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

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

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
- 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 means we trust equally the background and observation.
- But 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

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

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

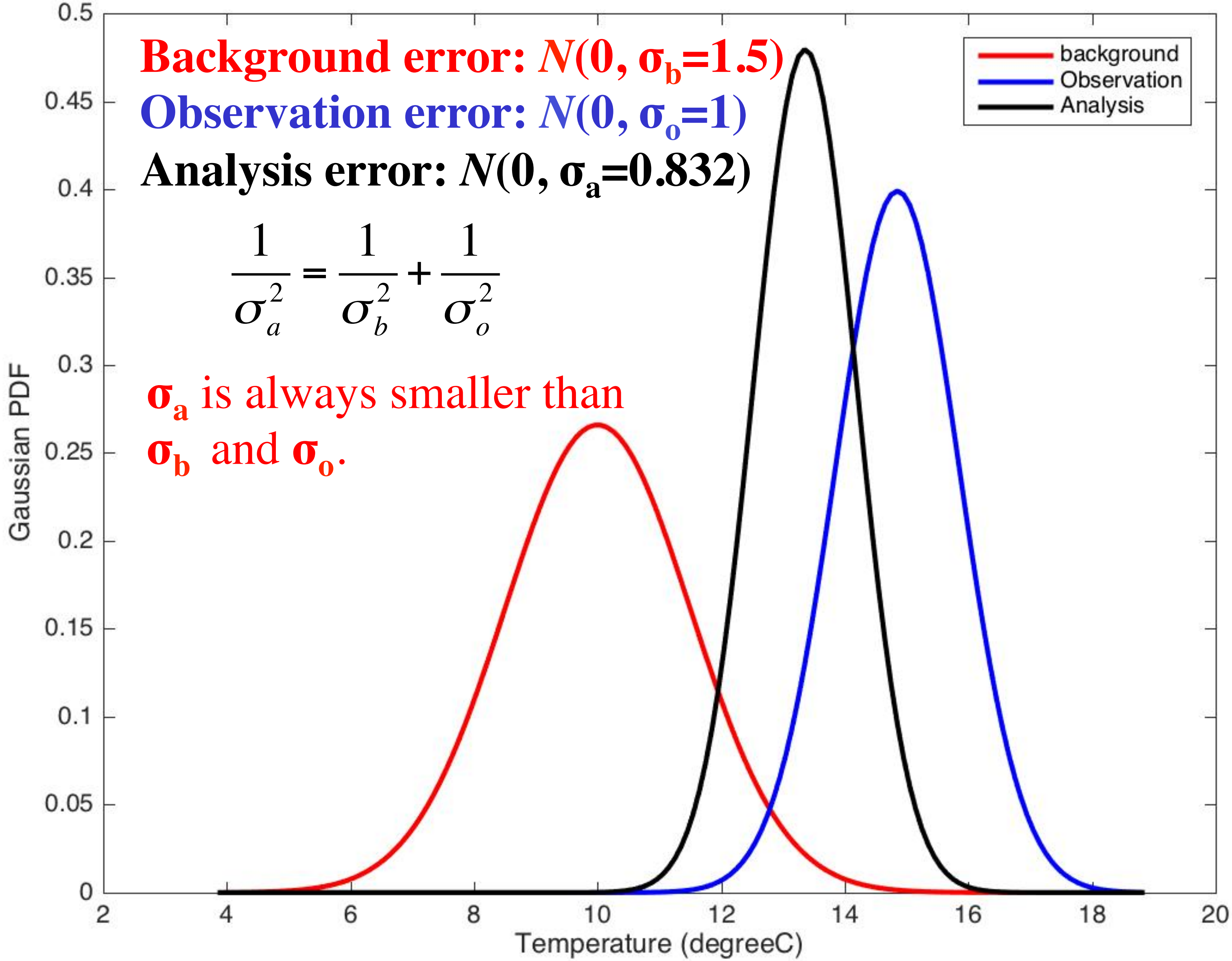
Scalar Case

- Minimize $$J(x) = \frac{1}{2}\frac{(x-x_b)^2}{\sigma_b^2} + \frac{1}{2}\frac{(x-y)^2}{\sigma_o^2}$$

- Is actually equivalent to maximize a Gaussian PDF

$$ce^{-J(x)}$$

Assume errors of X_b and y are unbiased


Background error: N(0, σb=1.5)
Observation error: N(0, σo=1)
Analysis error: N(0, σa=0.832)
1/σa² = 1/σb² + 1/σo²
σa is always smaller than
σb and σo.
background
Observation
Analysis
Gaussian PDF
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
2
4
6
8
10
12
14
16
18
20
Temperature (degreeC)

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*

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

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

Analysis increment for two variables

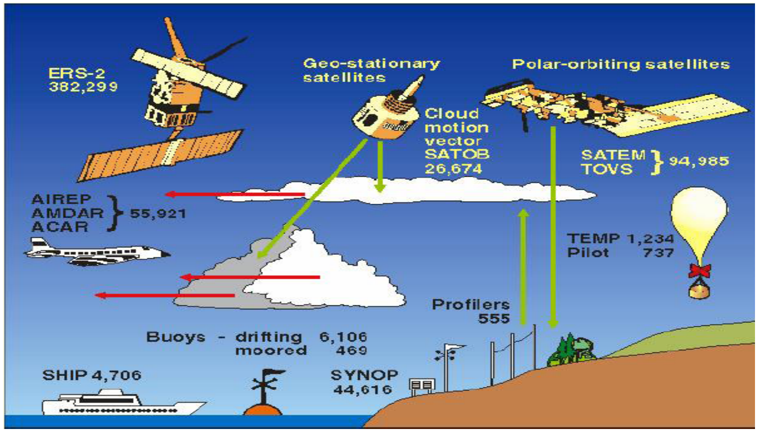
$$x_1^a - x_1^b = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_o^2}(y_1 - x_1^b)$$

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

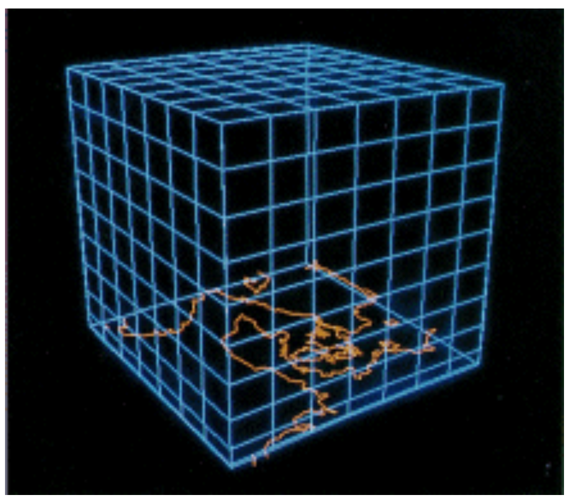
Unobserved variable x_2 gets updated through the error correlation *c* in the background error covariance.

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

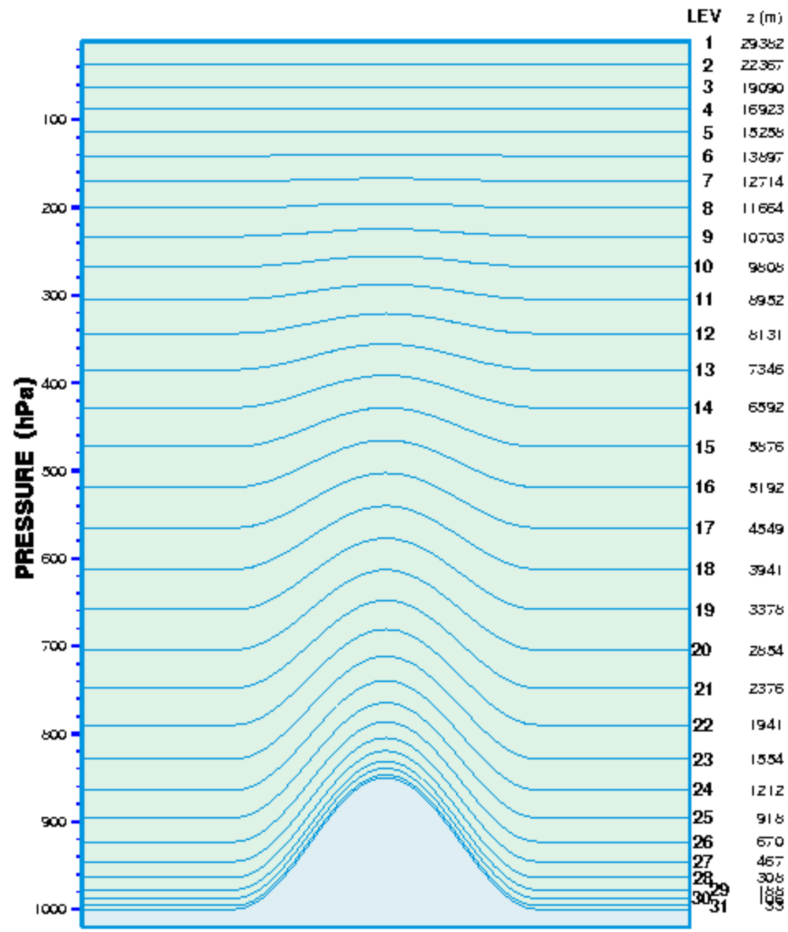
General Case



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

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

Observation error covariance

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

General Case: cost function

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

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

Minimize J(x) is equivalent to maximize a multi-dimensional 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:

