{p}_2^f + \cdots + w_{N_E}\boldsymbol{p}_{N_E}^f$$

$$\boldsymbol{w} = \left(\boldsymbol{I} + \boldsymbol{Z}^T\boldsymbol{Z}\right)^{-1/2}\zeta$$

• If the column-vectors are smooth, the analysis increment will be smooth

• Due to the $\boldsymbol{P}_f$ dynamic localization, need to improve the initial error covariance fast

Impact of dynamics:

• **smoothness of the analysis**

• **localization of error covariance**

Model error and parameter estimation in MLEF

State augmentation approach

$$\mathbf{z}_n = \begin{bmatrix} \mathbf{x}_n \\ \mathbf{b}_n \\ \boldsymbol{\gamma}_n \end{bmatrix} = \begin{bmatrix} M_{n,n-1} & (1-\alpha^n)G & 0 \\ 0 & \mathbf{I} & 0 \\ 0 & 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{n\text{-}1} \\ \mathbf{b}_{n\text{-}1} \\ \boldsymbol{\gamma}_{n-1} \end{bmatrix} + \begin{bmatrix} \alpha^n \\ 0 \\ 0 \end{bmatrix} \Phi_0$$

$\mathbf{x}$ – state vector (initial conditions) ; $\mathbf{b}$ – model bias ; γ – empirical parameters

Augmented control variable:

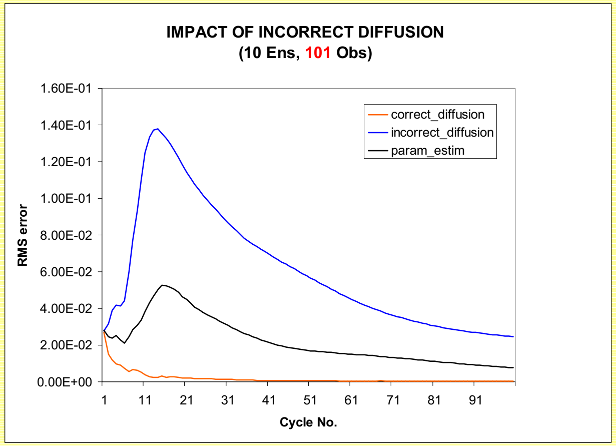
$$\mathbf{z} = (\mathbf{x}_0, \mathbf{b}, \gamma)$$

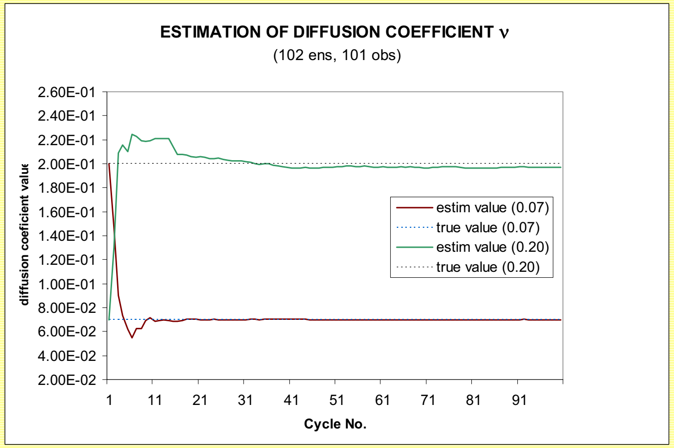
Augmented error covariance:

$$\mathbf{P} = \begin{pmatrix} P_{x_0,x_0} & P_{x_0,b} & P_{x_0,\gamma} \\ P_{b,x_0} & P_{b,b} & P_{b,\gamma} \\ P_{\gamma,x_0} & P_{\gamma,b} & P_{\gamma,\gamma} \end{pmatrix}$$

