Evaluation of Ensemble Data Assimilation for the Army Scale Meteorology

Milija Zupanski[1], Dusanka Zupanski[1],
Gustavo Carrio[2], and William Cotton[2]

[1]DoD Center for Geosciences-Atmospheric Research,
Cooperative Institute for Research in the Atmosphere
[2]Department of Atmospheric Science
Colorado State University, Fort Collins, CO

OUTLINE

- ***Research Objectives***
  - Develop an ensemble assimilation/prediction system to estimate the ***model state*** + ***uncertainties***
  - Evaluate the system in application with ***high*** model resolution, ***real*** observations
  - Impact of model error/parameter estimation

- ***Relevance to the Army scale Met problem***
  - Army microscale model(s) use real observations
  - Uncertainty of assimilation/prediction needed in DoD applications
  - Transfer of probabilistic (uncertainty) weather information to TDAs

- ***Research results and Deliverables***
  - Overview of the Maximum Likelihood Ensemble Filter (MLEF)
  - MLEF applications with mesoscale and LES RAMS models

- ***Future work***

RESEARCH OBJECTIVES

❑ *Model state + Uncertainties*

- Typically, DA methods estimate only model state, without reliable information about uncertainties of the state
- Uncertainty is often represented by a (co)variance matrix
- Ensemble methods estimate and update uncertainties
- Exploit the possibility to use ensemble methods in data assimilation, forecasting, a ***single algorithm with complete feed-back***

❑ *Model error/parameter estimation needed since*

- Models are imperfect
- Many unknown empirical parameters (physics, parameterization scheme)
- Any data assimilation needs to address this question
- Use to improve/correct the model performance, equations
- Ideally, want both the model error/parameter ***and*** its uncertainty

RESEARCH OBJECTIVES

- ***Real observations***
  - The most important source of information
  - Difficult to assimilate, wthout compensating for model error
  - Which part is the initial conditions, which part is the model error ?
  - Observation operators: ***transformation*** from model to observation
  - Can ensemble methods assimilate large number of observations (satellite, radar/lidar) ?

- ***High resolution models required for operations***
  - Are the ensemble methods applicable to the problem with high-dimensional model state?
  - Boundary layer meteorological applications
  - Operational weather forecasting
  - Are the ensemble methods computationally efficient in high resolution ?

