Ozone Data Assimilation with WRF-CHEM model

Milija Zupanski
Colorado State University

Seon Ki Park
Ewha Womans University

Overview:

- Goals and necessity of research
- Review of previous research
- Data and Methodology, Results, Future Plan
- Usefulness and Applicability of Research Results

Goals of research

- **Main goal of this multi-year project is to develop a chemical transport data assimilation (DA) system for assimilation of data from the environmental satellite in Korea.**

- Regional, high-resolution ensemble data assimilation will help maximize the utility of the future satellite chemistry data. Nonlinearities of satellite observation operators need to be taken into account.

- Develop and evaluate new *ensemble* system for chemistry data assimilation using simulated and real observations.

- First year: Develop basic interfaces between the regional coupled atmosphere-chemistry model and ensemble data assimilation. Assess the system in assimilation of simulated ozone observations, within an Observing System Simulation Experiments (OSSE) framework.

- Following years: Continue by adopting a general forward operator for assimilation of real satellite radiance from the environmental satellite in Korea.

Impact of atmospheric chemical constituents

- Trace gases and aerosols interact with climate and weather by their direct impact on radiation, and by indirect impacts on clouds.
- Relevant interactions between ozone, and weather and climate
- Implications on air quality and long-range pollution transport

- Need to improve understanding of atmospheric composition and to estimate distributions of surface sources and sinks of air pollution

Necessity of Research

- Atmospheric gases and aerosols have complex interactions that are impacted by natural and human sources (such as traffic, power generation, industry and agriculture). Important to use high-resolution coupled atmosphere-chemistry modeling in order to realistically simulate pollution transport.

- Major new information about atmospheric constituents comes from satellite measurements. New Korean Environmental satellite will provide high-resolution atmospheric chemistry measurements that will create a unique opportunity to improve our knowledge and eventually the prediction of atmospheric chemical constituents.

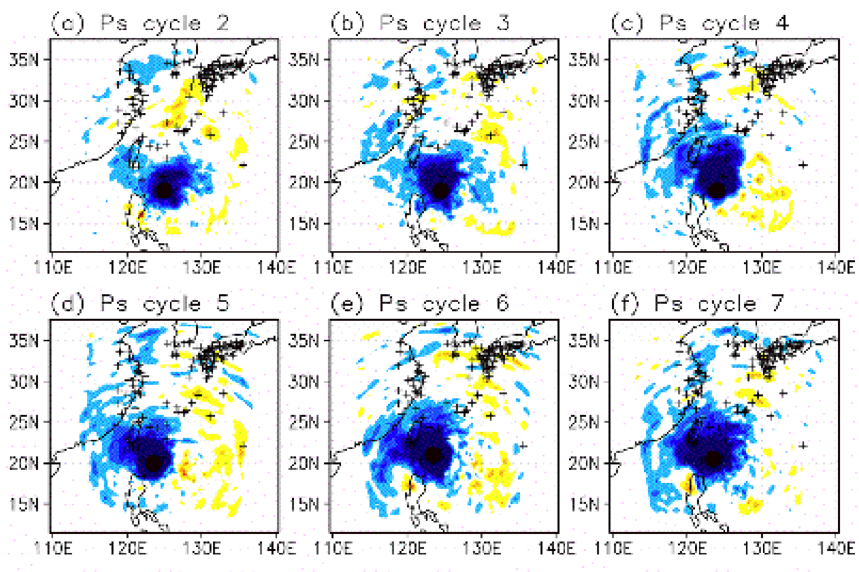
- Advanced data assimilation is required to address the challenging problem of utilizing the atmospheric chemistry information from new satellites.

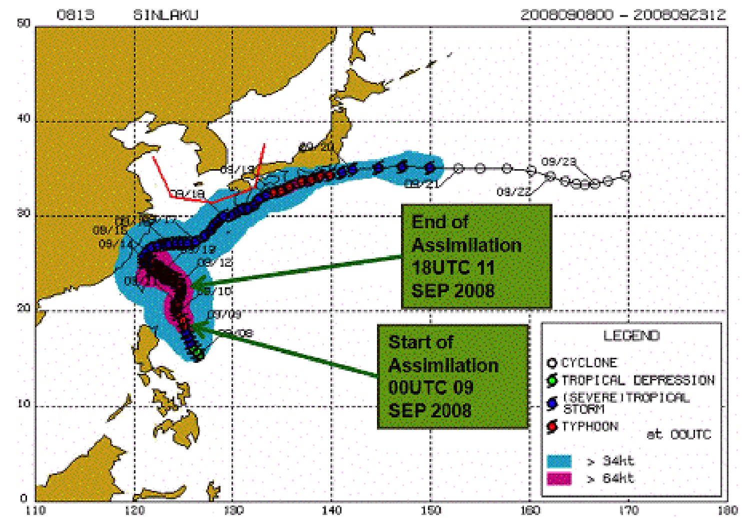
Review of previous research:

Data assimilation for Typhoon Sinlaku(2008) using MLEF+WRF

(*Kim et al. 2010, APJAS*)

Differences in surface pressure (hPa) between the experiments with and without data assimilation. Results for data assimilation cycles 2-7 are shown (from 0600 UTC 09 Sep 2008 to 1200 UTC 10 Sep 2008). Black circle indicates typhoon location.