Model error and parameter estimation with KdVB model

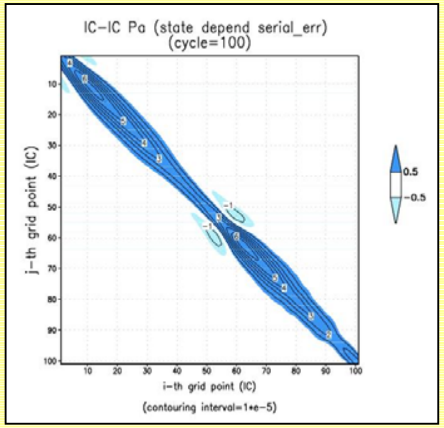
Parameter estimation (diffusion coefficient)

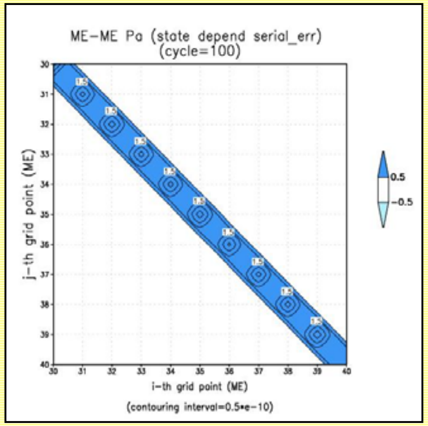


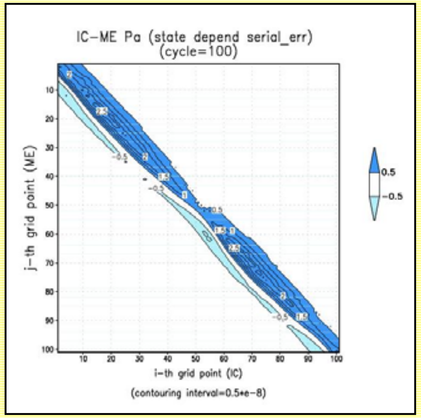


Error covariance block matrices

IC-IC **ME-ME** **IC-ME**



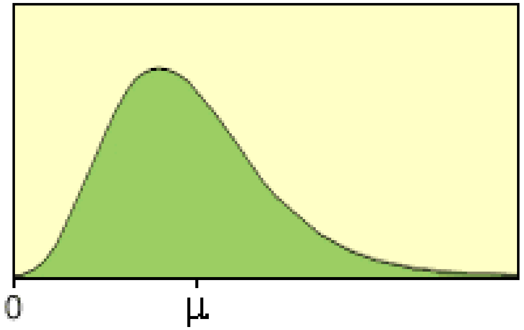




Significant cross-correlation between initial conditions and model error

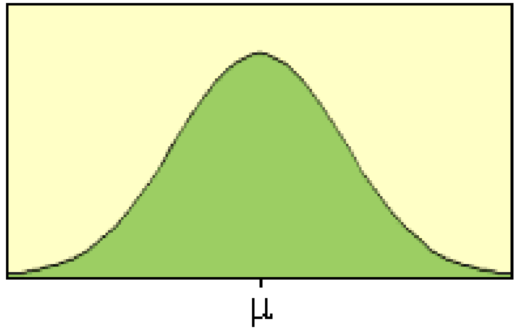
Non-Gaussian MLEF framework

Log-Normal PDF



$$\phi(x)=\begin{cases}\dfrac{exp\left(-\dfrac{1}{2}\left(\dfrac{log(x)-m}{s}\right)^2\right)}{xs\sqrt{2\pi}} & x>0\\ 0 & \text{otherwise}\end{cases}$$

Gaussian PDF



$$\phi(x)=\frac{exp\left(-\dfrac{1}{2}\left(\dfrac{x-\mu}{\sigma}\right)^2\right)}{\sigma\sqrt{2\pi}}$$

***Log-likelihood cost-function*:** $J(x)=-\ln[\phi(x)]$

Maximize posterior PDF ⇔ minimize cost function

Posterior conditional PDF $\phi(X\mid Y)=\dfrac{\phi(X)\phi(Y\mid X)}{\phi(Y)}$

$$\boldsymbol{max}\{\phi(X\mid Y)\}\propto \boldsymbol{max}\{\phi(X)\phi(Y\mid X)\}=\boldsymbol{min}\{-\ln[\phi(X)]-\ln[\phi(Y\mid X)]\}$$

