WRFDA Overview

Zhiquan (Jake) Liu

NCAR/MMM

WRFDA is a Data Assimilation system built within the WRF software framework, used for application in both research and operational environments….

Outline

- What is data assimilation
  - Scalar case
  - Two state variables case
  - General case
- Introduction to WRF Data Assimilation

Outline

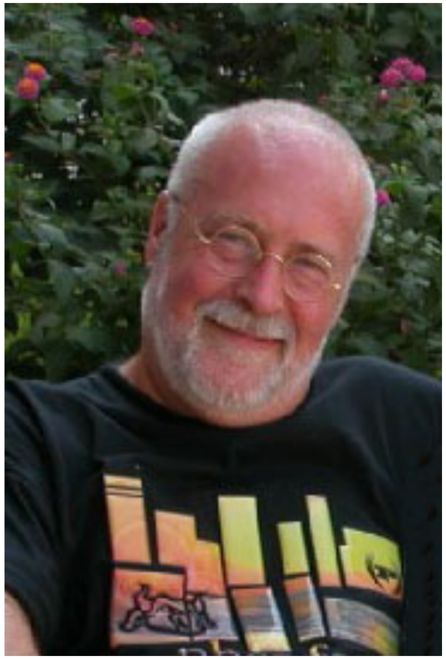
- What is data assimilation
  - Scalar case
  - Two state variables case
  - General case
- Introduction to WRF Data Assimilation

What is data assimilation?

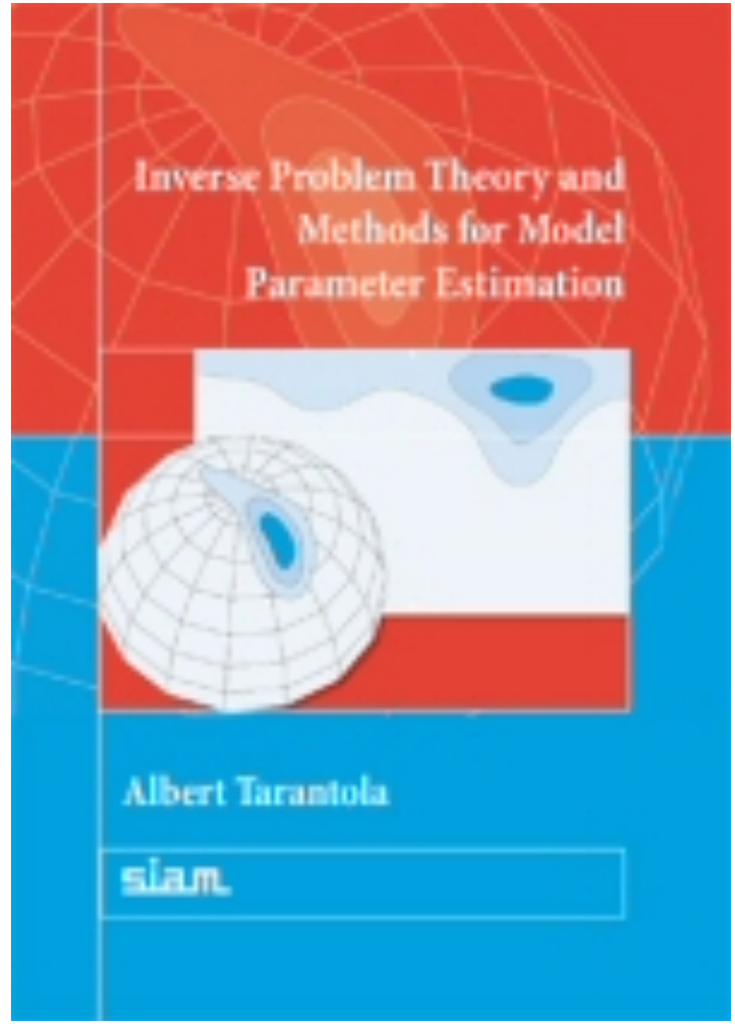
- A statistical method to obtain the best estimate of state variables, based upon
  - Probability theory, Bayes theorem
  - Optimal control, optimal estimation theory
  - Inverse problem theory

- In the atmospheric sciences, DA involves combining a model and observations, along with their respective errors characterization, to produce an *analysis* that can initialize a numerical weather prediction model (i.e., WRF)