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

This leads to analytical solution for the analysis increment:

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

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

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

Precision of Analysis with optimal B and R

$$\mathbf{A}^{-1} = \mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}$$

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

Or in another form: $\mathbf{A} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}$

With

$$\mathbf{K} = \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}$$

called Kalman gain matrix

Precision of analysis: more general formulation

$$\mathbf{A} = (\mathbf{I} - \mathbf{KH})\mathbf{B}_t(\mathbf{I} - \mathbf{KH})^{\mathrm{T}} + \mathbf{KR}_t\mathbf{K}^{\mathrm{T}}$$

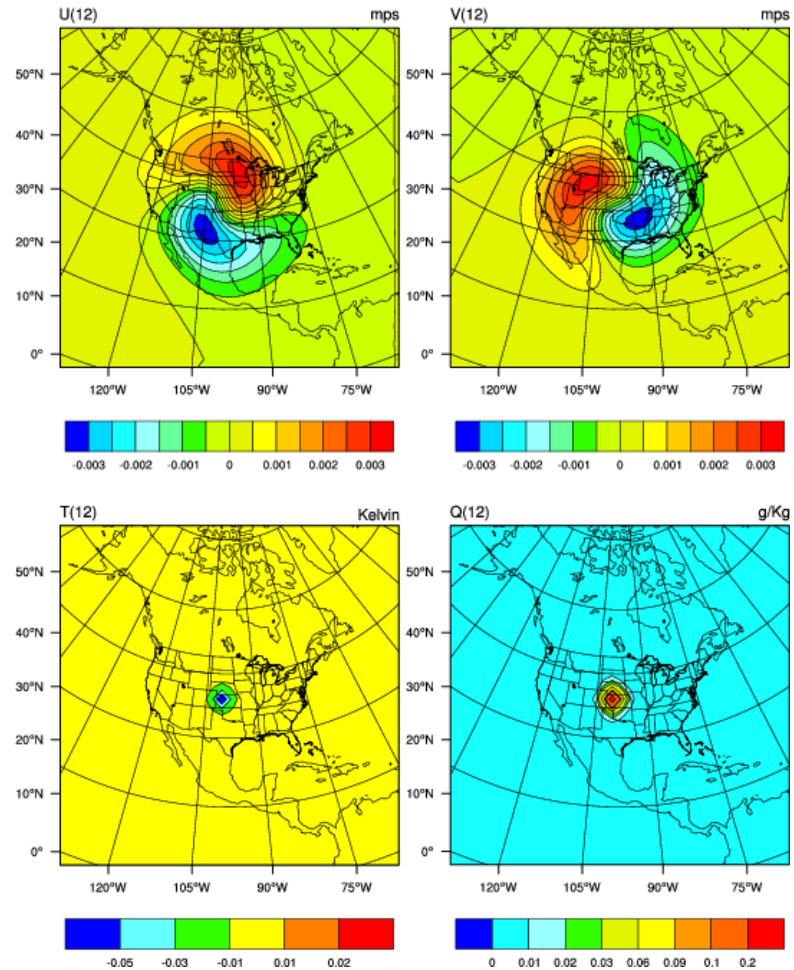
where B_t and R_t are "true" background and observation error covariances.

This formulation is valid for any given gain matrix K, which could be suboptimal (e.g., due to incorrect estimation/specification of B and R).

Analysis increment with a single humidity observation

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

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



It is generalization of previous two variables case:

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

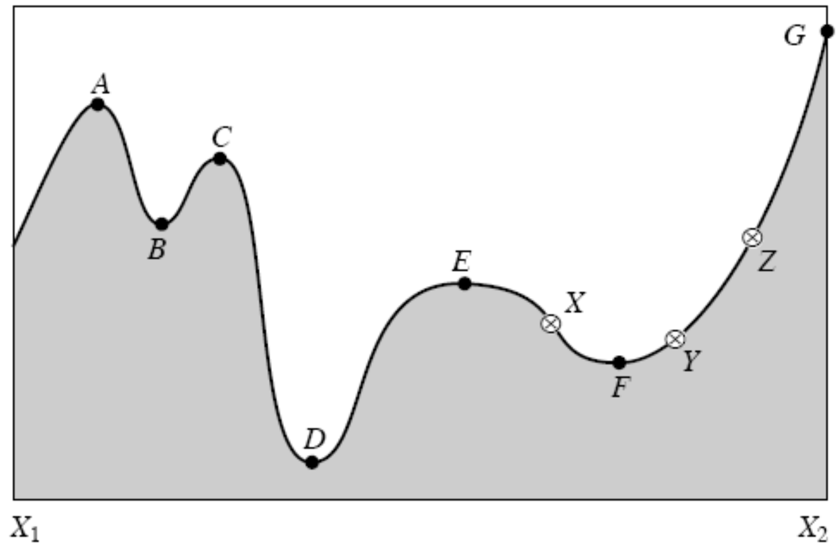
$$x_2^a - x_2^b = \frac{c\sigma_1\sigma_2}{\sigma_1^2 + \sigma_o^2}(y_1 - x_1^b)$$

cv_options=6 in WRFDA

Other Remarks

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

- J(x) can have multiple local minima. 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**.

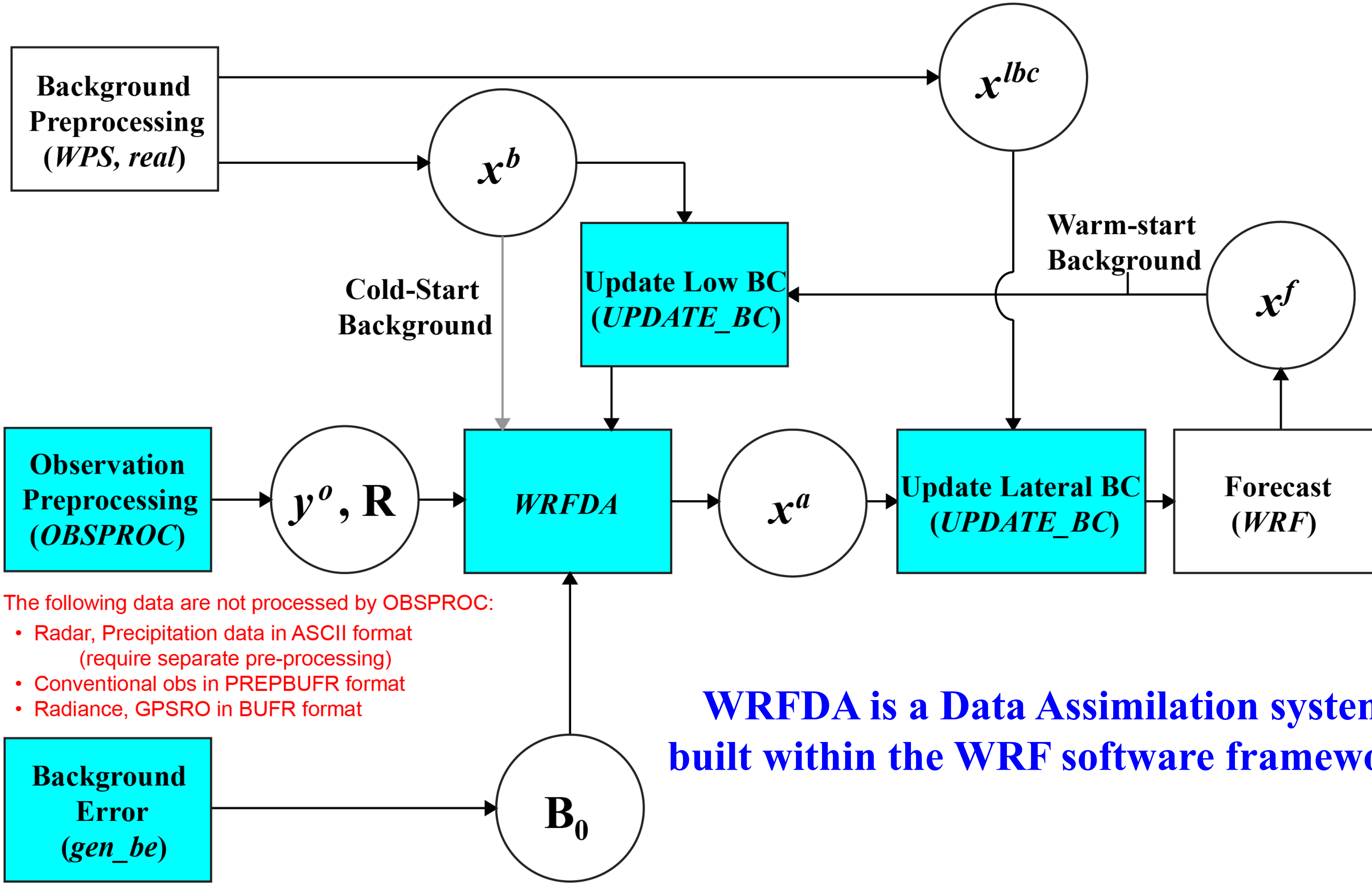
- Could use non-Gaussian PDF
  - Thus not a least square cost function
  - Difficult (usually slow) to solve; could transform into Gaussian problem via variable transform

Outline

- What is data assimilation
  - Scalar case
  - Two state variables case
  - General case

- Introduction to WRF Data Assimilation

WRFDA in the WRF Modeling System



The following data are not processed by OBSPROC:

- Radar, Precipitation data in ASCII format (require separate pre-processing)
- Conventional obs in PREPBUFR format
- Radiance, GPSRO in BUFR format

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

What WRFDA can do?