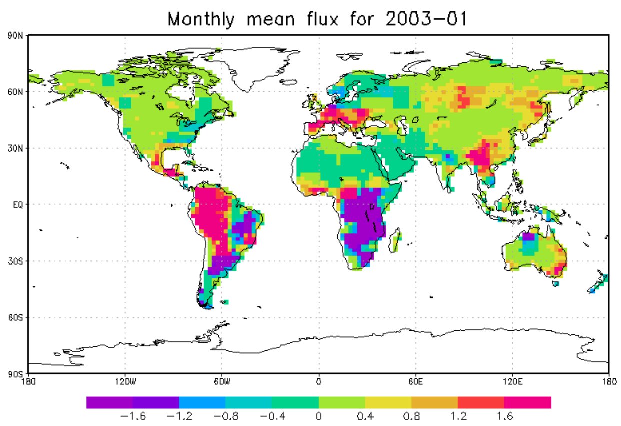
Typhoon is always located in the area where the pressure was reduced due to data assimilation (blue). Note switch in the blue/red dipole in cycle 4, when typhoon makes a turn towards east.

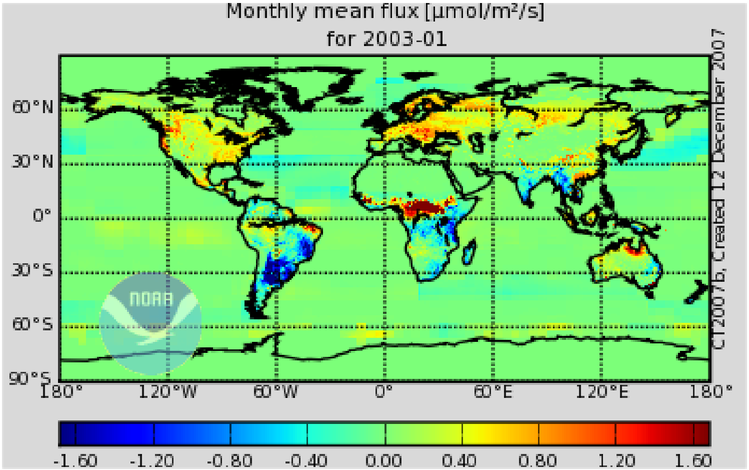
Review of previous research:

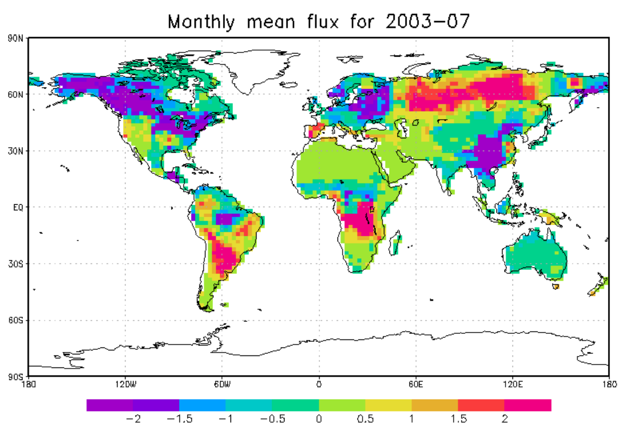
Carbon data assimilation - comparison of monthly mean fluxes (*Lokupitiya et al. 2008, JGR*)

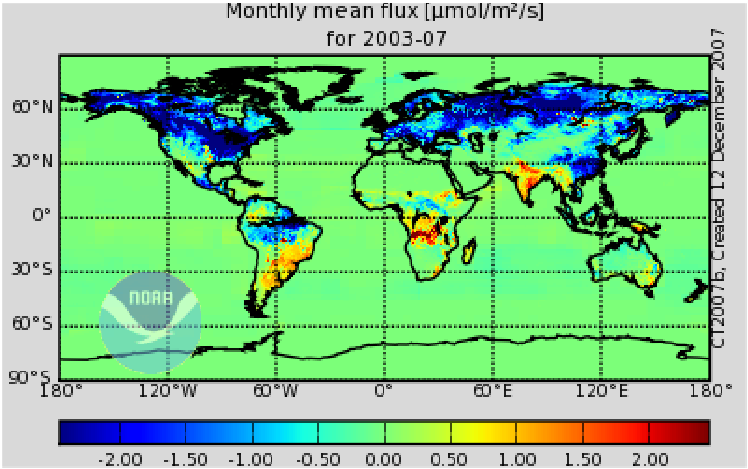
MLEF

Carbon Tracker









MLEF and Carbon Tracker (verification) produce comparable monthly fluxes

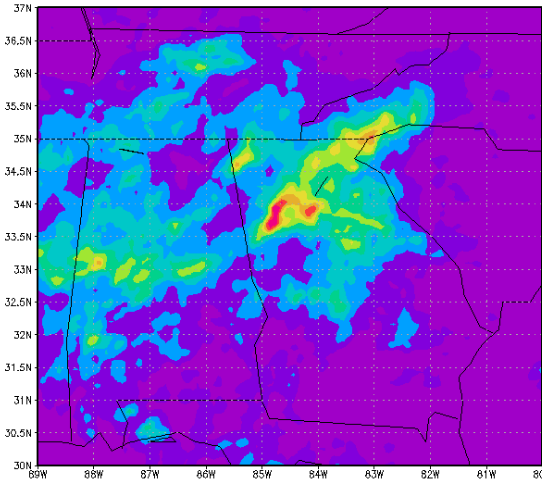
Review of previous research:

Assimilation of all-sky AMSR-E and TMI radiances (Zhang et al. 2011)