Non-Gaussian MLEF framework:
Lognormal height observation errors

Assume:

- **Gaussian prior probability distribution (state vector)** $\phi(X) \sim Gaussian$

- **Observations: Height (Lognormal), Wind (Gaussian)**

$\phi(Y_{HEIGHT} \mid X) \sim Lognormal$

$\phi(Y_{WIND} \mid X) \sim Gaussian$

$$\frac{\boldsymbol{y}_{HEIGHT}}{\mathcal{H}(\boldsymbol{x})} = \boldsymbol{\varepsilon} \Rightarrow \ln[\boldsymbol{y}_{HEIGHT}] - \ln[\mathcal{H}(\boldsymbol{x})] = \ln \boldsymbol{\varepsilon}$$

$$\boldsymbol{y}_{WIND} - \mathcal{H}(\boldsymbol{x}) = \boldsymbol{\varepsilon}$$

Minimize mixed Normal-Lognormal cost function:

$$J(\boldsymbol{x}) = \frac{1}{2}\left(\boldsymbol{x} - \boldsymbol{x}^b\right)^T \boldsymbol{P}_f^{-1}\left(\boldsymbol{x} - \boldsymbol{x}^b\right) + \frac{1}{2}\left(y_{WIND} - \mathcal{H}(\boldsymbol{x})\right)^T \boldsymbol{R}^{-1}\left(y_{WIND} - \mathcal{H}(\boldsymbol{x})\right)$$

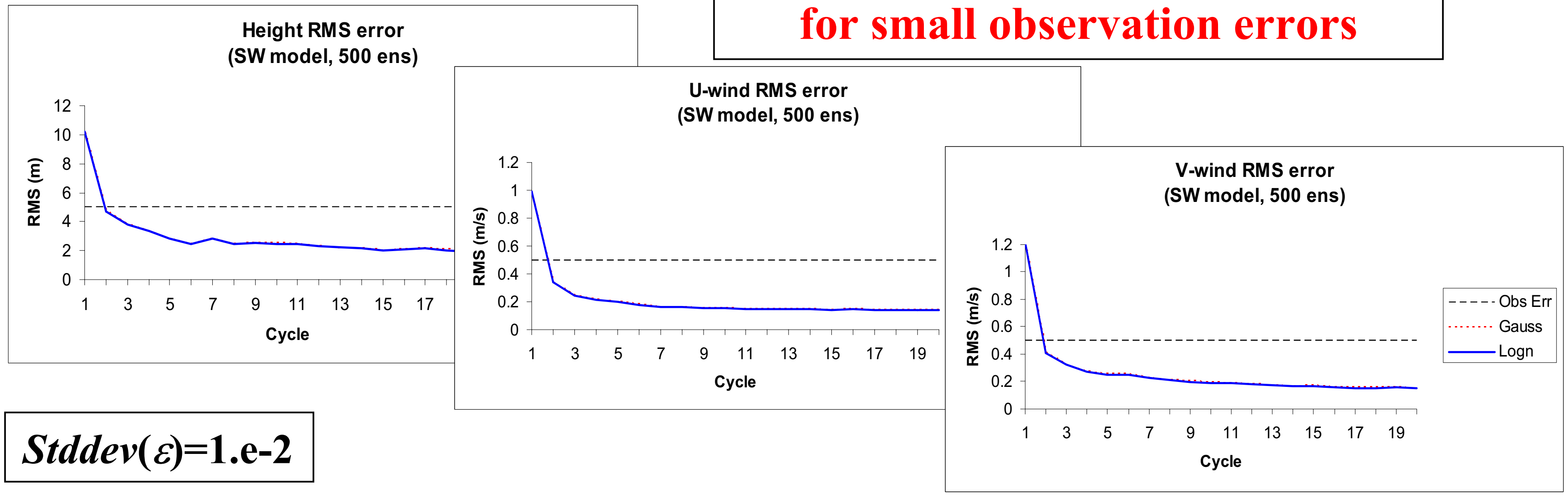
$$+ \frac{1}{2}\left(\ln\left[\frac{y_{HEIGHT}}{\mathcal{H}(\boldsymbol{x})}\right] - m\right)^T \boldsymbol{R}_S^{-1}\left(\ln\left[\frac{y_{HEIGHT}}{\mathcal{H}(\boldsymbol{x})}\right] - m\right) + \sum_{i=1}^{N_{obs}} \ln\left[\frac{y_{HEIGHT}}{\mathcal{H}(\boldsymbol{x})}\right]_i$$