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

DA algorithms currently available in WRFDA

- 3DVAR and FGAT
  - Different options for choice of control variables (e.g., Psi/Chi or U/V) and background error covariance modeling (e.g., vertical EOF or vertical recursive filter)
- 4DVAR
  - TL/Adjoint (i.e., WRFPlus code) of WRF up-to-date with WRF
  - Allow LBC control variable and Jc-DFI
- Hybrid-3DEnVar
  - Can run in dual-resolution mode
  - Can ingest ensemble from global or regional sources
- ETKF: for generating ensemble analysis

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 (reflectivity/retrieved rainwater, and radial-wind)**
  - Satellite temperature/humidity/thickness profiles
  - GPS refractivity (e.g. COSMIC)
  - **Stage IV precipitation/rain rate data (4D-Var only)**
- **Radiances:**

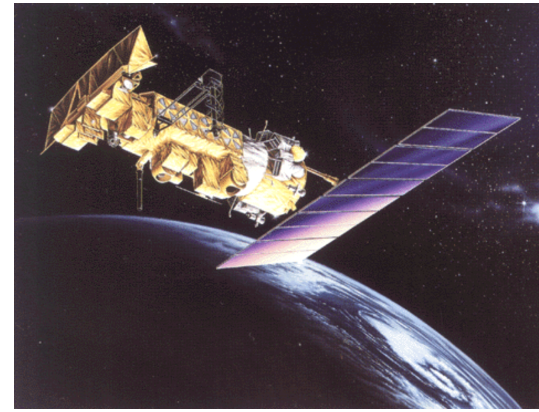
HIRS	NOAA-16, NOAA-17, NOAA-18, NOAA-19, METOP-A
AMSU-A	NOAA-15, NOAA-16, NOAA-18, NOAA-19, EOS-Aqua, METOP-A, METOP-B
AMSU-B	NOAA-15, NOAA-16, NOAA-17
MHS	NOAA-18, NOAA-19, METOP-A, METOP-B
AIRS	EOS-Aqua
SSMIS	DMSP-16, DMSP-17, DMSP-18
IASI	METOP-A, METOP-B
ATMS	Suomi-NPP
MWTS	FY-3
MWHS	FY-3
SEVIRI	METEOSAT
AMSR2	**GCOM-W1 (new in V3.8)**

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

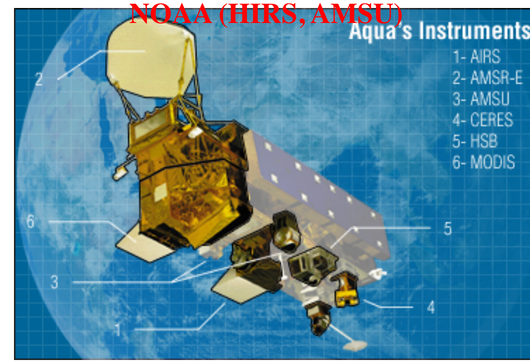
- **Little_r (ascii), HDF, Binary**
- **NOAA MADIS (netcdf),**
- **NCEP PrepBufr,**
- **NCEP radiance bufr**

WRFDA
Radiance Assimilation

- Two RTM interfaces
  - RTTOV or CRTM
- Variational Bias Correction
- Modular code design to ease adding new satellite sensors
- Capability for cloudy radiance DA

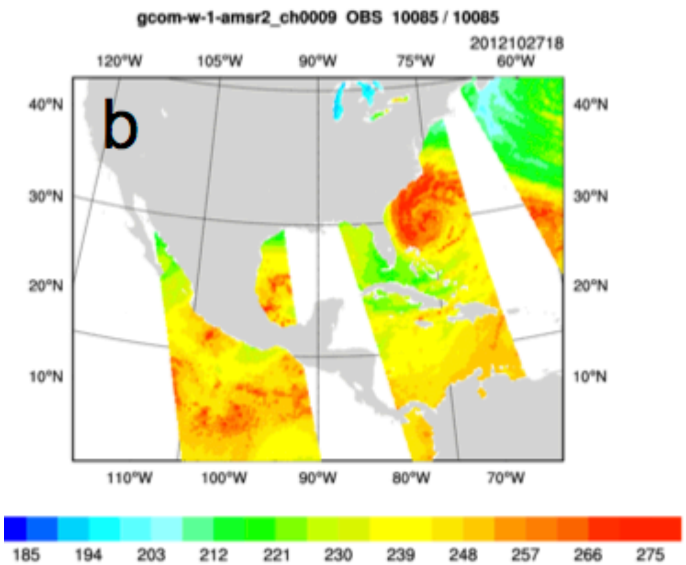


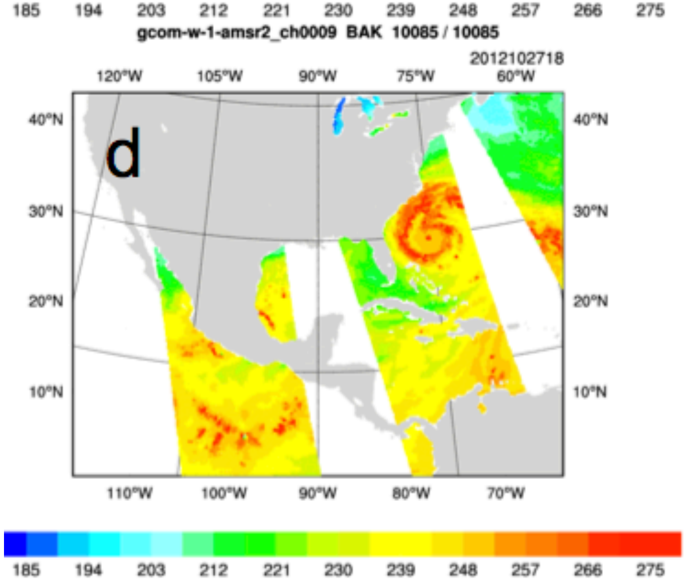
NOAA (HIRS, AMSU)

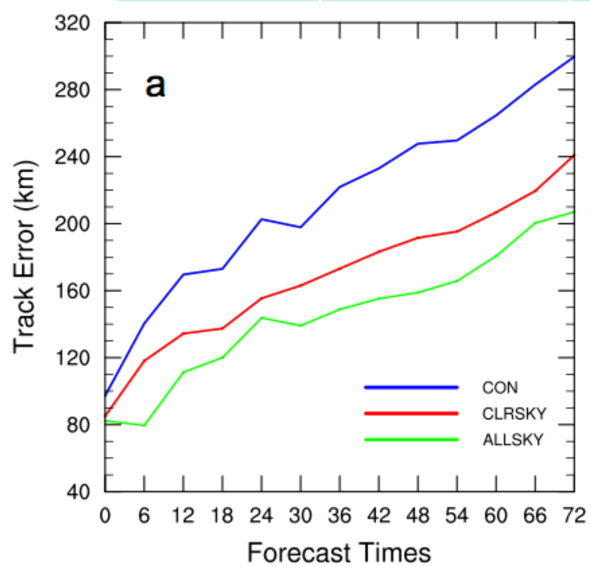


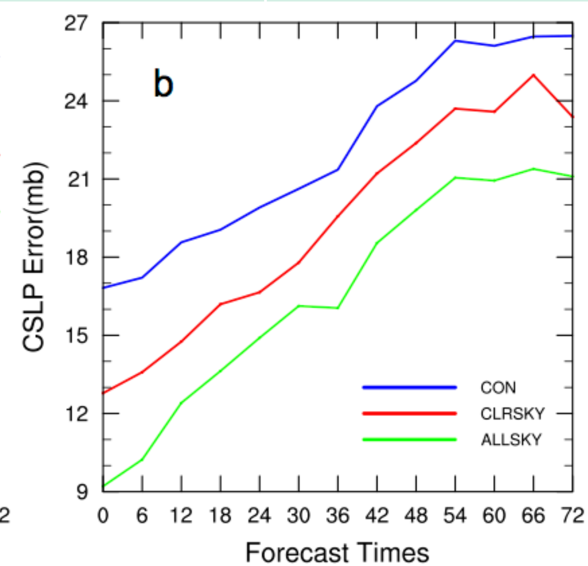
New: all-sky radiance DA: AMSR2

Channel	Frequency (GHz)	Polarization	Footprint (along scan* along track)
1,2	6.925	V,H	35*61 km
3,4	7.3	V,H	35*61 km
5,6	10.65	V,H	24*41 km
7,8	18.7	V,H	13*22 km
9,10	23.8	V,H	15*26 km
11,12	36.5	V,H	7*12 km
13,14	89.0	V,H	3*5 km









WRFDA: Beta releases

WRFDA is undergoing continuous development as more capabilities are added, both by NCAR scientists and community contributors. On this page you can find pre-release versions of new capabilities. As these capabilities are new and not fully tested, we appreciate any feedback you can offer us: contact us through wrfhelp or the WRFDA webmaster.