A freely available book



Albert Tarantola



Scalar Case

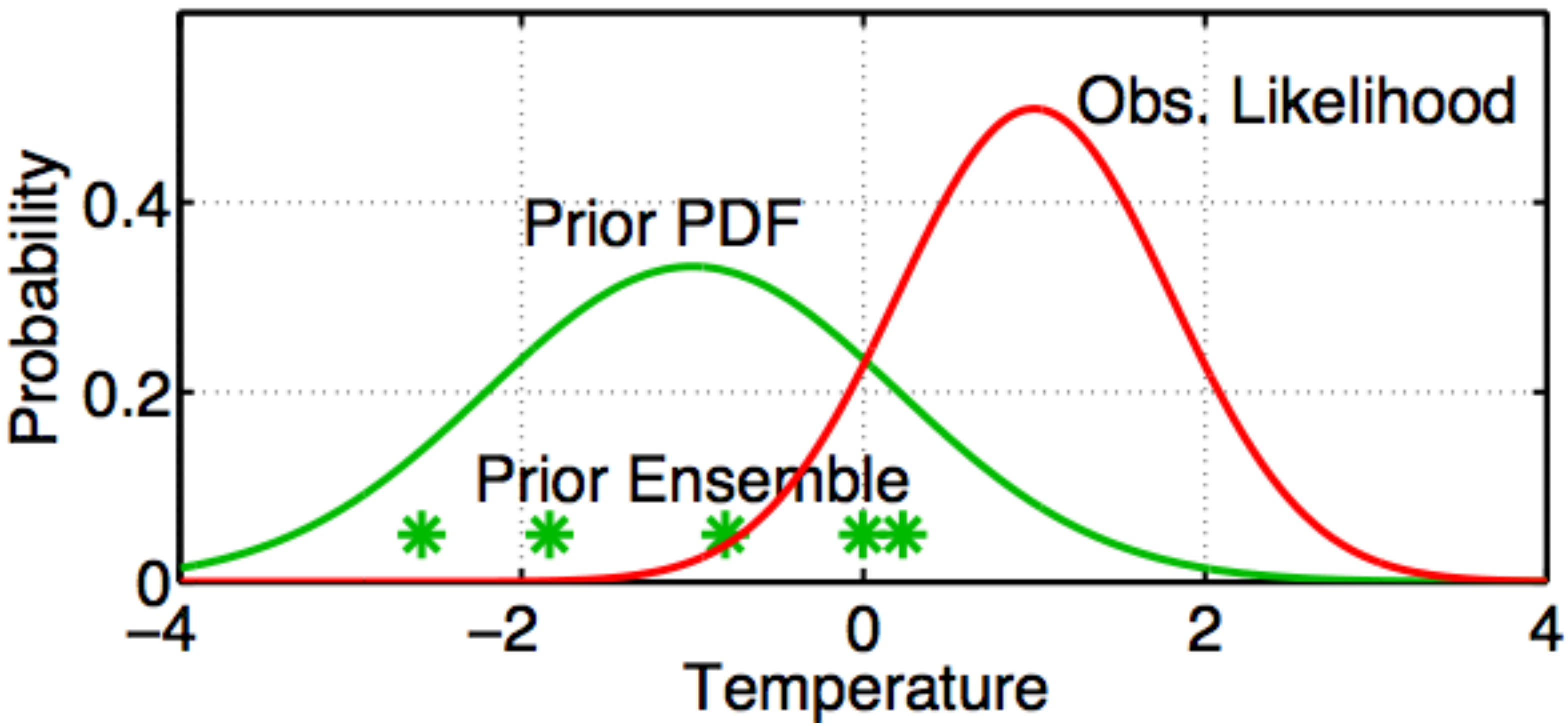
- State variable to estimate "x", e.g., consider today's temperature of Boulder at 12 UTC.
- Now we have a "background" (or "prior") information x_b of x, which is from a 6-h GFS or WRF forecast initiated from 06 UTC today.
- We also have an observation y of x at a surface station in Boulder, measured at 12 UTC.
- What is the best estimate (analysis) x_a of x?

Scalar Case

- We can simply average them: $x_a = \frac{1}{2}(x_b + y)$
  - This implies we trust equally the background and observation.

- But what if their accuracy is different and we have some estimation of their errors
  - e.g., for background, we have statistics (e.g., mean and variance) of $x_b - y$ from the past
  - For observation, we have instrument error information from manufacturer

Assume we got Gaussian error statistics for both background and observation



background error: $N(0, \sigma_b)$
observation error: $N(0, \sigma_o)$

Scalar Case

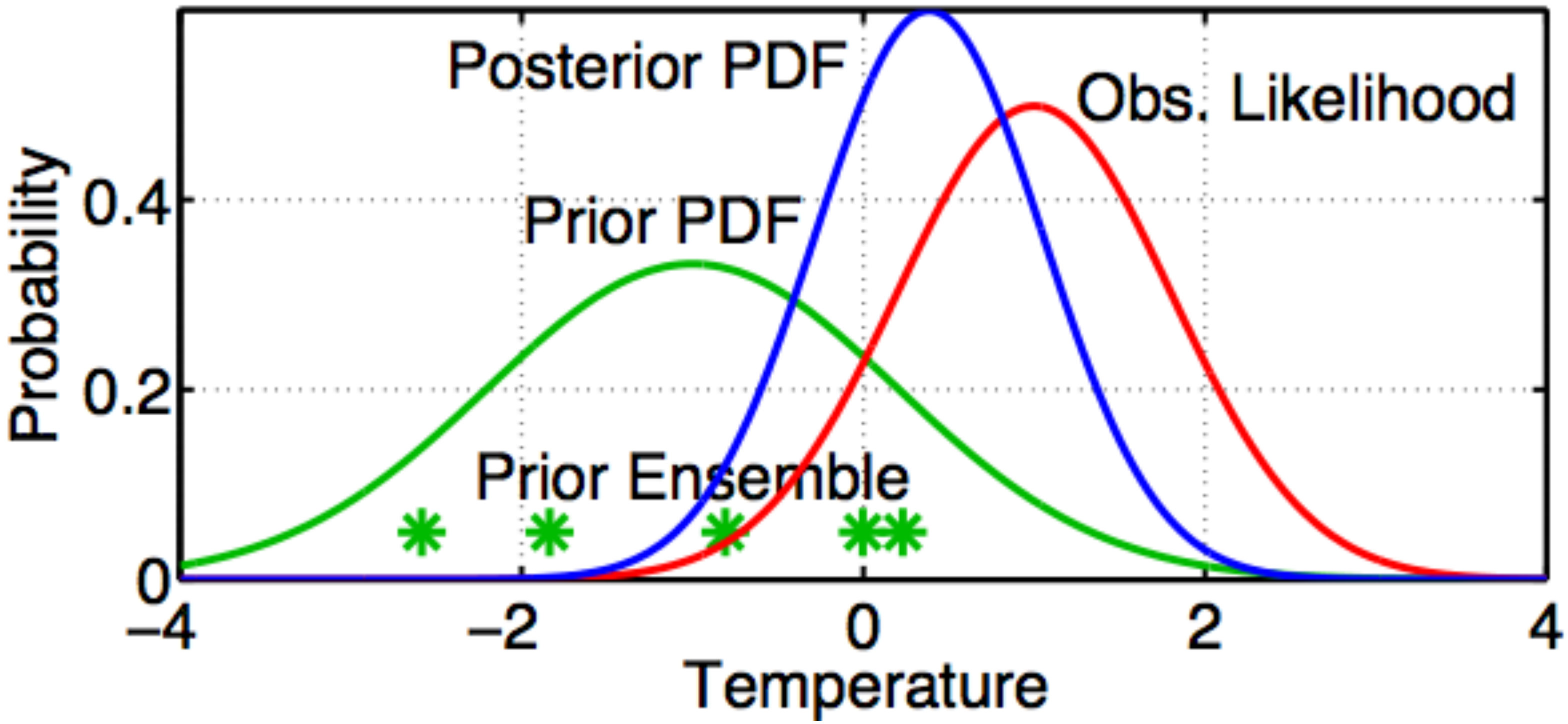
- Then we can do a weighted mean: $x_a = ax_b + by$ in a least square sense, i.e.,
  - Minimize $J(x) = \frac{1}{2}\frac{(x-x_b)^2}{\sigma_b^2} + \frac{1}{2}\frac{(x-y)^2}{\sigma_o^2}$
  - Requires $\frac{dJ(x)}{dx} = \frac{(x-x_b)}{\sigma_b^2} + \frac{(x-y)}{\sigma_o^2} = 0$
  - Then we can easily get