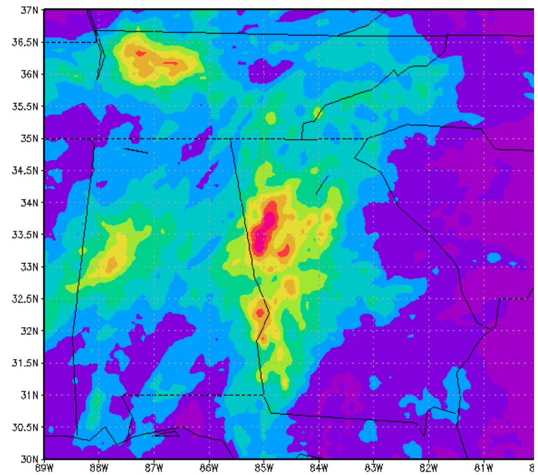
Surface precipitation short-term forecasts verification

Accumulated rain during 15-22 September 2009
in the Southeast flood region - 3-hour forecasts

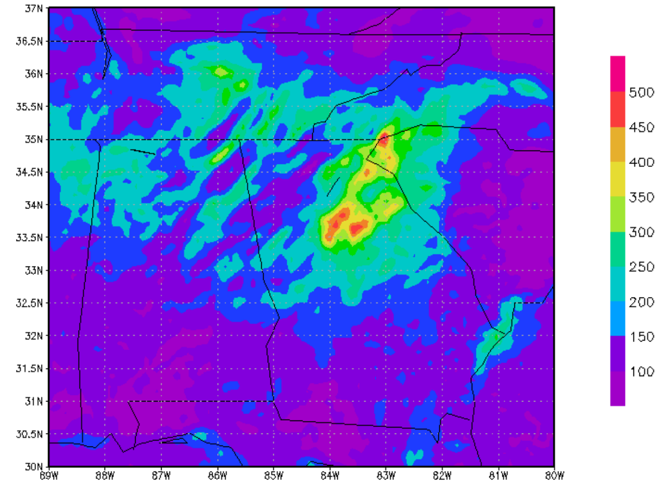
Ground-based Verification
(NOAA Stage IV data)



3DVAR, no AMSR-E,TMI
(WRF-GSI)



EDAS, with AMSR-E, TMI
(WRF-EDAS)



Assimilation of precipitation-affected radiance
improves short-term precipitation forecasts, in spatial pattern and intensity

Review of previous research:

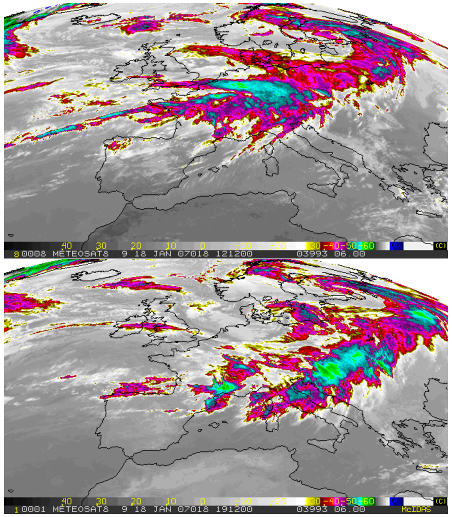
Assimilation of all-sky MSG SEVIRI IR radiances (Zupanski et al. 2011b)

- Kyrill: an extratropical wind storm in Europe in January 2007
- Data assimilation cycle is 1 h
- Control variable = (T, q, Qcloud, Qrain, Qice, Qsnow, Qgraupel)
- WRF model with 15 km horizontal resolution (300x300x40)
- All-sky radiative transfer based on CRTM and SHDOM
- Maximum Likelihood Ensemble Filter (MLEF)
- Ensemble size is 48 members

METEOSAT imagery 18 Jan 2008

12:12 UTC

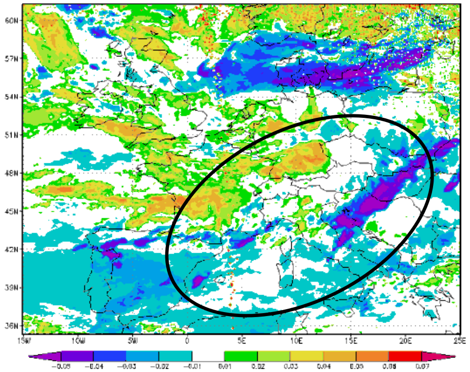
19:12 UTC



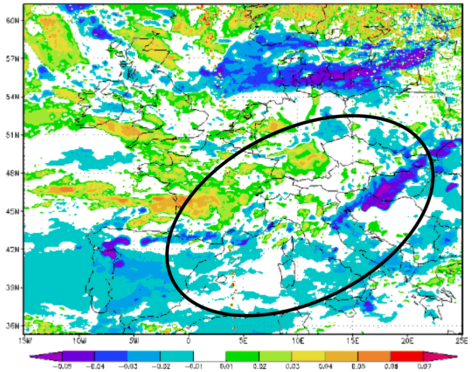
Fast-moving storm

MSG SEVIRI 10.80 μm (W m^{-2} sr^{-1} cm) valid 16Z 18 Jan 2007

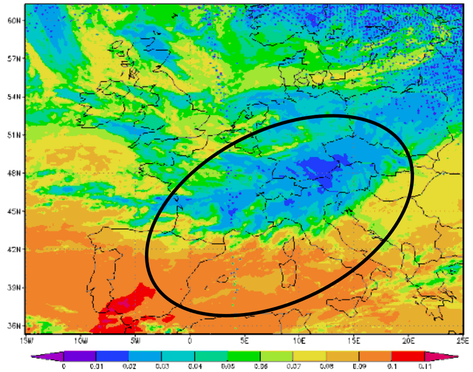
FG - OBS



ANL - OBS



OBS (10.80 μm)