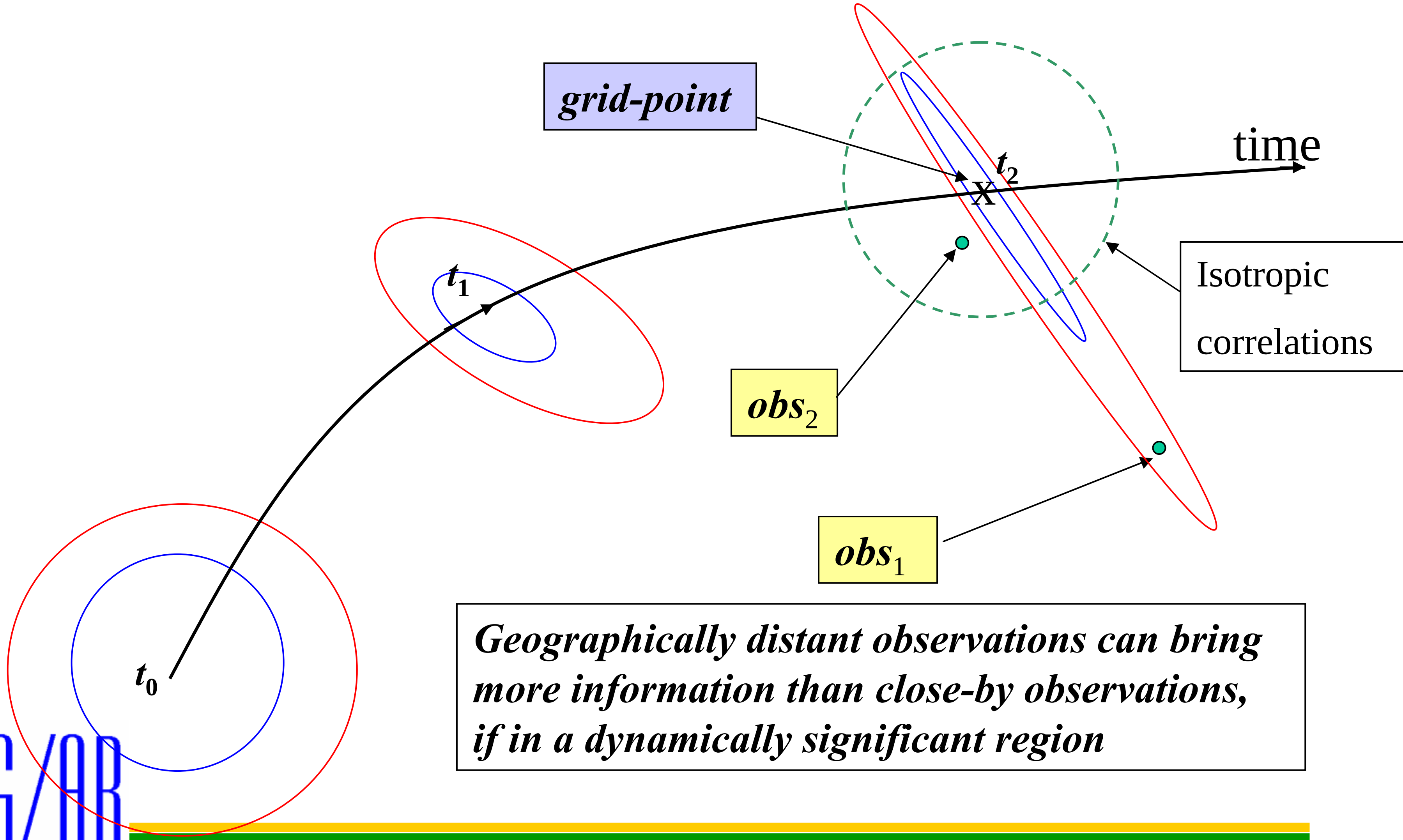
RELEVANCE TO THE ARMY SCALE METEOROLOGY

- ***Goal is to develop an ensemble assimilation/prediction system for use with Army microscale models and real observations***
  - Uncertainty estimate (Tactical Decision Aids)

- ***Critical requirements***
  - Ensemble-based methodology must work with **high-dimensional** Army meteorological models
  - Ensemble system has to be able to estimate and update the **model errors** and **empirical parameters**, needed in realistic Army applications
  - Computationally **efficient** system: use of Army High Performance Computing (**HPC**)
  - Spin-up time

- ***Current CGAR Research Status:***
  - Elements of the APP (paragraphs V.d.1 and 2) have been completed
  - Item Vd3 is being addressed at this time

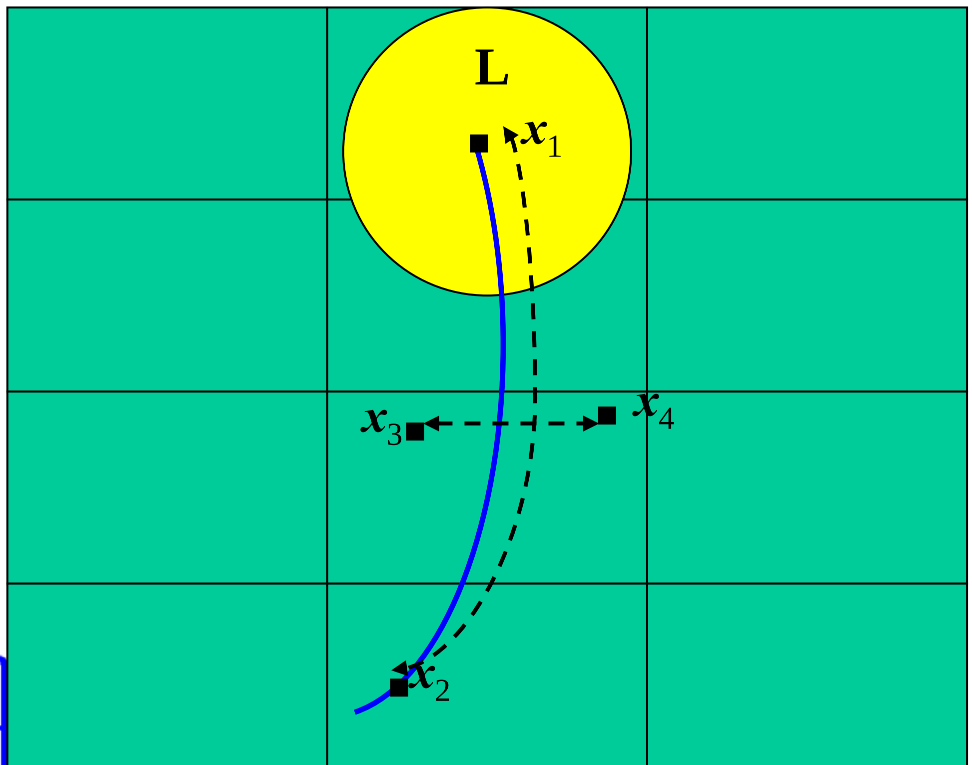
Ensembles: Flow-dependent forecast error covariance and spread of information from observations