$$x_a = \frac{\sigma_o^2}{\sigma_b^2+\sigma_o^2}x_b + \frac{\sigma_b^2}{\sigma_b^2+\sigma_o^2}y$$

  - We can also write in the form of **analysis increment**

$$x_a - x_b = \frac{\sigma_b^2}{\sigma_b^2+\sigma_o^2}(y - x_b)$$

Innovation



- Analysis (posterior) error PDF: $N(0, \sigma_a^2)$

$$\frac{1}{\sigma_a^2} = \frac{1}{\sigma_b^2} + \frac{1}{\sigma_o^2}$$

Precision: inverse of error variance

So σ_a^2 is always smaller than σ_b^2 and σ_o^2 (only in a statistical sense, but for a sin gle realization, analysis is not necessar ily more accurate than background).

Two state variables case

- Consider two state variables to estimate: Boulder and Denver's temperatures x_1 and x_2 at 12 UTC today.

- Background from 6-h forecast: x_1^b and x_2^b
  - and their **error covariance** with **correlation** *c*, ***which is extremely important in data assimilation*** (see lecture by Rizvi)

$$\mathbf{B} = \begin{bmatrix} \sigma_1^2 & c\sigma_1\sigma_2 \\ c\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} 1 & c \\ c & 1 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$$

- We only have an observation y_1 at a Boulder station and its error variance σ_o^2

Analysis increment for two variables

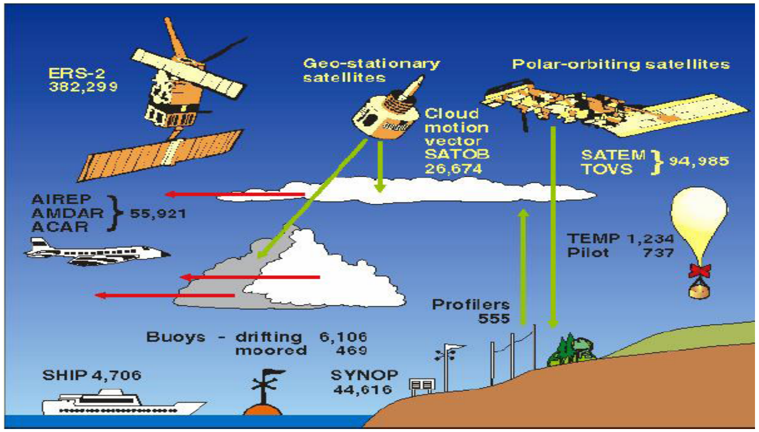
$$x_1^a - x_1^b = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_o^2}(y_1 - x_1^b)$$ **Boulder**

$$x_2^a - x_2^b = \frac{c\sigma_1\sigma_2}{\sigma_1^2 + \sigma_o^2}(y_1 - x_1^b)$$ **Denver**

Unobserved variable x_2 gets updated through the error correlation *c* in the background error covariance.

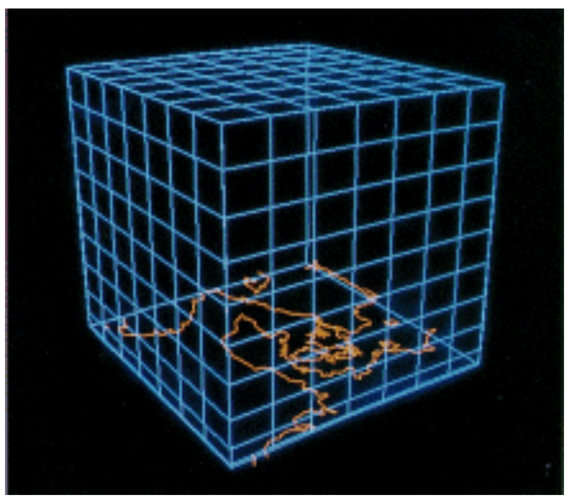
This correlation can be correlation between two locations (spatial), two variables (multivariate), or two times (temporal).