Reduction of errors due to data assimilation of MSG SEVIRI radiances

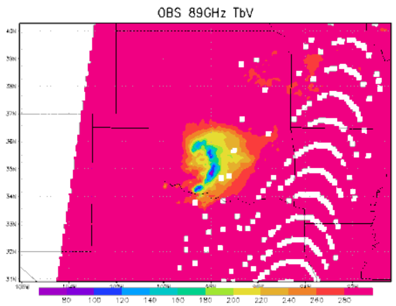
Relevance to this project:

All-sky radiance observation information content

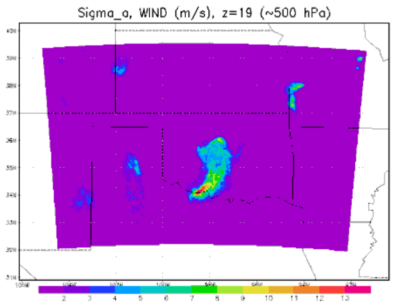
MW radiances: AMSR-E data assimilation (Erin, 2007) - WRF 3km

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

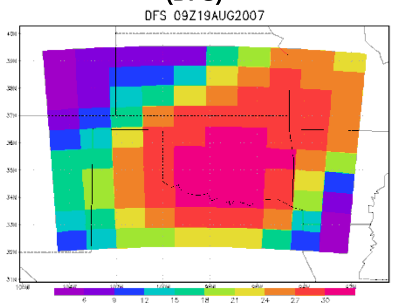
OBS 89v GHz Tb



Wind analysis uncertainty (500 hPa)



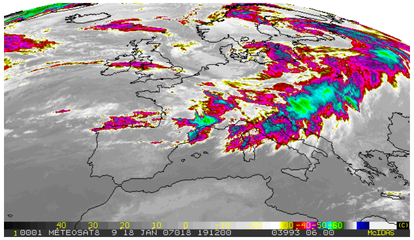
Degrees of Freedom for Signal (DFS)



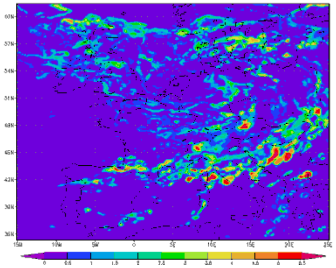
IR radiances: Assimilation of synthetic GOES-R ABI (10.35 mm) all-sky radiances (Kyrill, 2007) - WRF 15km

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

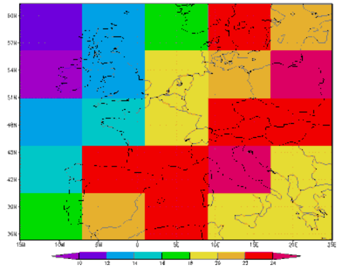
METEOSAT Imagery valid at 19:12 UTC 18 Jan 2007



Cloud ice analysis uncertainy



Degrees of Freedom for Signal (DFS)



MLEF is capable of extracting maximum information from MW and IR all-sky radiances

Overview of previous results

- MLEF ensemble data assimilation has been successfully applied with Weather Research and Forecast (WRF) atmospheric model.

- Applications to :
  - Typhoon tracking and intensity
  - Severe weather
  - Wind storms
  - Extreme precipitation and flooding
  - Carbon tracking

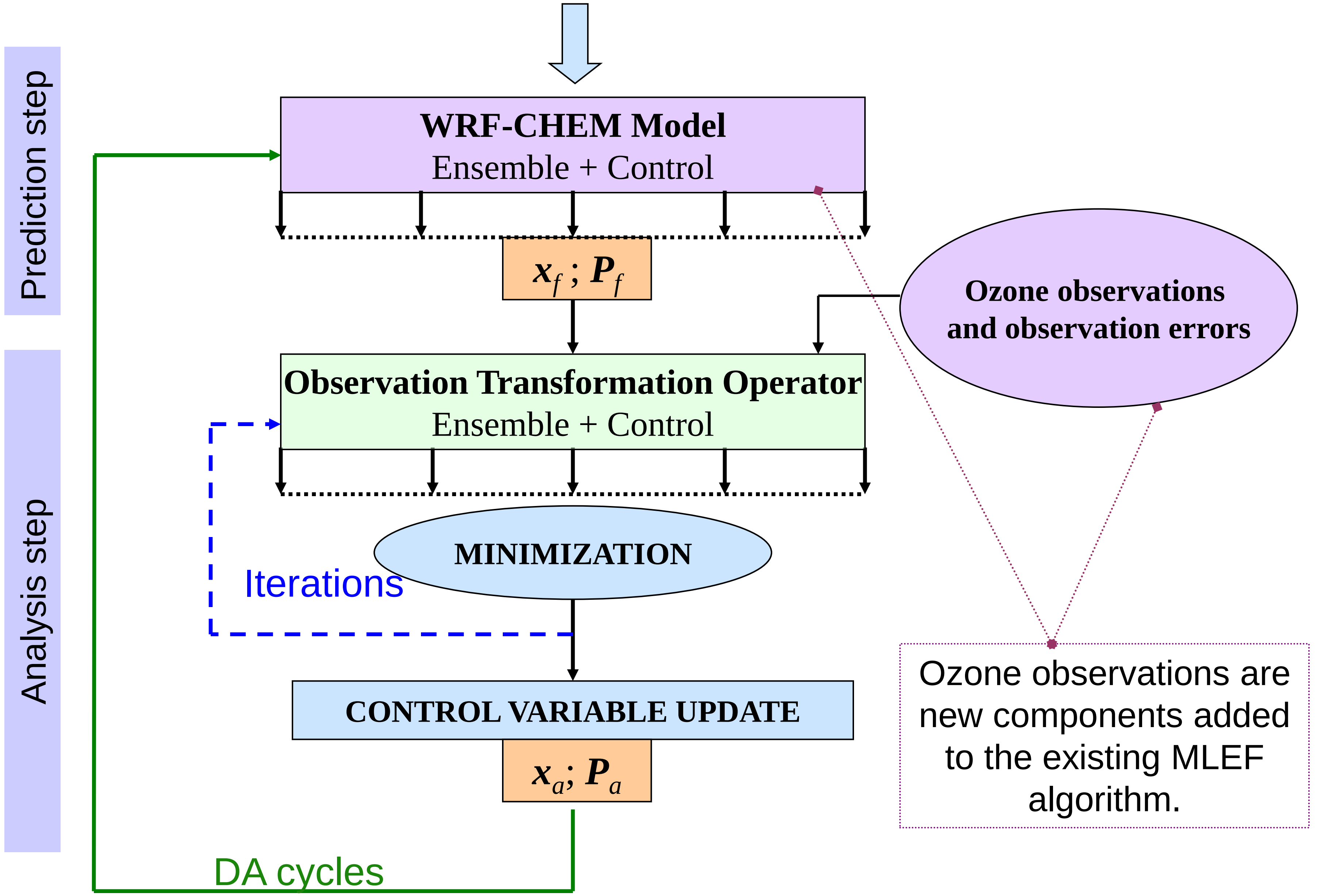
- Challenging nonlinear data assimilation of all-sky infrared and microwave satellite radiances has been successfully accomplished with MLEF.