Below is a list of the current beta releases we have available.

AMSR2 Cloudy Radiance assimilation

Typically, with radiance assimilation in WRFDA, pixels which are determined to have clouds in them are rejected. However, we have developed the ability to assimilate cloud-affected radiance observations with the JAXA GCOM-W1 AMSR2 instrument. This capability is described in this PDF guide, and in the following publication:

> Chun Yang, Zhiquan Liu, Jamie Bresch, Syed R. H. Rizvi, Xiang-Yu Huang and Jinzhong Min, 2016: AMSR2 all-sky radiance assimilation and its impact on the analysis and forecast of Hurricane Sandy with a limited-area data assimilation system. *Tellus A*, **68**, 30917, doi:10.3402/tellusa.v68.30917.

Download pre-release code

To download beta release code:

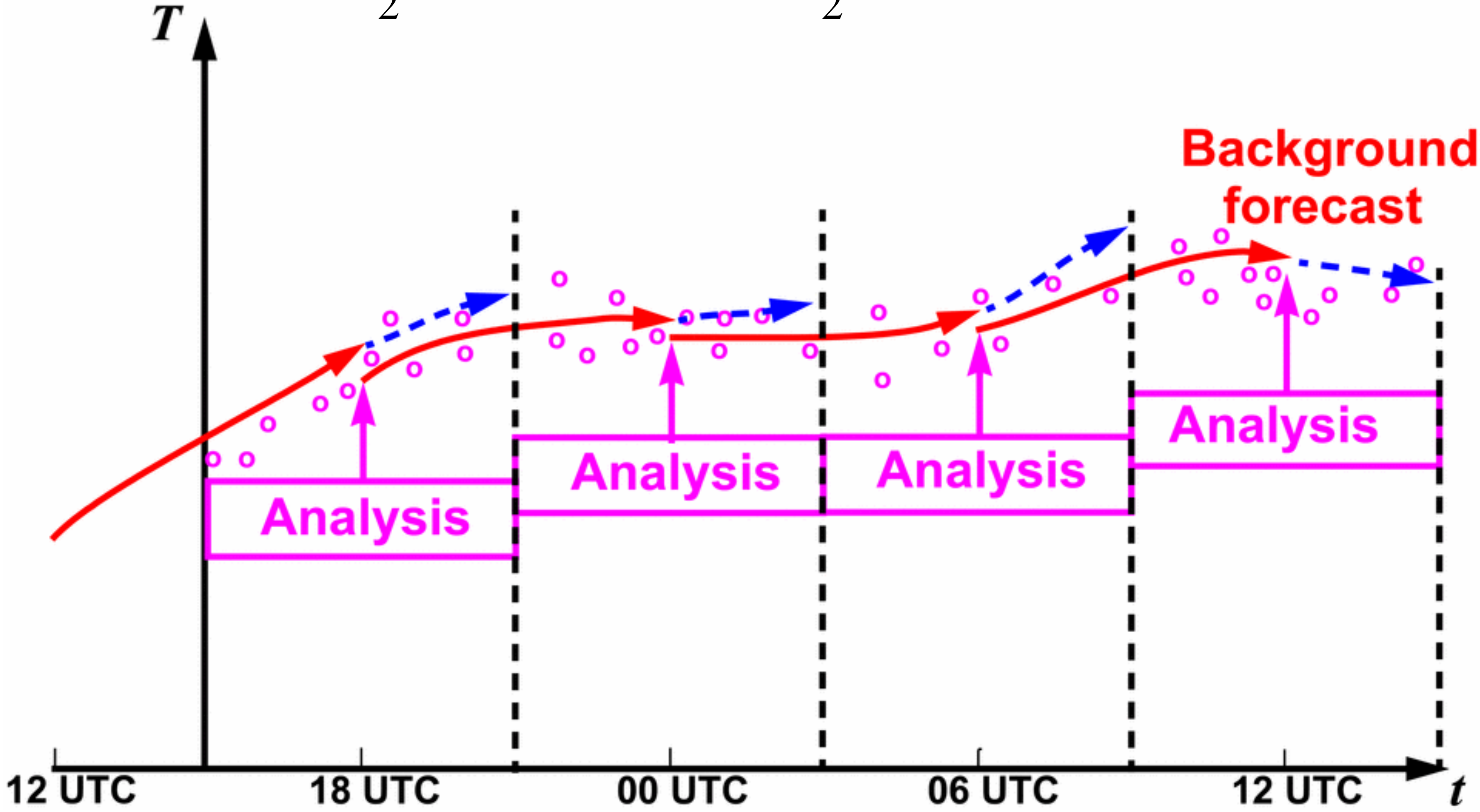
Fill out the registration form by clicking **'New Users'** below, or select **'Returning Users'** if you have already registered to download WRF or WRFDA in the past. You will be redirected to a page where you can download a tar file with the code you are interested in.

New Users

Returning Users

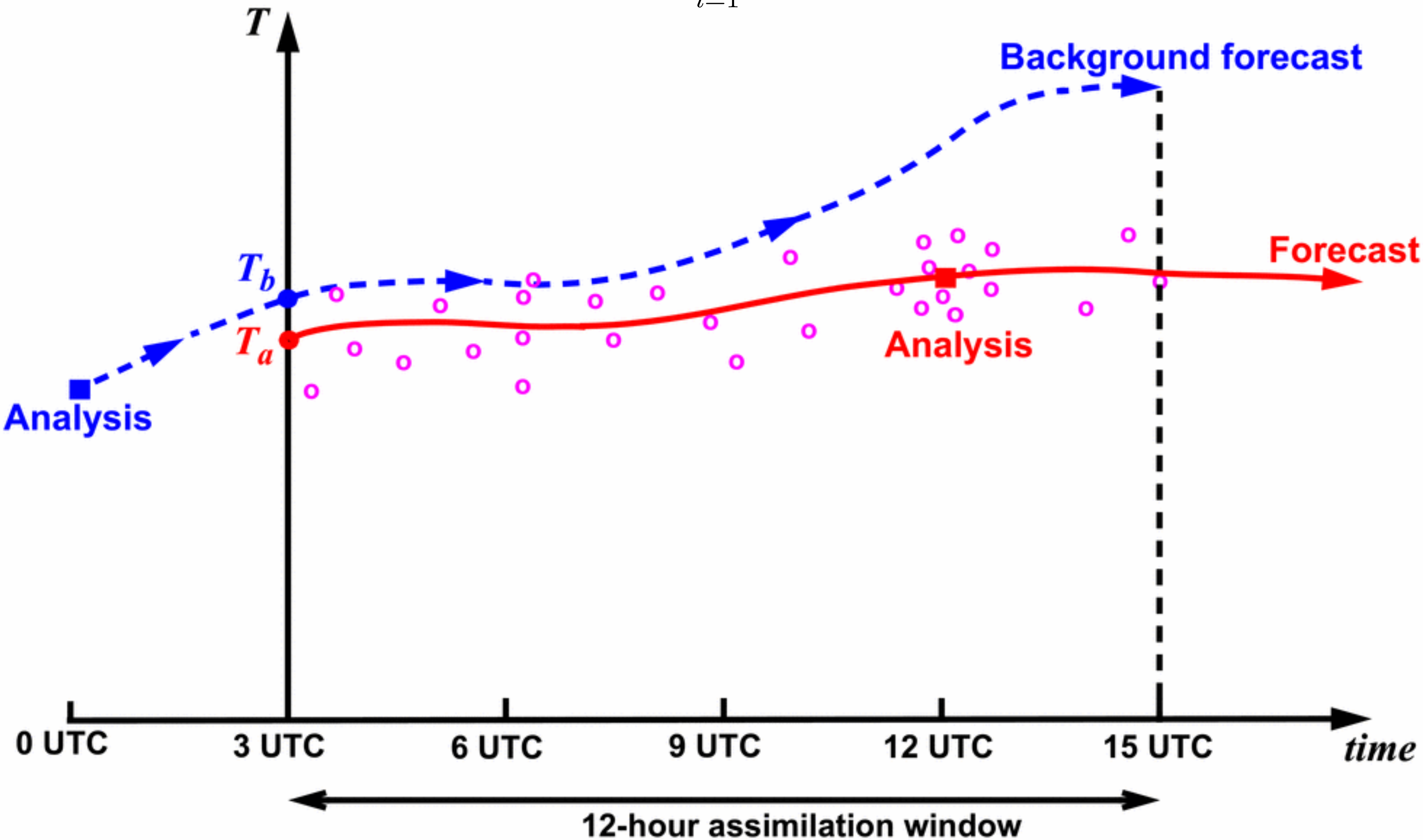
3DVAR (Barker et al. 2004)