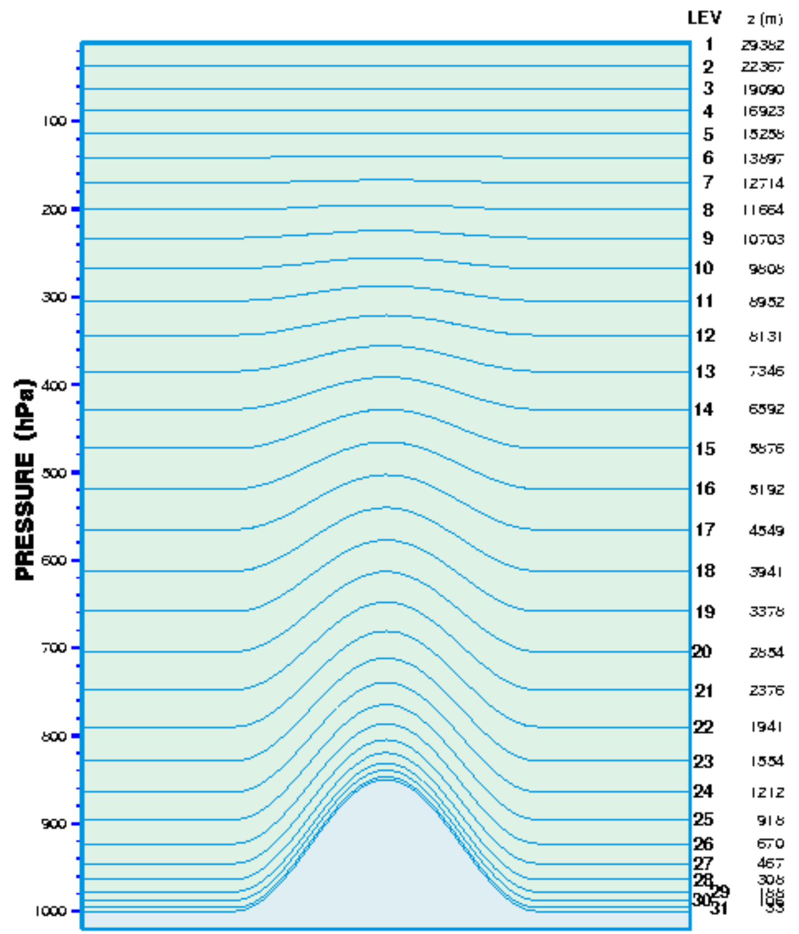
General NWP Case



Observations
y^o, ~10^5-10^6

Model state
x, ~10^7





Vertical resolution of the DMI-HIRLAM system

General Case: vector and matrix notation

state vector

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

observation vector

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

background error covariance

$$\mathbf{B} = \begin{bmatrix} \sigma_1^2 & c_{12}\sigma_1\sigma_2 & \dots & \dots \\ c_{12}\sigma_1\sigma_2 & \sigma_2^2 & \dots & \dots \\ \dots & \dots & \ddots & \dots \\ \dots & \dots & \dots & \sigma_m^2 \end{bmatrix}$$

$$J(x) = \frac{1}{2}(x - x^b)^T \mathbf{B}^{-1}(x - x^b) + \frac{1}{2}[\mathbf{H}x - y]^T \mathbf{R}^{-1}[\mathbf{H}x - y]$$

Observation error covariance

$$\mathbf{R} = \begin{bmatrix} \sigma_{o1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{o2}^2 & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & \dots & \dots & \sigma_{on}^2 \end{bmatrix}$$

H [n x m] maps x to y space, e. g., interpolation.
Terminology in DA: observation operator

Minimize J is equivalent to maximize a Gaussian PDF

Constant * $e^{-J(x)}$

General case: analytical solution

Again, minimize J requires its gradient (a vector) with respect to x equal to zero:

$$\nabla J_{\mathbf{x}}(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_{\mathbf{b}}) - \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{y} - \mathbf{H}\mathbf{x}] = 0$$

Transpose of H: adjoint operator

This leads to analytical solution for the analysis increment:

$$x^a - x^b = \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}[y - \mathbf{H}x^b]$$

Analog to 2 variables case: $x_2^a - x_2^b = \dfrac{c\sigma_1\sigma_2}{\sigma_1^2 + \sigma_o^2}(y_1 - x_1^b)$

$\mathbf{HBH}^{\mathrm{T}}$: projection of background error covariance in observation space

$\mathbf{BH}^{\mathrm{T}}$: projection of background error covariance in background-observation space

Precision of Analysis

$$\mathbf{A}^{-1} = \mathbf{B}^{-1} + \mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{H} = (\mathbf{I} - \mathbf{KH})\mathbf{B}$$

= Hessian: the second order derivative of cost function

Generalization of scalar case $\frac{1}{\sigma_a^2} = \frac{1}{\sigma_b^2} + \frac{1}{\sigma_o^2}$