Strategy and Methodology

- Develop a chemistry data assimilation system based on ensemble data assimilation and a high-resolution regional atmosphere-chemistry model.
- Begin with assimilation of simulated observations to prepare for assimilation of real satellite observations from the future Korean environmental satellite

- *Coupled modeling system*: Weather Research and Forecasting Chemistry (WRF-CHEM) model.
- *Ensemble data assimilation system*: Maximum Likelihood Ensemble Filter (MLEF - Zupanski 2005; Zupanski and Zupanski 2006; Zupanski et al. 2008).
- One of the unique characteristics of the MLEF is that it includes an unconstrained iterative minimization of the cost function with implicit Hessian preconditioning. This system can address highly nonlinear observation operators used for satellite chemistry observations.

MLEF-WRF-CHEM Flow Diagram



MLEF Analysis: Generalization of Kalman Filter to include nonlinear observation operators

Control theory viewpoint:

In standard KF, the analysis is obtained by minimizing a quadratic cost function (i.e. with linear observation operators)

Generalize KF to include *nonlinear observation* operators:

- **Minimize nonlinear (i.e. non-quadratic) cost function**
- **Use best applicable minimization method**
- **Build data assimilation around minimization**

$$J(x) = \frac{1}{2}\left(x - x^f\right)^T P_f^{-1}\left(x - x^f\right) + \frac{1}{2}\left(y - \mathcal{K}(x)\right)^T R^{-1}\left(y - \mathcal{K}(x)\right)$$

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

MLEF Analysis: Generalization of Kalman Filter to include nonlinear model operators

$$P_f^{1/2} = MP_a^{1/2} \quad \Rightarrow \quad \begin{bmatrix} p_1^f & p_2^f & \cdots & p_n^f \end{bmatrix} = \begin{bmatrix} Mp_1^a & Mp_2^a & \cdots & Mp_n^a \end{bmatrix}$$

In KF, the forecast error column is a forecast of the analysis error column

Since $\{p_1^a \quad p_2^a \quad \cdots \quad p_n^a\}$ **spans the analysis uncertainty subspace, one can say that uncertainty is transported in time by a (linear) model** M

Generalize KF to include nonlinear forecast model:

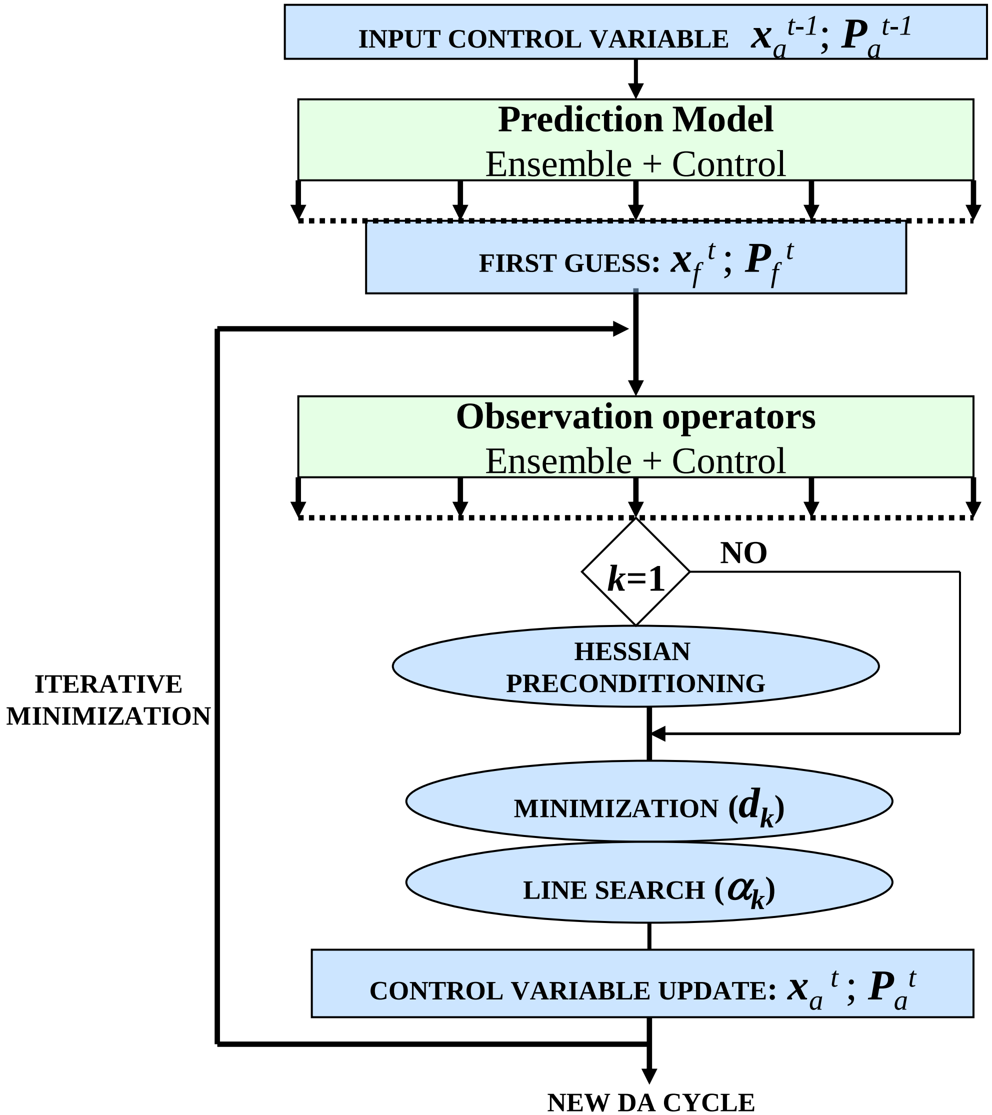
Transport uncertainty in time by a *nonlinear* model $\mathcal{M}$ (one span vector at a time)