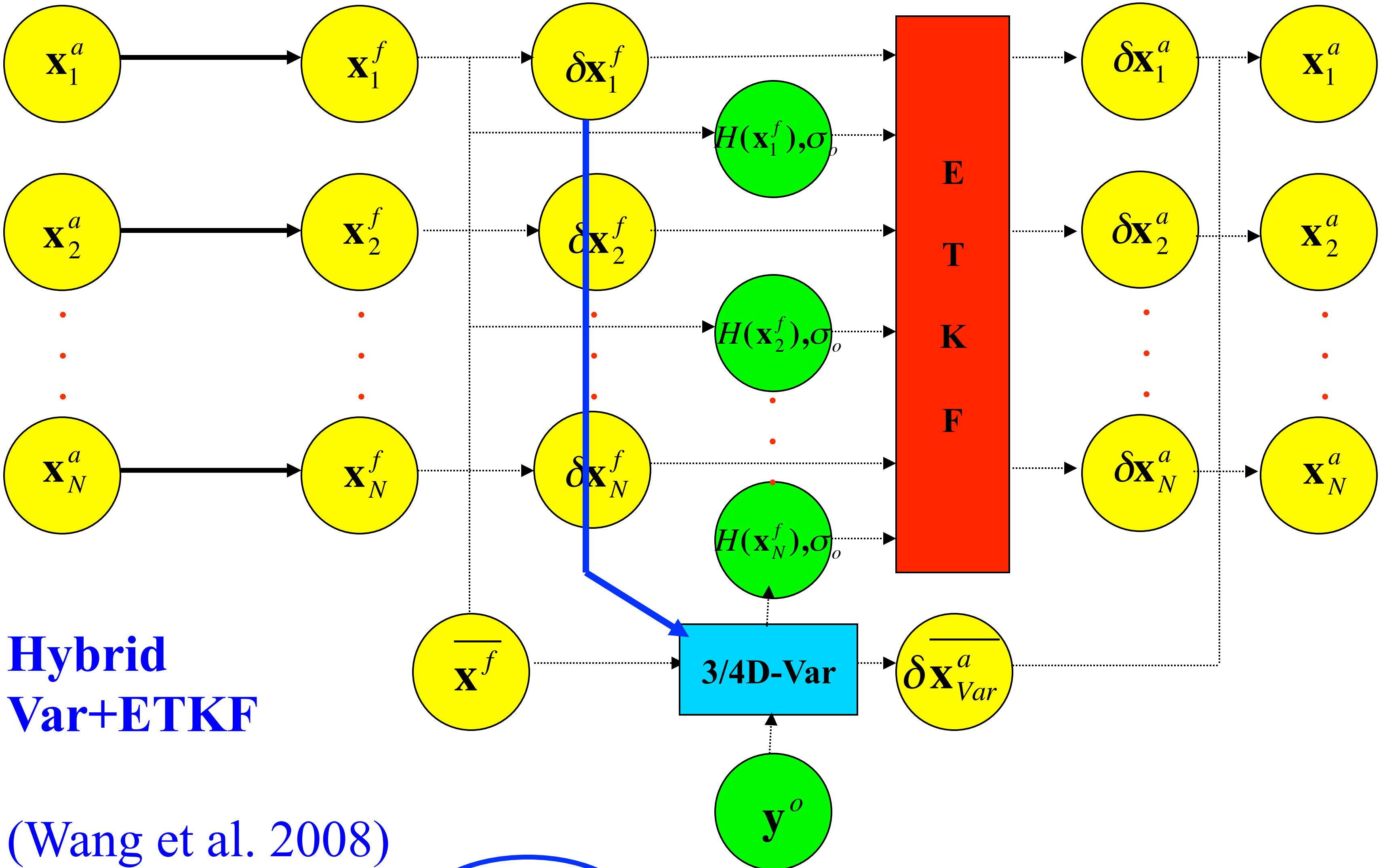
$$J(x) = \frac{1}{2}(x - x_b)^{\mathrm{T}} \mathrm{B}^{-1}(x - x_b) + \frac{1}{2}[H(x) - y]^{\mathrm{T}} \mathrm{R}^{-1}[H(x) - y]$$



4DVAR (Huang et al. 2009)

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2}\sum_{i=1}^{N}[H_i(M_i(\mathbf{x}_0)) - \mathbf{y}_i]^{\mathrm{T}}\mathbf{R}_i^{-1}[H_i(M_i(\mathbf{x}_0)) - \mathbf{y}_i]$$