With

$$\mathbf{K} = \mathbf{BH}^{T}(\mathbf{HBH}^{T} + \mathbf{R})^{-1}$$

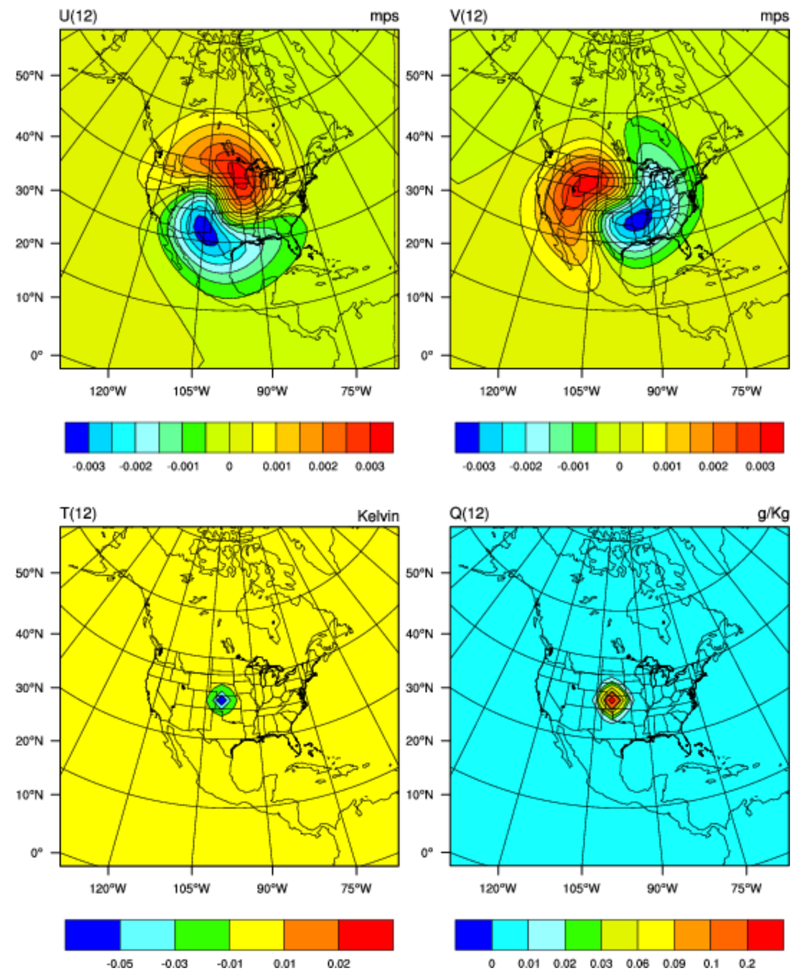
called Kalman gain matrix

Analysis increment with a single humidity observation

$$x^a - x^b = \mathbf{BH}^\mathrm{T}(\mathbf{HBH}^\mathrm{T} + \mathbf{R})^{-1}[y - \mathbf{H}x^b]$$

$$\downarrow$$

$$x_l^a - x_l^b = \frac{c_{lk}\sigma_l\sigma_k}{\sigma_k^2 + \sigma_{ok}^2}(y_k - x_k^b)$$



It is generalization of previous two variables case:

$$x_1^a - x_1^b = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_o^2}(y_1 - x_1^b)$$

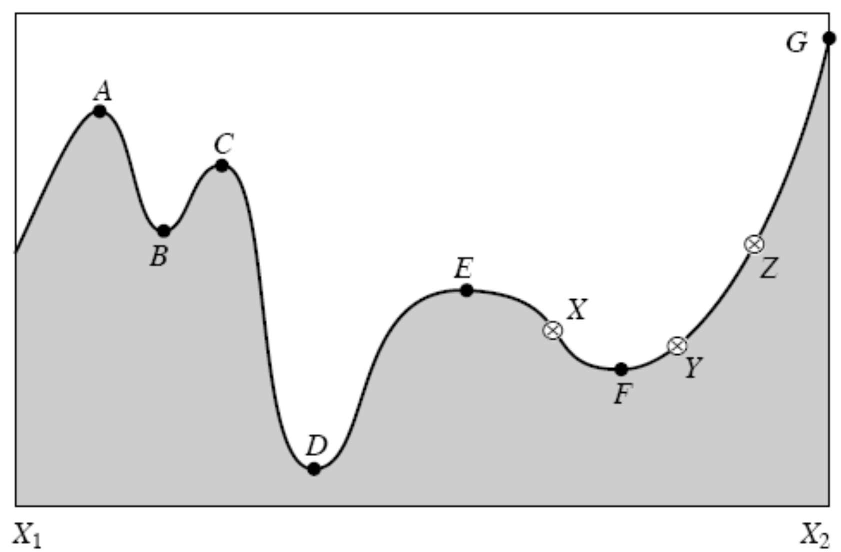
$$x_2^a - x_2^b = \frac{c\sigma_1\sigma_2}{\sigma_1^2 + \sigma_o^2}(y_1 - x_1^b)$$

cv_options=6 in WRFDA

Other Remarks

- Observation operator can be non-linear and thus analysis error PDF is not necessarily Gaussian

- For non-linear problem, J(x) can have multiple local minimum. Final solution of least square depends on starting point of iteration, e.g., choose the background x_b as the first guess.



Other Remarks

- **B** matrix is of very large dimension, explicit inverse of **B** is impossible, substantial efforts in data assimilation were given to the estimation and modeling of **B**.

- **B** shall be spatially-varied and time-evolving according to weather regime.