Geographically distant observations can bring more information than close-by observations, if in a dynamically significant region

Dynamical localization vs. prescribed structure

- Can we prescribe correlations ?
- Can we afford to neglect distant correlations ?
- Degrees of freedom
- Computational efficiency



Weak correlations across frontal zone

$x_3 \leftarrow\rightarrow x_4$ is small

Strong correlations along frontal zone

$x_1 \leftarrow\rightarrow x_2$ is large

Nonlinear observation operators

- **The KF, EKF, EnKF solution identical to the minimization of a quadratic cost-function**

$$\boldsymbol{x}^a = \boldsymbol{x}^f + \boldsymbol{P}_f \boldsymbol{H}^T \left(\boldsymbol{H}\boldsymbol{P}_f \boldsymbol{H}^T + \boldsymbol{R}\right)^{-1} [\boldsymbol{y} - \mathcal{H}(\boldsymbol{x}^f)]$$

$$\boldsymbol{x}^a = \boldsymbol{x}^f + \alpha \left(\frac{\partial^2 J}{\partial \boldsymbol{x}^2}\right)^{-1} \left(-\frac{\partial J}{\partial \boldsymbol{x}}\right) = \boldsymbol{x}^f + \alpha \boldsymbol{P}_f \boldsymbol{H}^T \left(\boldsymbol{H}\boldsymbol{P}_f \boldsymbol{H}^T + \boldsymbol{R}\right)^{-1} [\boldsymbol{y} - \mathcal{H}(\boldsymbol{x}^f)]$$

- ***Two strategies for nonlinear observation operators*:**

(1) Use ***linear*** KF solution, combined with nonlinear operators in covariance calculation

- EnKF algorithms

(2) Directly search for solution of ***nonlinear*** problem by ***minimizing non-quadratic cost function***

- Maximum Likelihood Ensemble Filter (MLEF) – *conditional mode*

MAXIMUM LIKELIHOOD ENSEMBLE FILTER (MLEF)

(Zupanski, 2005, *MWR*; Zupanski and Zupanski, 2005, *MWR*)

- ***Non-sample based ensemble method***
  - Nonlinear approach to estimating the basis of a dynamically unstable subspace
  - Reduces to Kalman Filter with N_{state} ensembles (for Gaussian PDF and linear models, superior nonlinear solution otherwise)
  - **No need for artificial error covariance inflation, localization** i.e. assimilation methods all use "invented" covariance parameters – except MLEF

- ***Analysis***
  - Non-differentiable minimization of the cost function in ensemble subspace
  - Iterative solution for nonlinear models and observation operators
  - Analysis error covariance from the inverse Hessian (minimization)

- ***Forecast***
  - The previous cycle posterior error covariance used for the initial perturbations in ensemble forecasting
  - Implicit dynamic localization of error covariances

MLEF Analysis Step

Use the forecast (prior) error covariance square-root

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

Ensemble size

$$\boldsymbol{p}_i^f = \begin{pmatrix} p_{1,i}^f \\ p_{2,i}^f \\ \vdots \\ p_{S,i}^f \end{pmatrix}$$

State-space dimension

Minimize cost function in the subspace ***spanned by ensemble perturbations***

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

Similar to variational, however:

- **Non-differentiable iterative minimization with superior preconditioning**
- **Solution in ensemble subspace (reduced rank)**
- **Analysis uncertainty estimate**

Hessian preconditioning in MLEF

Change of variable (Hessian preconditioning)

$$x_k = x^f + P_f^{1/2}\left[I + \left(Z^b\right)^T Z^b\right]^{-1/2} \zeta_k$$

Ensemble-size matrix

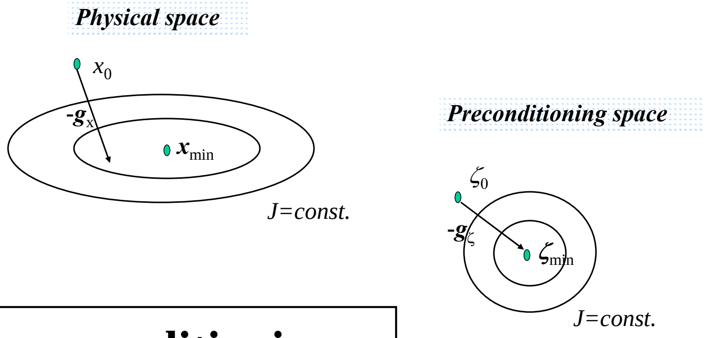
$$Z^b = \left[z_1^b \quad z_2^b \quad \cdot \quad \cdot \quad z_{N_E}^b\right]$$

$$z_i^b = R^{-1/2}\left[\mathcal{H}(x^f + p_i^f) - \mathcal{H}(x^f)\right]$$

Background (first guess)

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

k – iteration index



Use ETKF transformation in Hessian preconditioning

Analysis (posterior) error covariance

$$P_a^{1/2} = P_f^{1/2}(I + Z^T Z)^{-1/2}$$

$$P_a^{1/2} = [\,p_1^a \quad p_2^a \quad \cdots \quad p_{N_E}^a\,]$$

$$z_i = R^{-1/2}\mathcal{H}(x^a + p_i^f) - R^{-1/2}\mathcal{H}(x^a)$$

Analyisis (minimizer)

- **Analysis error covariance estimated from minimization algorithm**
- **At the minimum (x_{min}=x^a) use**

Inverse Hessian = Analysis error covariance

Forecast (prior) error covariance

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

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

$$\boldsymbol{p}_i^f = \begin{pmatrix} p_{1,i}^f \\ p_{2,i}^f \\ \vdots \\ p_{S,i}^f \end{pmatrix}$$

- **Control vector is the *most likely* forecast**
- **Square-root used in the algorithm**
- **No sampling of error covariance**

MLEF has been successfully applied to variety of models and shows a robust, consistent behavior

OVERVIEW OF MLEF

- ***General method for estimating the state + uncertainty***

- ***Algorithmically simple and efficient – HPC computing***

- ***Non-differentiable minimization – handles discontinuous processes*** (*convection, boundary layer*)

- ***Not confined by isotropic, or other covariance limitations***

- ***Better use of data, since distant dynamically important correlations are allowed***

- ***Augmented control variable approach:***
    - *model errors, parameters, initial conditions can be all adjusted*
    - *cross-correlations between these components calculated*

MLEF APPLICATIONS WITH MESOSCALE AND LES RAMS MODELS

- ***Empirical model estimation using the mesoscale RAMS model***
  - POSTER session: "*A general ensemble-based approach to data assimilation, model error, and parameter estimation*" by Dusanka Zupanski and Milija Zupanski
  - MLEF can be used to *learn* about, and eventually *correct* model errors, by utilizing the information from observations

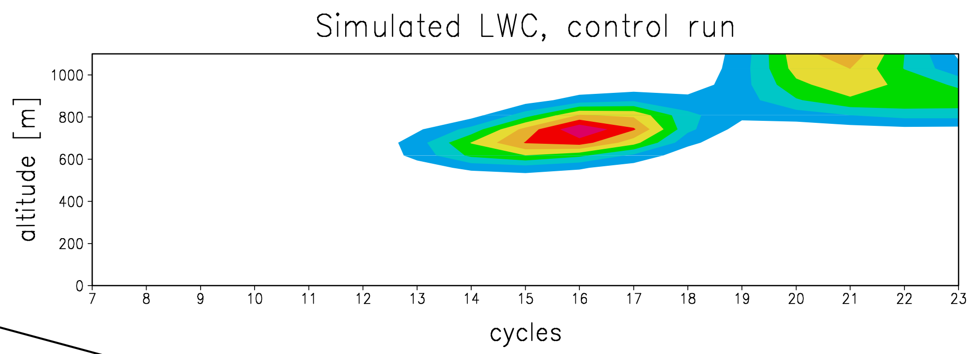
- ***Assimilation of real observations using the LES RAMS model***
  - Difficult-to-assimilate boundary-layer cloud observations
  - Important implications for the Army scale Meteorology
  - LES embedded into the mesoscale RAMS