$$x^f = \mathcal{M}(x^a) \qquad x_i^f = \mathcal{M}(x^a + p_i^a)$$

$$p_i^f = x_i^f - x^f = \mathcal{M}(x^a + p_i^a) - \mathcal{M}(x^a)$$

Each uncertainty column vector is a member of an "ensemble" (i.e. span)

Ensemble data assimilation based on Control Theory: General formulation of the MLEF



- A hybrid between EnKF and variational data assimilation
- Full-rank or reduced-rank
- Deterministic first guess forecast
- Analysis is the maximum of a posterior pdf
- Nonlinear analysis solution by an iterative minimization
- Improved minimization efficiency by an implicit Hessian preconditioning

MLEF UNIQUE AND SPECIAL FEATURES

- **Fully nonlinear ensemble data assimilation / forecasting system**
- **Analysis obtained by an iterative minimization of nonlinear cost function**
- **Advanced Hessian preconditioning using a complete information from prediction model and observations**
- **Standard unconstrained minimization algorithms with Gateaux differential substituted by its finite-difference representation are used**
- **Reduced growth of Local Lyapunov Vectors (LLV) using observations (Carrassi et al. 2009)**
- **Object-oriented programming (flexible for adding/deleting modules)**
- **Parallel computation capability (MPI)**

MLEF Equations

Forecast:

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

$$P_f^{1/2} = \begin{bmatrix} p_1^f & p_2^f & \dots & p_n^f \end{bmatrix} \qquad p_i^f = \mathcal{M}(x^a + p_i^a) - \mathcal{M}(x^a)$$

Change of variable (Hessian preconditioning):

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

$$Z(x) = \begin{bmatrix} z_1(x) & z_2(x) & \dots & z_n(x) \end{bmatrix} \qquad z_i(x) = R^{-1/2}\left[\mathcal{K}(x + p_i^f) - \mathcal{K}(x)\right]$$

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

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 in MLEF

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

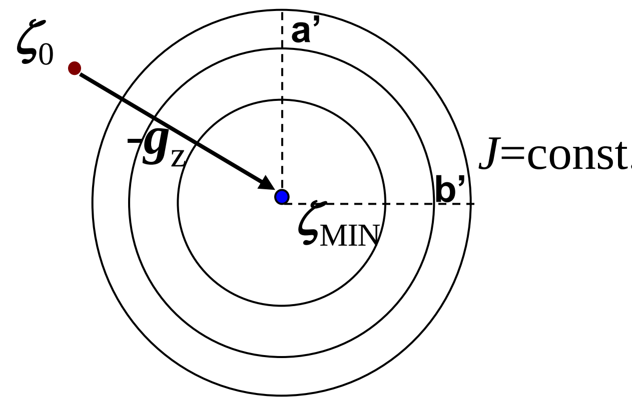
Hessian matrix: $$Q = \frac{\partial^2 J}{\partial x^2} = P_f^{-1} + K^T R^{-1} K$$

$$\boxed{Q = EE^T}$$

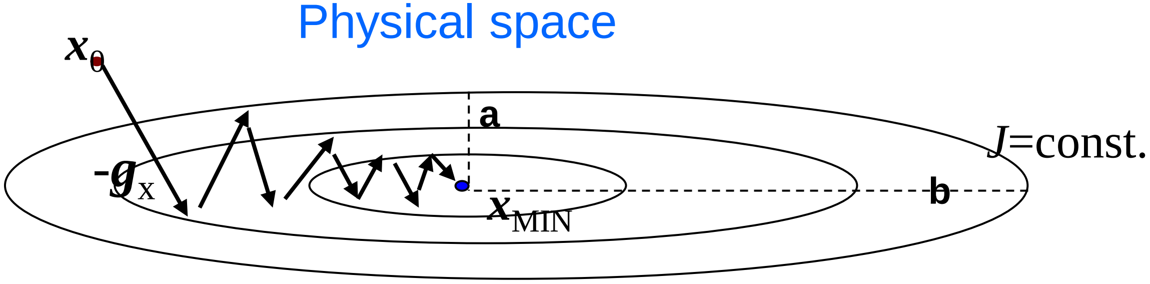
Optimal Hessian preconditioning: $$x - x_f = E^{-T}\zeta$$ $$E^{-T} = P_f^{1/2}\left(I + Z(x^f)^T Z(x^f)\right)^{-1/2}$$