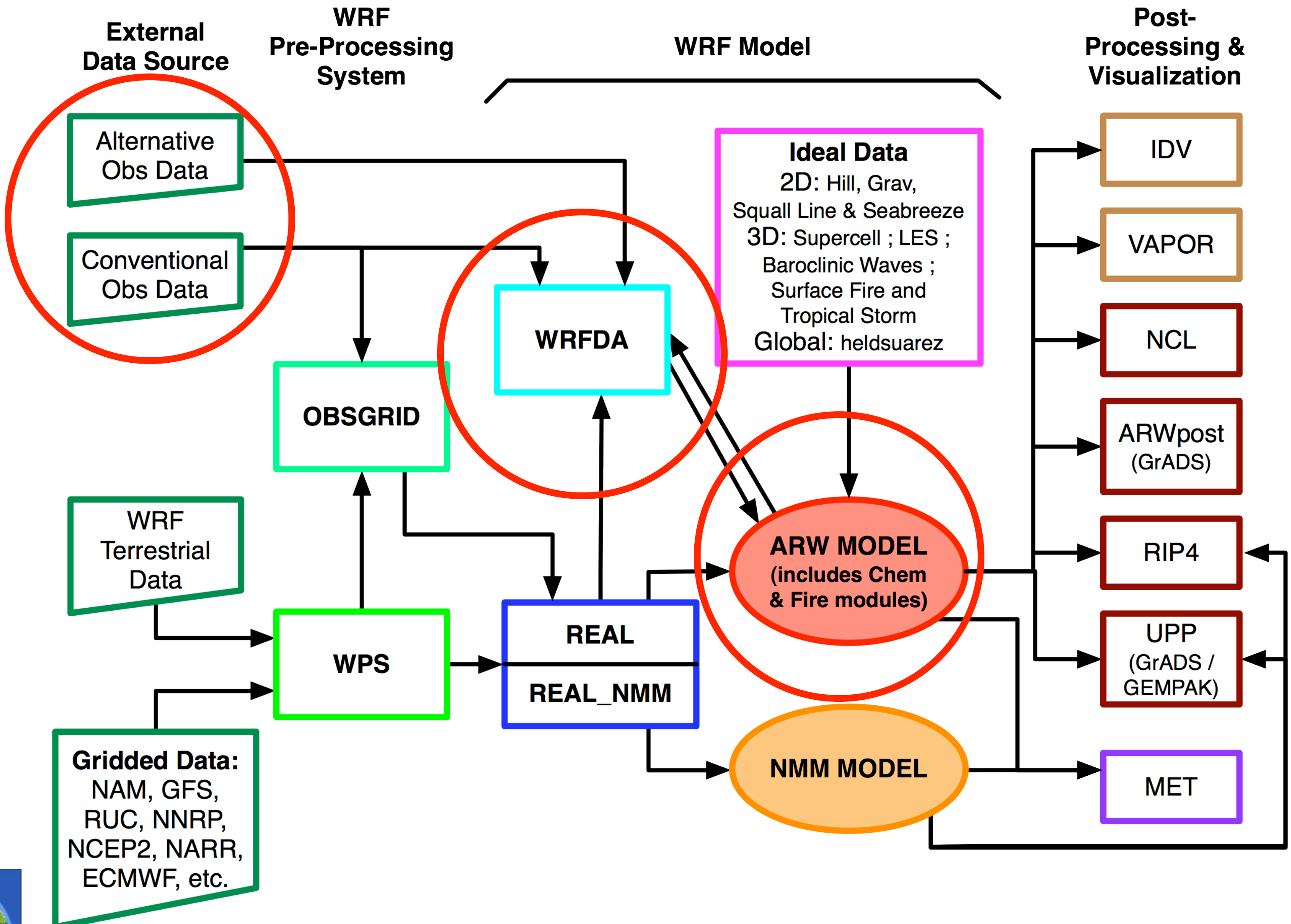
- Analysis can be sub-optimal if using inaccurate estimate of **B** and **R**.

- Could use non-Gaussian PDF
  - Thus not a least square cost function
  - But difficult (usually slow) to solve

Outline

- What is data assimilation
  - Scalar case
  - Two state variables case
  - General case
- Introduction to WRF Data Assimilation

WRFDA in WRF Modeling System



What WRFDA can do?

- Provide Initial conditions for the WRF model forecast
- Verification and validation via difference b.w. obs and model
  - See the last Lecture by Kavulich
- Observing system design, monitoring and assessment
- Reanalysis
- Better understanding:
  - Data assimilation methods
  - Model errors
  - Data errors
  - …

Assimilation methods

- Empirical methods
  - Successive Correction Method (SCM)
  - Nudging
  - Physical Initialisation (PI), Latent Heat Nudging (LHN)
- Statistical methods
  - Optimal Interpolation (OI)
  - 3-Dimensional VARiational data assimilation (3DVAR)
  - 4-Dimensional VARiational data assimilation (4DVAR)
- Advanced methods
  - Extended Kalman Filter (EKF)
  - Ensemble Kalman Filter (EnKF)
  - Hybrid VAR/Ens DA

WRFDA is a Data Assimilation system built within the WRF software framework, …

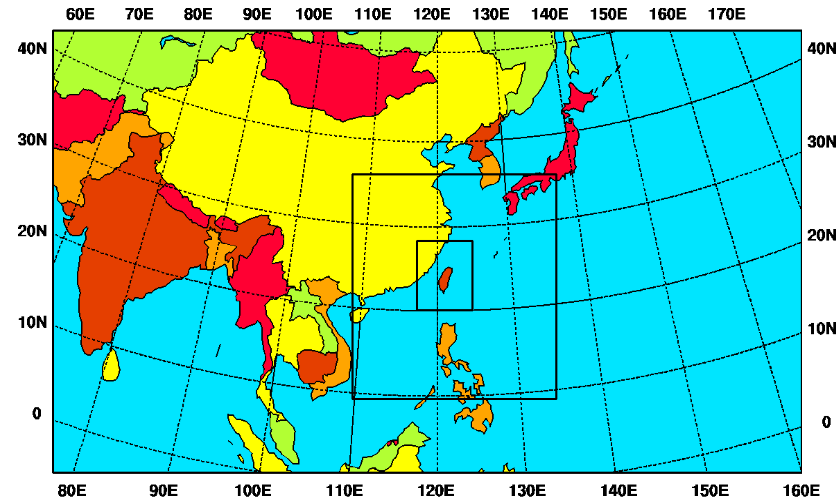
- **Goal:** Community WRF DA system for
  - research/operations, and
  - deterministic/probabilistic applications.