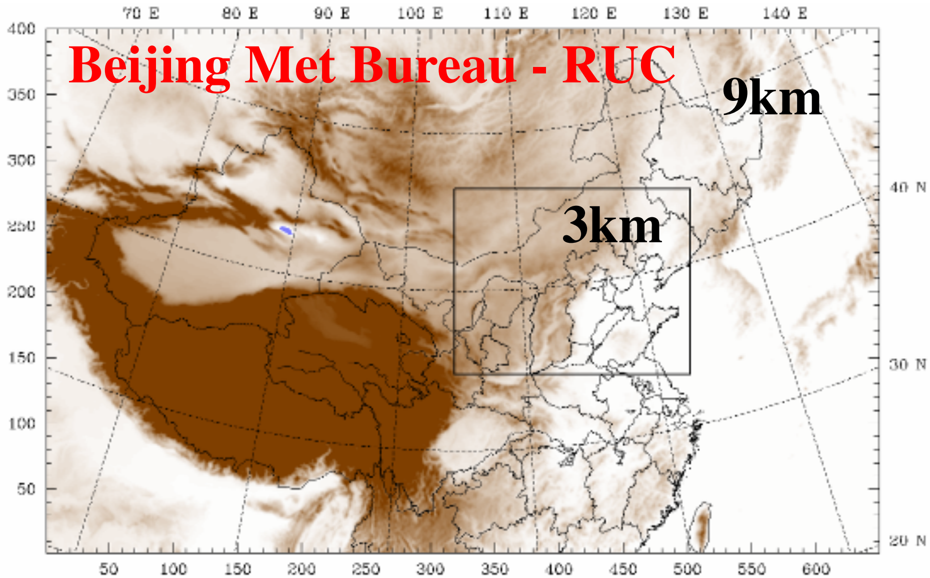


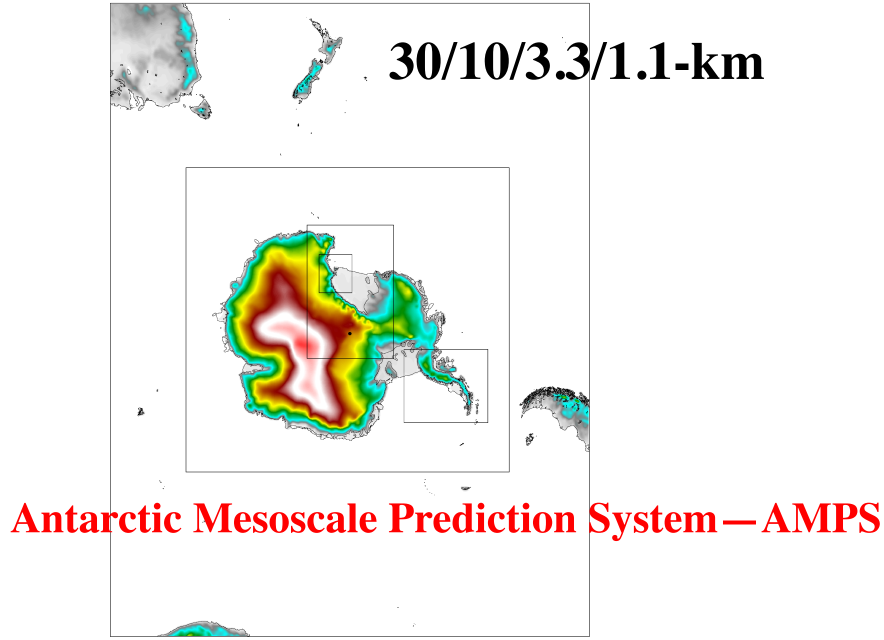


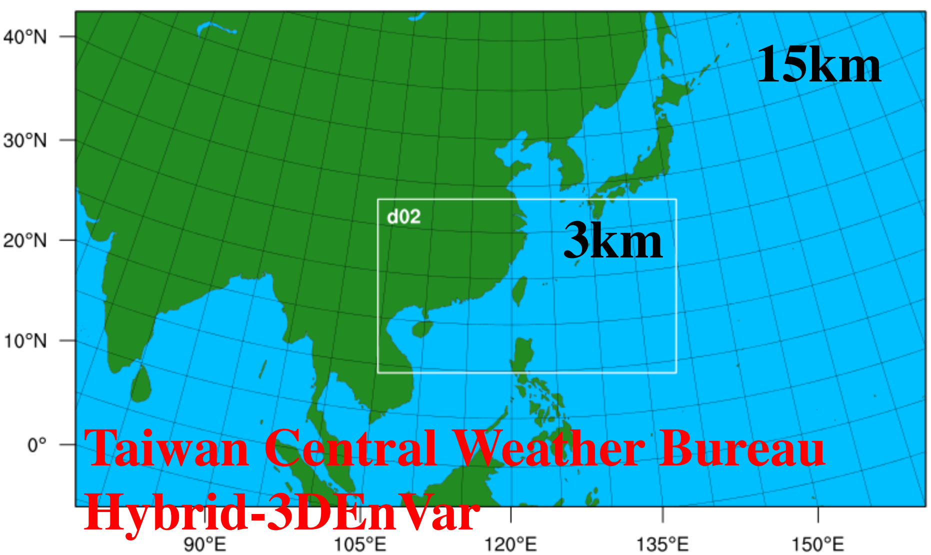
Hybrid
Var+ETKF

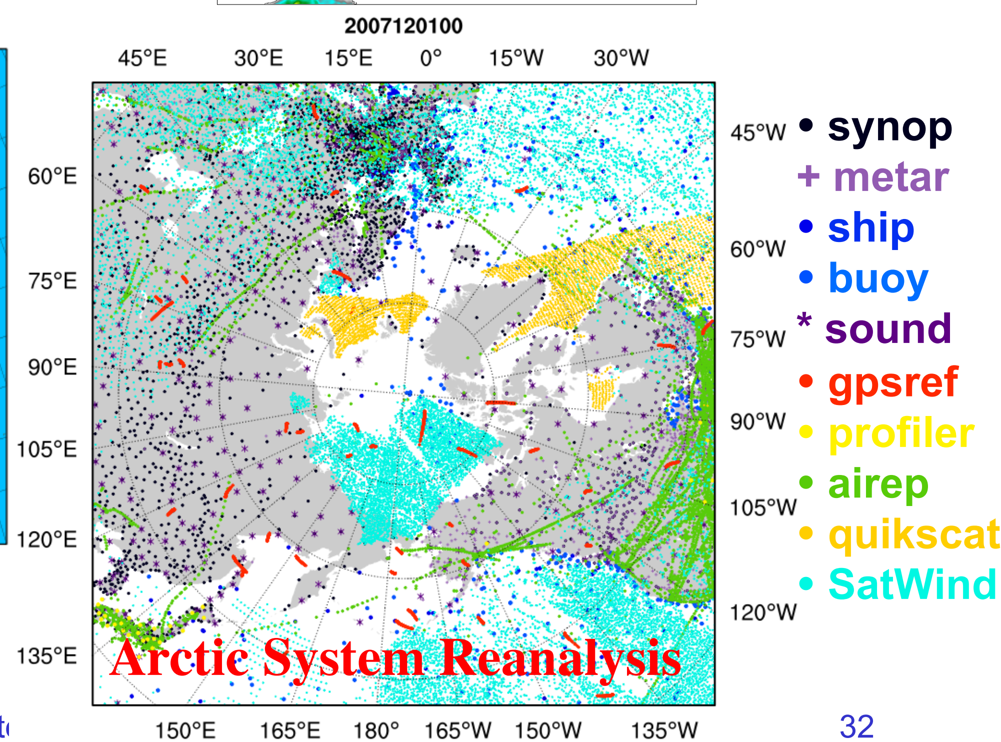
(Wang et al. 2008)

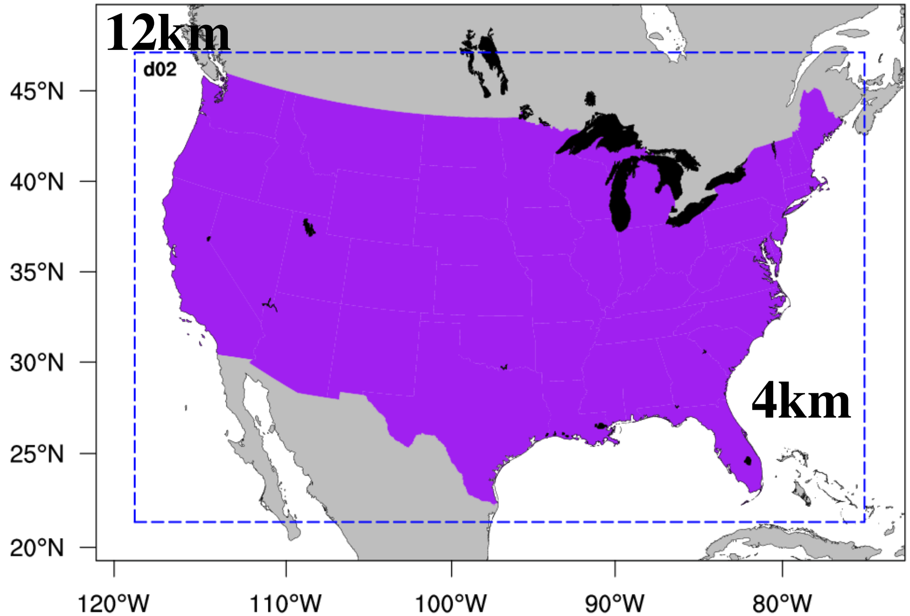
$$J = \frac{W_b}{2}\mathbf{v}^T\mathbf{v} + \frac{W_\alpha}{2}\mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \frac{1}{2}\sum_{i=0}^{n}\left[\mathbf{d}_i - \mathbf{H}_i\mathbf{M}_i\mathbf{U}\mathbf{v}\right]^T \mathbf{R}_i^{-1}\left[\mathbf{d}_i - \mathbf{H}_i\mathbf{M}_i\mathbf{U}\mathbf{v}\right]$$