$$\boxed{Q_\zeta = E^{-1}QE^{-T} = E^{-1}EE^T E^{-T} = I}$$

Preconditioning space

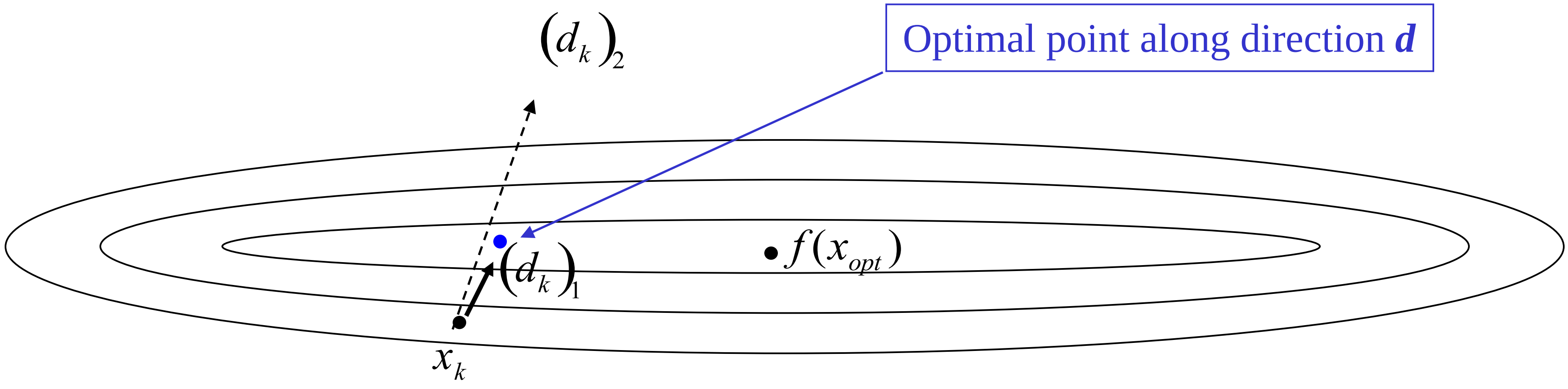


Physical space



Hessian preconditioning is critical for efficient minimization.

Relevance of line search



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

An intuitive approach is α=1 (Newton's method, EnKF). **However ...**

$J(x_k + (d_k)_1)$ may not be the optimal (minimum) value

$J(x_k + (d_k)_2)$ could be worse then the starting point

Therefore, ***d*** is not sufficient: **Need a size parameter α for nonlinear cost function**.

MLEF employs an advanced line search based on satisfying the Wolfe conditions.

Finite-difference representation of Gateaux differentials

Standard Taylor expansion of cost function using Gateaux differentials:

$$J(w+\delta w) = J(w) + DJ(w;\delta w) + \frac{1}{2!}D^2 J(w;\delta w) + \ldots + \frac{1}{k!}D^k J(w;\delta w) + O(\|\delta w\|^{k+1})$$

$$DJ(w;\delta w) = \left(\frac{\partial J}{\partial w}\right)\delta w = \lim_{t\to 0}\frac{J(x+t\delta w) - J(x)}{t}$$

Define finite-difference (FD) representation of Gateaux differentials:

$$D_{FD}J(w;\delta w;t) = \frac{J(x+t\delta w) - J(x)}{t} \qquad (0 < t \le 1)$$

$$D_{FD}J(w;\delta w;t) \xrightarrow[t\to 0]{} DJ(w;\delta w)$$

Expansion of cost function using finite differences with $t=1$ (MLEF):

$$J(w+\delta w) = J(w) + D_{FD}J(w;\delta w;1) + \frac{1}{2!}D^2_{FD}J(w;\delta w;1)$$

- No additional nonlinear terms for $t=1$! All components included.

- The value of t defines the degree of nonlinearity in FD differentials

Finite-difference representation of Gateaux differentials in MLEF minimization

Use as first derivative $D_{FD}J$

$$D_{FD}J = w - Z(x)^T R^{-1/2}\left[y - \mathcal{K}(x)\right]$$

Use as second derivative $D_{FD}^2 J$

$$D_{FD}^2 J = I + Z(x)^T Z(x)$$

Note that this is the same form as if using true G-differentials

$$Z(x) = \left[z_1(x) \quad z_2(x) \quad \cdots \quad z_n(x)\right]$$

Standard

$$Z(x) = R^{-1/2} K \delta x$$

MLEF

$$z_i(x) = R^{-1/2}\left[\mathcal{K}(x + \delta x_i) - \mathcal{K}(x)\right]$$

Robustness of nonlinear CG and BFGS algorithms improved with FD representation !

(Zupanski et al. 2008)

Results/Accomplishments

Mid-term accomplishments:

- The latest version of WRF-CHEM (V3.3) installed on Ewha computer.
- The latest version of WRF Preprocessing System (WPS) installed on Ewha computer.
- MLEF algorithm installed on Ewha computer.
- WRF-CHEM, WPS and MLEF are compiled and interfaced.