- **DA Techniques:**
  - 3D-Var (Lecture by Schwartz)
  - 4D-Var (Lecture by Liu)
  - Ensemble Transformed Kalman Filter
  - Hybrid-3DVAR (Lecture by Schwartz)

- **Support:**
  - NCAR/MMM via wrfhelp@ucar.edu

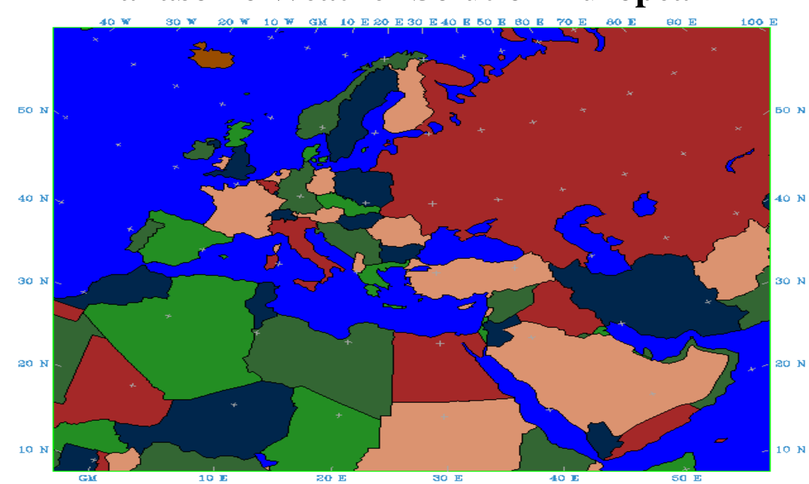
- **Observations:** Conv.+Sat.+Radar(+bogus)
  - Lectures by Bresch and Sun.

Domain area of CWB WRF



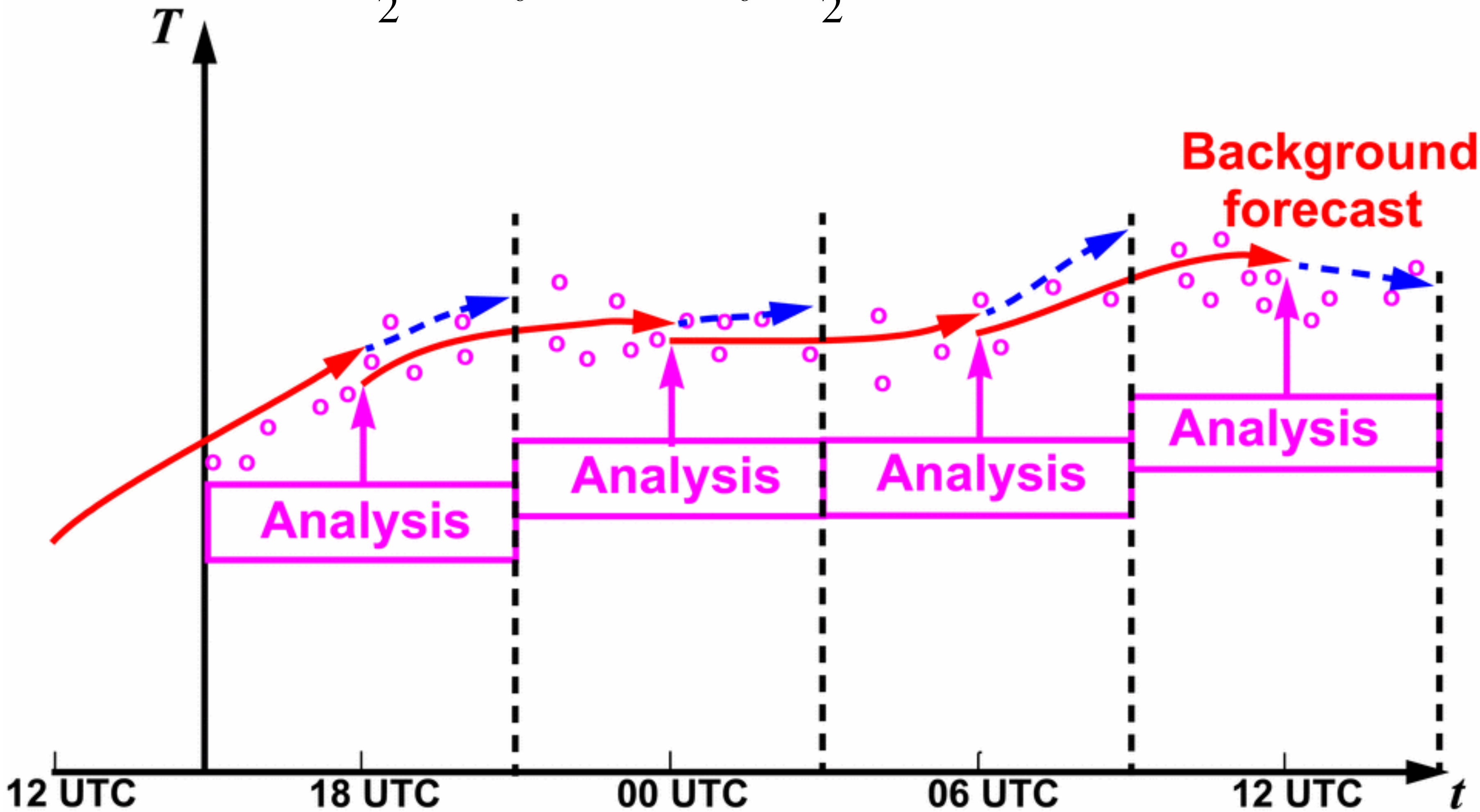
Both operations run in hybrid-3DVAR mode

Panasonic Weather Solution European



3DVAR

$$J(x) = \frac{1}{2}(x - x_b)^{\mathrm{T}} \mathrm{B}^{-1}(x - x_b) + \frac{1}{2}[H(x) - y]^{\mathrm{T}} \mathrm{R}^{-1}[H(x) - y]$$