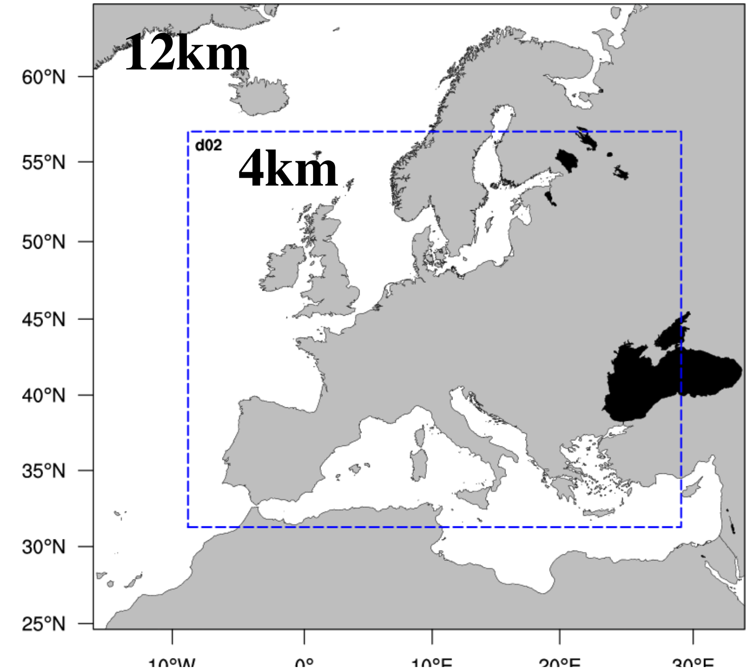

Beijing Met Bureau - RUC
9km
3km


30/10/3.3/1.1-km
Antarctic Mesoscale Prediction System—AMPS

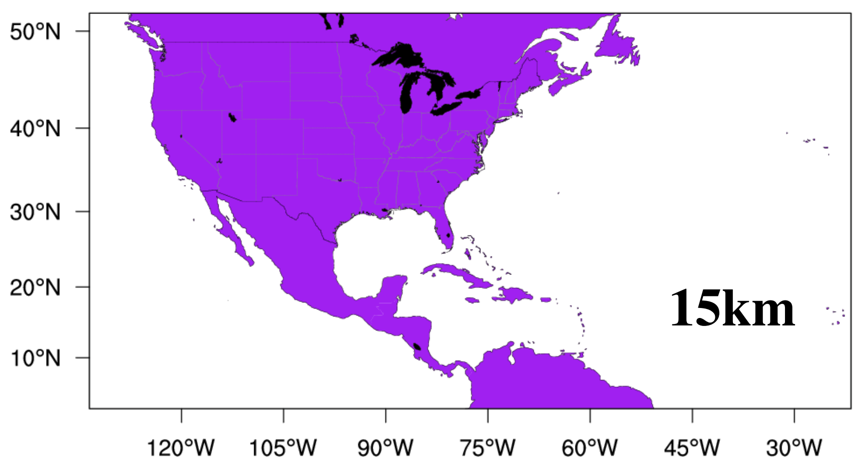

15km
d02
3km
Taiwan Central Weather Bureau
Hybrid-3DEnVar

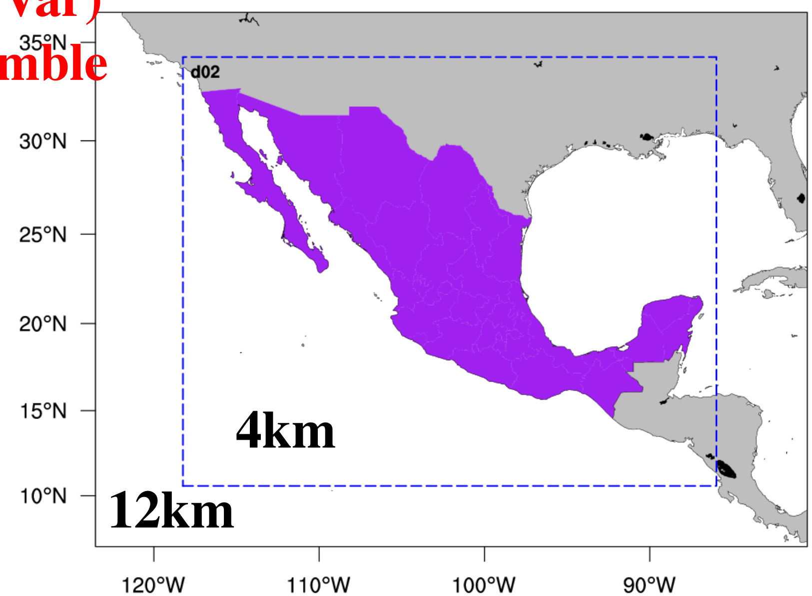

2007120100
• synop
+ metar
• ship
• buoy
* sound
• gpsref
• profiler
• airep
• quikscat
• SatWind
Arctic System Reanalysis





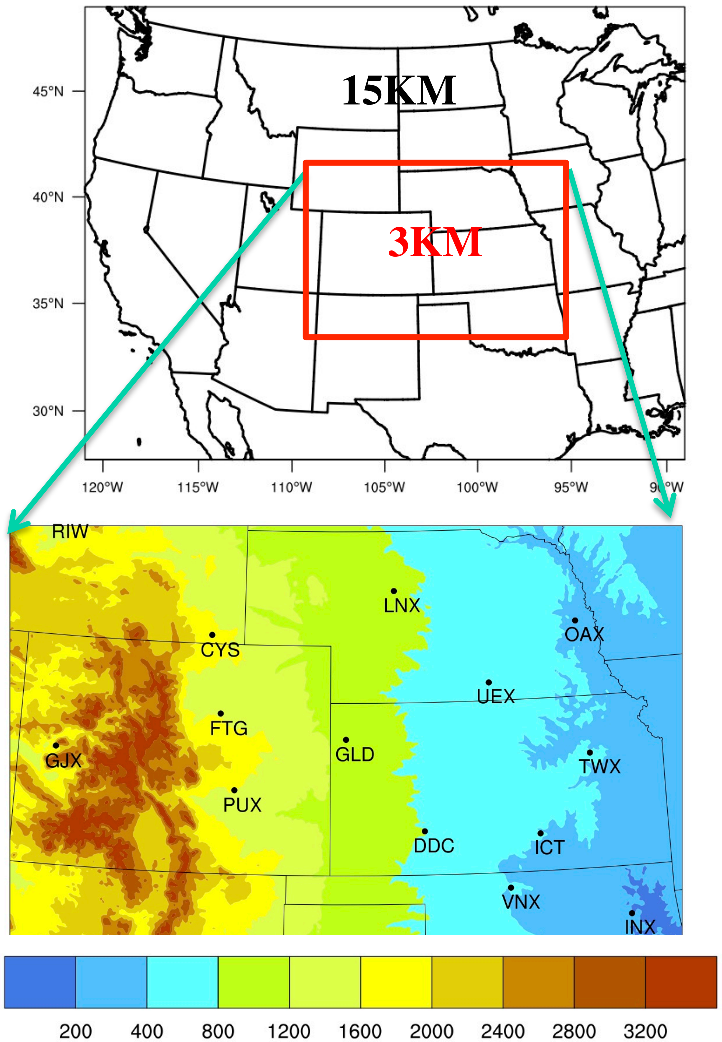
Operational domains by PWS (hybrid-3DEnVar)
Ensemble covariances from global GFS ensemble