ASSIMILATION OF REAL BOUNDARY-LAYER CLOUD OBSERVATIONS USING THE LES RAMS

- **23 2-h DA cycles: 18UTC 2 May 1998 – 00 UTC 5 May 1998**
  (Mixed phase Arctic boundary layer cloud at Sheba site)
- **Experiments initialized with typical *clean aerosol* concentrations**
- **May 4 was abnormal: *high* IFN and CCN above the inversion**
- **Δx= 50m, Δz_{max} = 30m (2d domain: 50col, 40lev), Δt=2s, Nens=48**
- **Sophisticated microphysics in RAMS/LES**
- **Control variables: Θ_il, u, v, w, N_x, R_x (8 species), IFN, CCN**
  (dim= 22 variables x 50 columns x 40 levels = **44,000**)
- **Radar/lidar/aircraft observations (retrievals) of IWP, LWP**

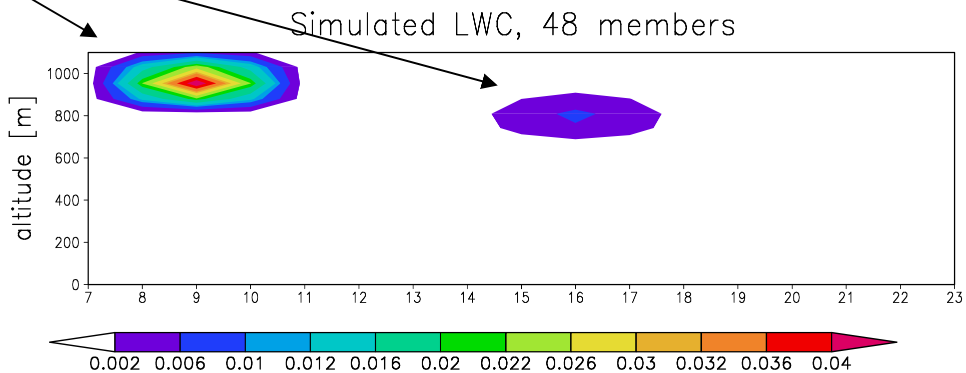
LIQUID WATER CONTENT ASSIMILATION

NO ASSIMILATION



Better timing of maxima

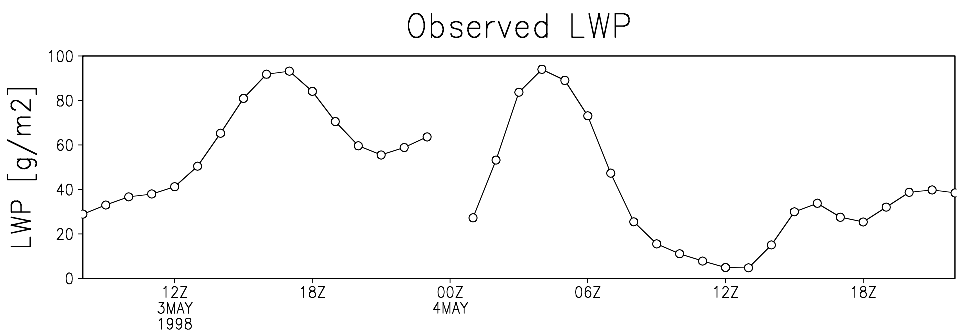
MLEF ASSIMILATION



Vertical *structure* of the analysis:

LWC

OBSERVATIONS



Vertically *integrated* observations:

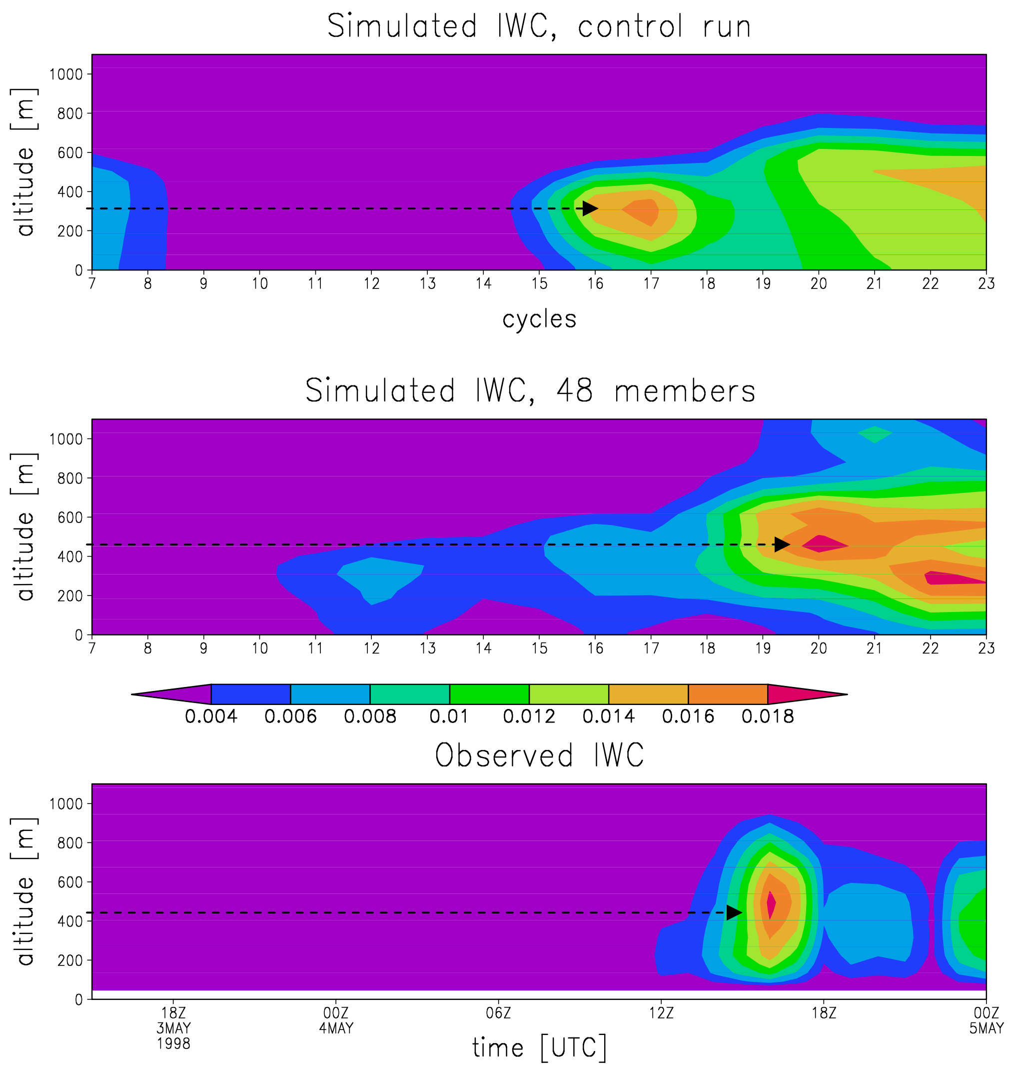
LWP

ICE WATER CONTENT ASSIMILATION

NO ASSIMILATION

MLEF ASSIMILATION

OBSERVATIONS

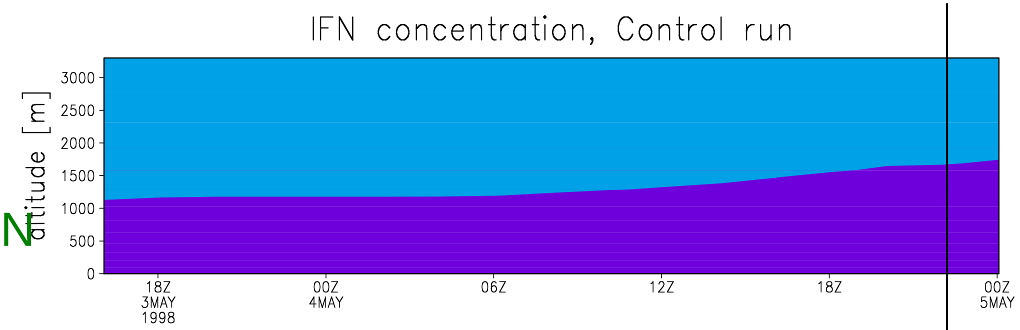


Improved timing and ***locations*** of the maxima

NO vertical structure is assimilated (IWP)

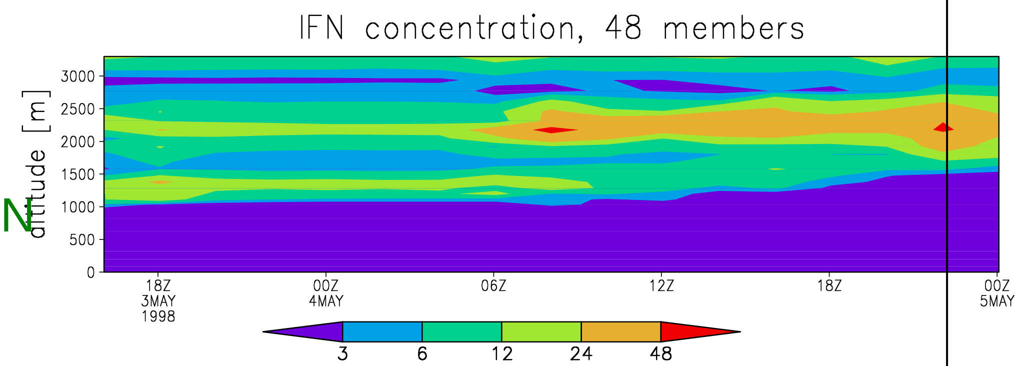
ICE FORMING NUCLEI (IFN) CONCENTRATION

NO ASSIMILATION



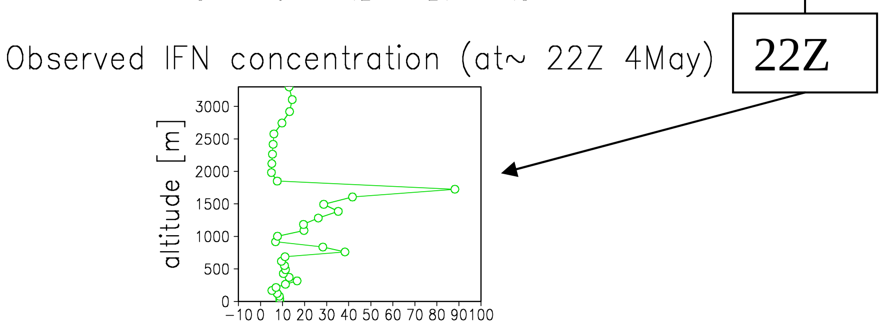
IFN↓ below inversion as cloud forms

MLEF ASSIMILATION



IFN↑ above the inversion, as observed

OBSERVATIONS



Independent observation

SUMMARY AND CONCLUSIONS

✓ ***MLEF is shown to efficiently estimate the empirical model parameters with mesoscale model (50 ensemble members)***

- Model deficiency estimation should be an important component of assimilation/prediction

✓ ***Completed MLEF experiments of real observations assimilation with LES model***

- Boundary-layer Arctic cloud
- Only 48 ensemble members
- Vertically integrated Radar/Lidar LWP, IWP observations
- Improved vertical structure of LWC, IWC, IFN, CCN

✓ ***MLEF does work with real observations, high resolution models, and can successfully estimate model error/empirical parameters***

FUTURE WORK

- ***Use Army microscale model(s)***
  - Univ. Purdue model
  - Univ. Wisconsin RAMS
  - WRF ?

- ***Use real observations***
  - Radiosonde, surface, lidar
  - Collected over the WSMR
  - Develop observation operators as needed

- ***Army HPC***
  - Perform real observation experiments with Army microscale models on Army HPC
  - Exploit parallel programming
  - Evaluate the computational efficiency of ensemble system