WRFDA Observations

- **In-Situ:**
  - SYNOP
  - METAR
  - SHIP
  - BUOY
  - TEMP
  - PIBAL
  - AIREP, AIREP humidity
  - TAMDAR
- **Bogus:**
  - TC bogus
  - Global bogus

- **Remotely sensed retrievals:**
  - Atmospheric Motion Vectors (geo/polar)
  - SATEM thickness
  - Ground-based GPS TPW or ZTD
  - SSM/I oceanic surface wind speed and TPW
  - Scatterometer oceanic surface winds
  - Wind Profiler
  - **Radar data (enhancements in V3.7)**
  - Satellite temperature/humidity/thickness profiles
  - GPS refractivity (e.g. COSMIC)
  - Stage IV precipitation/rain rate data (4D-Var)

- **Radiances: can use RTTOV_11.1 or 11.2 (new in V3.7) or CRTM_2.1.3:**

– HIRS	NOAA-16, NOAA-17, NOAA-18, NOAA-19, METOP-A
– AMSU-A	NOAA-15, NOAA-16, NOAA-18, NOAA-19, EOS-Aqua, METOP-A, METOP-B
– AMSU-B	NOAA-15, NOAA-16, NOAA-17
– MHS	NOAA-18, NOAA-19, METOP-A, METOP-B
– AIRS	EOS-Aqua
– SSMIS	DMSP-16, DMSP-17, DMSP-18
– IASI	METOP-A, METOP-B
– ATMS	Suomi-NPP
– MWTS	FY-3
– MWHS	FY-3
– SEVIRI	METEOSAT

WRFDA is flexible to allow assimilation of different formats of observations:

- **Little_r (ascii), HDF, Binary**
- **NOAA MADIS (netcdf),**
- **NCEP PrepBufr,**
- **NCEP radiance bufr**

www2.mmm.ucar.edu/wrf/users/wrfda

WRFDA USERS PAGE

Home | System | User Support | Download | Publications & Documentation | Links | Internal

Search

WRFDA Home

- WRFDA News
- Public Domain Notice
- Contact Us
- WRF Users Page

WRF Data Assimilation System Users Page

Welcome to the page for users of the Weather Research and Forecasting (WRF) model data assimilation system (WRFDA). The WRFDA system is in the public domain and is freely available for community use. It is designed to be a flexible, state-of-the-art atmospheric data assimilation system that is portable and efficient on available parallel computing platforms. WRFDA is suitable for use in a broad range of applications, across scales ranging from kilometers for regional and mesoscale modeling to thousands of kilometers for global scale modeling.

The Mesoscale and Microscale Meteorology (MMM) Laboratory of NCAR currently maintains and supports a subset of the overall WRF code (Version 3) that includes:

- WRF Software Framework (WSF)
- Advanced Research WRF (ARW) dynamic solver, including one-way, two-way nesting and moving nests, grid and observation nudging
- WRF Pre-Processing System (WPS)
- **WRF Data Assimilation System (WRFDA)** *(found on this site)*
- Numerous physics packages contributed by WRF partners and the research community

Other components of the WRF system will be supported for community use in the future, depending on interest and available resources.

Quick links:

LATEST WRFDA RELEASE

WRFDA Version 3.7
(Released April 20, 2015)

UPCOMING EVENTS

June 15–19, 2015

2015 WRF Workshop, NCAR Center Green Campus, Boulder, CO, USA.
Registration is now open!

August 3–7, 2015

2015 WRFDA New User Tutorial, NCAR Foothills Laboratory, Boulder, CO, USA.
Registration is now full!

WHAT'S NEW

April 20, 2015

WRFDA Version 3.7 has been released. View release notes.

April 3, 2015

Registration is now open for the WRF Workshop and the WRFDA tutorial.

2015 WRFDA Tutorial Agenda

Wednesday - August 5, 2015

08:00-08:30	Registration	
08:30-09:00	**Welcome and Participants' Introduction**	Zhiquan Liu
09:00-10:00	**Overview of WRF Data Assimilation**	Zhiquan Liu
10:00-10:20	Coffee Break	
10:20-11:10	**WRFDA Software and Compilation**	Michael Kavulich
11:10-12:00	**Observations (1): Conventional Obs Pre-Processing**	Jamie Bresch
12:00-13:00	Lunch	
13:00-14:00	**Algorithm (1): 3DVAR Setup, Run and Diagnostics**	Craig Schwartz
14:00-15:00	**Algorithm (2): Background Error Modeling and Estimation**	Syed Rizvi
15:00-15:20	Coffee Break	
15:20-15:30	**Introduction to practice sessions**	Michael Kavulich
15:30-18:00	**Practice Session 1 (OBSPROC, 3DVAR, GEN_BE, single-ob tests)**	

Thursday - August 6, 2015

09:00-10:00	**Observations (2): Radiance data assimilation**	Jamie Bresch
10:00-10:20	Coffee Break	
10:20-11:00	**Algorithm (3): 4DVAR**	Zhiquan Liu
11:00-12:30	**Practice Session 2 (Radiance, 4DVAR)**	
12:30-13:30	lunch	
13:30-14:20	**Algorithm (4): Hybrid Variational/Ensemble**	Craig Schwartz
14:20-15:10	**Observations (3): Radar Data Assimilation**	Jenny Sun
15:10-15:30	Coffee Break	
15:30-16:10	**WRFDA Tools and Verification Package**	Michael Kavulich
16:10-16:30	**Wrap-up discussion**	Zhiquan Liu
16:30-18:00	**Practice Session 3 (hybrid, radar, tools)**	

Friday - August 7, 2015

08:00-12:00	**Advanced practice session (WRF/WRFDA cycling, FGAT, FSO, advanced lessons)**	