Higher nonlinearity of the cost function compared to the Gaussian

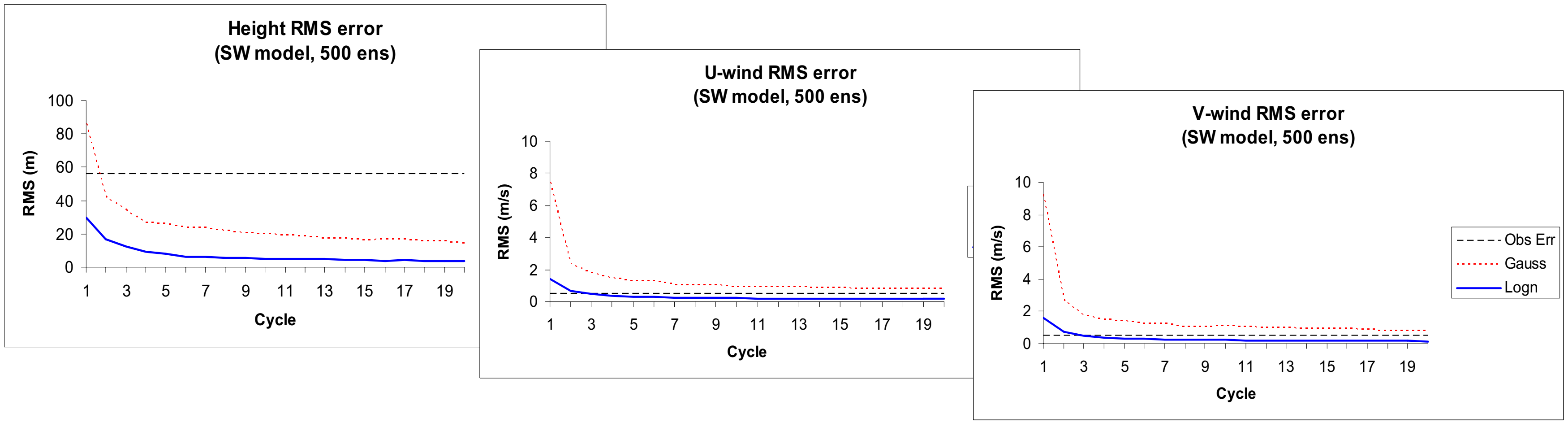
Impact of Lognormal observation erros:
Analysis RMS errors

Stddev(ε)=1.e-3

Gaussian framework works only for small observation errors



Stddev(ε)=1.e-2

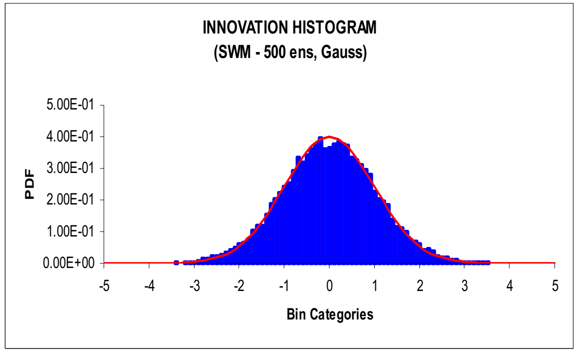


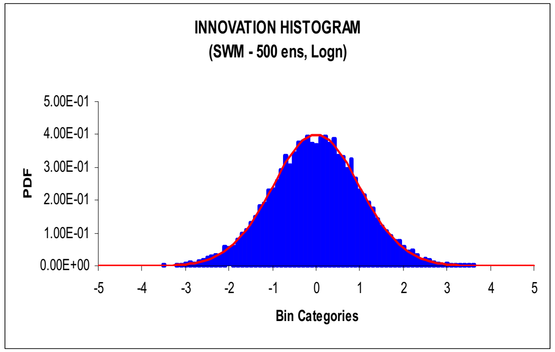
Lognormal framework works for all error magnitudes