Remaining tasks (until the end of the project):

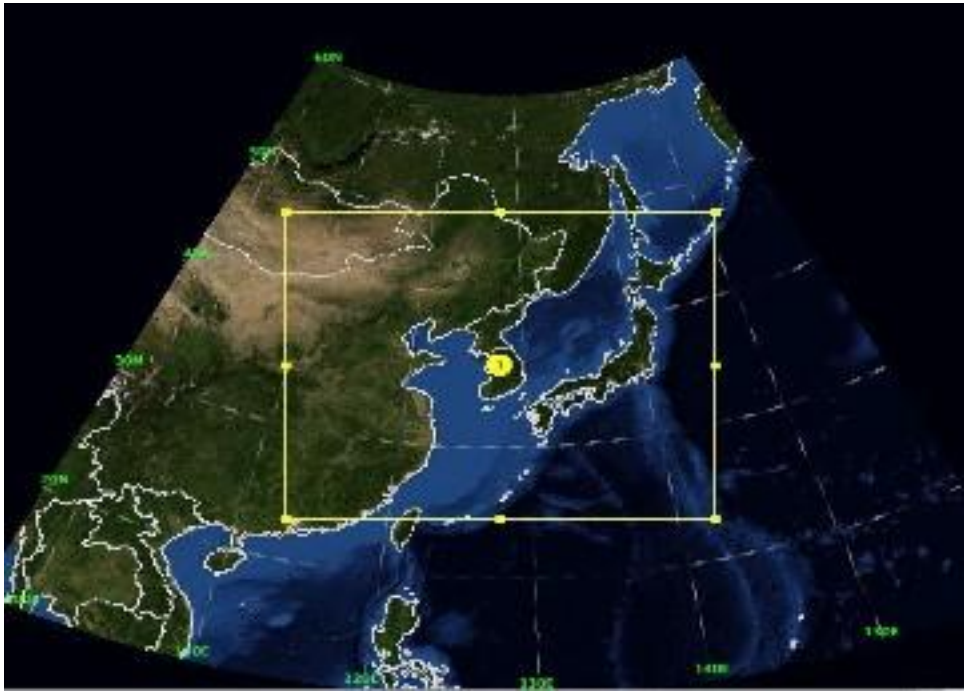
- Create simulated O_3 observations for OSSE using WRF-CHEM.
- Evaluate the performance of the MLEF-WRF-CHEM algorithm in assimilation of simulated ozone data.

- All tasks are on track

- Accomplishment of all tasks is expected by the end of the project

Experimental setup

- WRF-CHEM model centered over Korea
- WRF-CHEM model resolution 27 km / 28 layers (131x111x28)
- Automatic processing of model files (lateral BC, IC) from NCEP global model (GFS) at 6-hour intervals
- Simulated observations of ozone at model grid points at 6-hour intervals
- Only one minimization iterations since observation operator is an identity



WRF-CHEM model domain

Future plans (long term goals)

- Develop/adopt an observation operator for aerosol/ozone. This may include satellite and/or other observations.

- Assimilate real satellite observations from the Korean environmental satellite. Demonstrate the system’s capability to process such observations.

- Evaluate the MLEF-WRF-CHEM system in high-resolution chemical transport DA. Focus on the impact of the Korean environmental satellite observations.

Usefulness and Applicability of Research Results

- Maximize the utility of high-resolution observations from Korean environmental satellite using advanced data assimilation and prediction system

- Prepare for satellite launching by developing the components required for an efficient and thorough processing of environment satellite observations

- Improve our knowledge of geographical coverage and concentration of atmospheric chemistry constituents over Korean peninsula and surrounding areas (information content analysis)

- Useful for future tracking of air-pollution sources and sinks

References cited

Carrassi, A., S. Vannitsem, D. Zupanski, and M. Zupanski, 2009: The Maximum Likelihood Ensemble Filter performances in chaotic systems. *Tellus A*, **61**, 587-600.

Kim, H.-H., S.-K. Park, D. Zupanski, and M. Zupanski, 2010: Uncertainty Analysis Using the WRF Maximum Likelihood Ensemble Filter System and Comparison with Dropwindsonde Observations in Typhoon Sinlaku (2008). *Asia-Pacific J. Atmos. Sci.*, **46**, 317-325.

Lokupitiya, R. S., D. Zupanski, A. S. Denning, S. R. Kawa, K. R. Gurney, and M. Zupanski, 2008: Estimation of Global CO_2 Fluxes at Regional Scale Using the Maximum Likelihood Ensemble Filter. *J. Geophys. Res.*, **113**, D20110, doi:10.1029/2007JD009679.

Zupanski D. and M. Zupanski, 2006: Model error estimation employing an ensemble data assimilation approach. *Mon. Wea. Rev.*, **134**, 1337-1354.

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

Zupanski D., M. Zupanski, L. D. Grasso, R. Brummer, I. Jankov, D. Lindsey and M. Sengupta, and M. DeMaria, 2011b: Assimilating synthetic GOES-R radiances in cloudy conditions using an ensemble-based method. *Int. J. Remote Sensing*, doi:10.1080/01431161.2011.572094.

Zhang, S. Q., M. Zupanski, and A. Y. Hou, 2011: Assimilation of precipitation affected radiances in a cloud-resolving WRF ensemble data assimilation system. To be submitted to *Mon. Wea. Rev*.

Zupanski, M., 2005: Maximum likelihood ensemble filter: Theoretical aspects. *Mon. Wea. Rev.*, **133**, 1710–1726.

Zupanski, M., I. M. Navon, and D. Zupanski, 2008: The maximum likelihood ensemble filter as a non-differentiable minimization algorithm. *Quart. J. Roy. Meteor. Soc*. **134**, 1039-1050.

More references and information about the MLEF can be found at
http://www.cira.colostate.edu/projects/ensemble