Impact of Lognormal height observation errros:
Innovation histogram

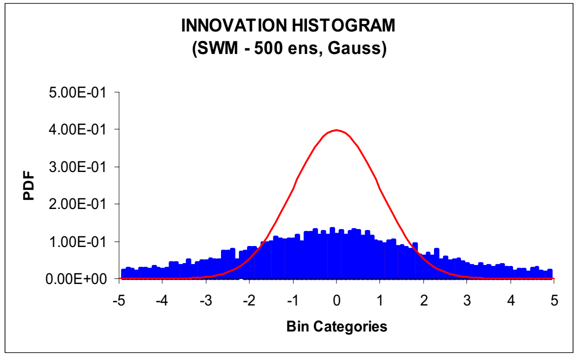
Gaussian framework | **Lognormal framework**

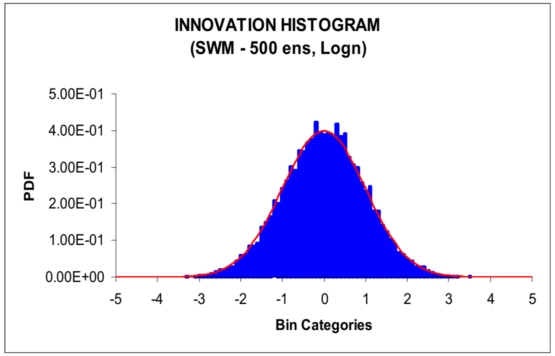
Stddev(ε)=1.e-3





Stddev(ε)=1.e-2





Height innovations → $y - \mathcal{H}(\mathbf{x})$ | $\ln\left[\frac{y}{\mathcal{H}(\mathbf{x})}\right] - m$

Generalized non-Gaussian framework can handle Gaussian, Lognormal, or mixed PDF errors !

Ongoing work

❑ **NCEP Global Forecasting System model and real data**

- Apply MLEF with GFS and compare with other EnKF algorithms
- Compare (and combine) the conditional mean and mode PDF estimates
- Develop and test *double-resolution* MLEF
  - Operational (high) resolution control analysis/forecast
  - Low (coarse) resolution ensembles – error covariance

❑ **Carbon Cycle data assimilation**

- Apply MLEF with fully coupled SiB-CASA-RAMS atmospheric-carbon-biomass model
- Assimilate carbon concentration globally and locally
- Empirical parameter estimation in carbon models

❑ **GOES-R Risk Reduction**

- Evaluate the impact of GOES-R measurements in applications to severe weather and tropical cyclones
- Information content of GOES-R measurements

Future Research Directions

Algorithmic and theoretical aspects:

Development of a fully non-Gaussian algorithm
- Allow for non-Gaussian *state variable* components
- Generalized algorithm with a list of PDFs

Representativeness error
- How well the ensembles represent the truth
- Improve robustness of the algorithm

Model bias and parameter estimation
- Improve prediction models by learning about its errors and uncertainties
- Develop as a probabilistic tool for new model development

Initiation of ensemble perturbations ("cold" start)
- Competition between the rate of covariance localization and the development of a representative ensemble basis
- Kardar-Parisi-Zhang equation + correlations
- Tikhonov regularization as a criteria for ensemble size reduction

  (B. Uzunoglu: reduction from 1000 ensembles to about 200 in few cycles !)

Future Research Directions

Applications:

Climate models

- Assimilation of clouds and precipitation – MMF (Multiscale Modeling Framework)
- Coupled climate-ocean models
- Predictability

Microscale models

- Boundary layer, 100m-500m horizontal resolution
- Probabilistic transfer and interaction between scales

Carbon data assimilation

- Exploit assimilation of new measurements (OCO-Orbiting Carbon Observatory)

Cross-over the existing scientific boundaries

- Apply probabilistic assimilation/prediction to other science disciplines
- General, adaptive, algorithmically simple algorithm opens new possibilities