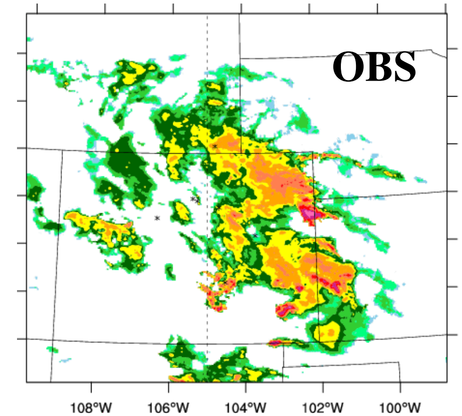


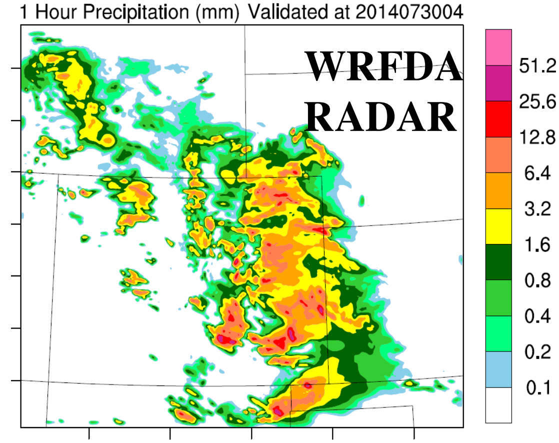
Radar DA for hydrological application

STEP Hydromet Real Time Exp. during spring time

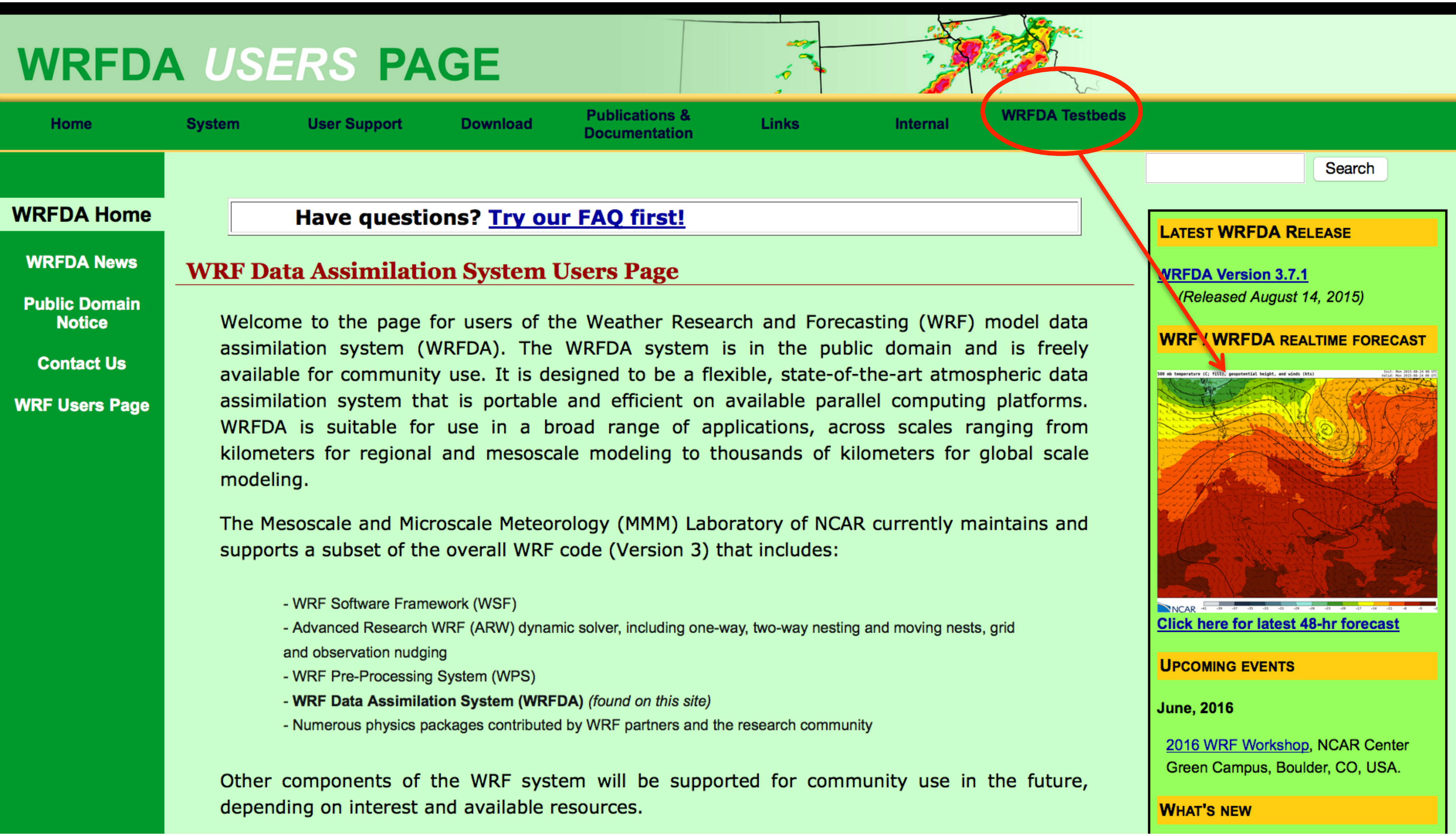


- The goal is to improve lo cal-scale QPF in coupled hydromet system
- < 1 h rapid update
- Radar radial velocity an d reflectivity assimilation
- High resolution vs. ense mble
- Impact of terrain
- Improved results in capt uring localized storms





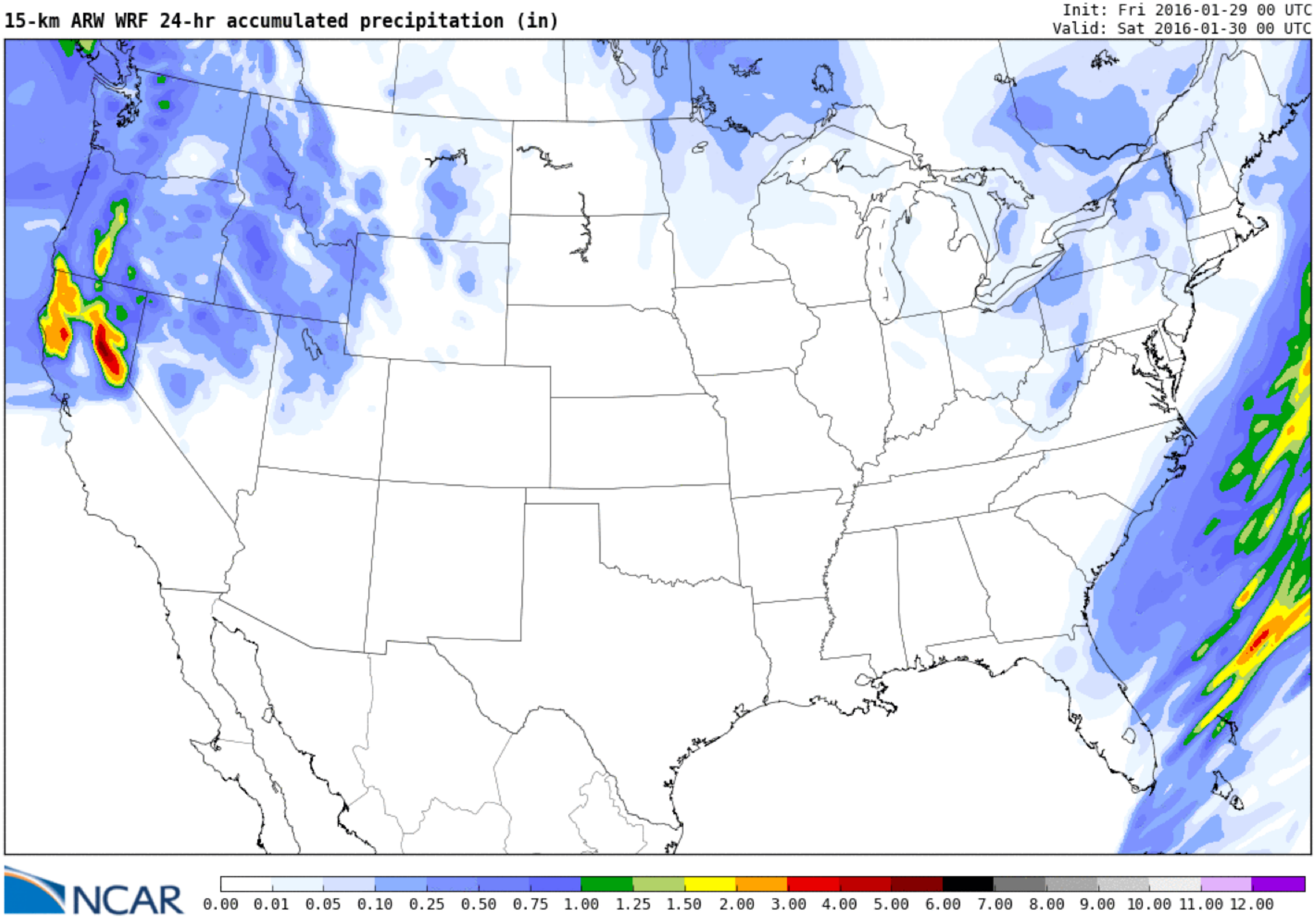
Real-Time WRF/WRFDA-hybrid analysis/forecast over CONUS



NCAR WRF/WRFDA Initialized: 00 UTC Fri 29 Jan 2016

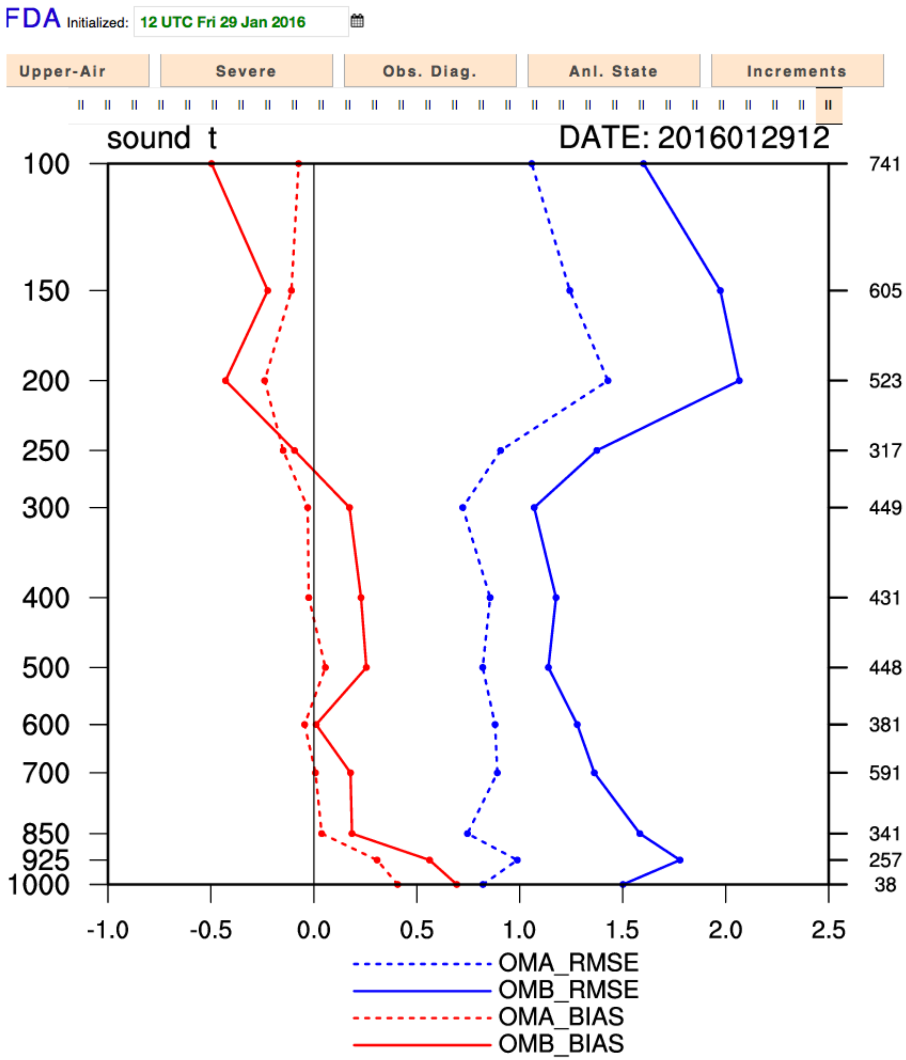
Surface/Precip | Upper-Air | Severe | Obs. Diag. | Anl. State | Increments

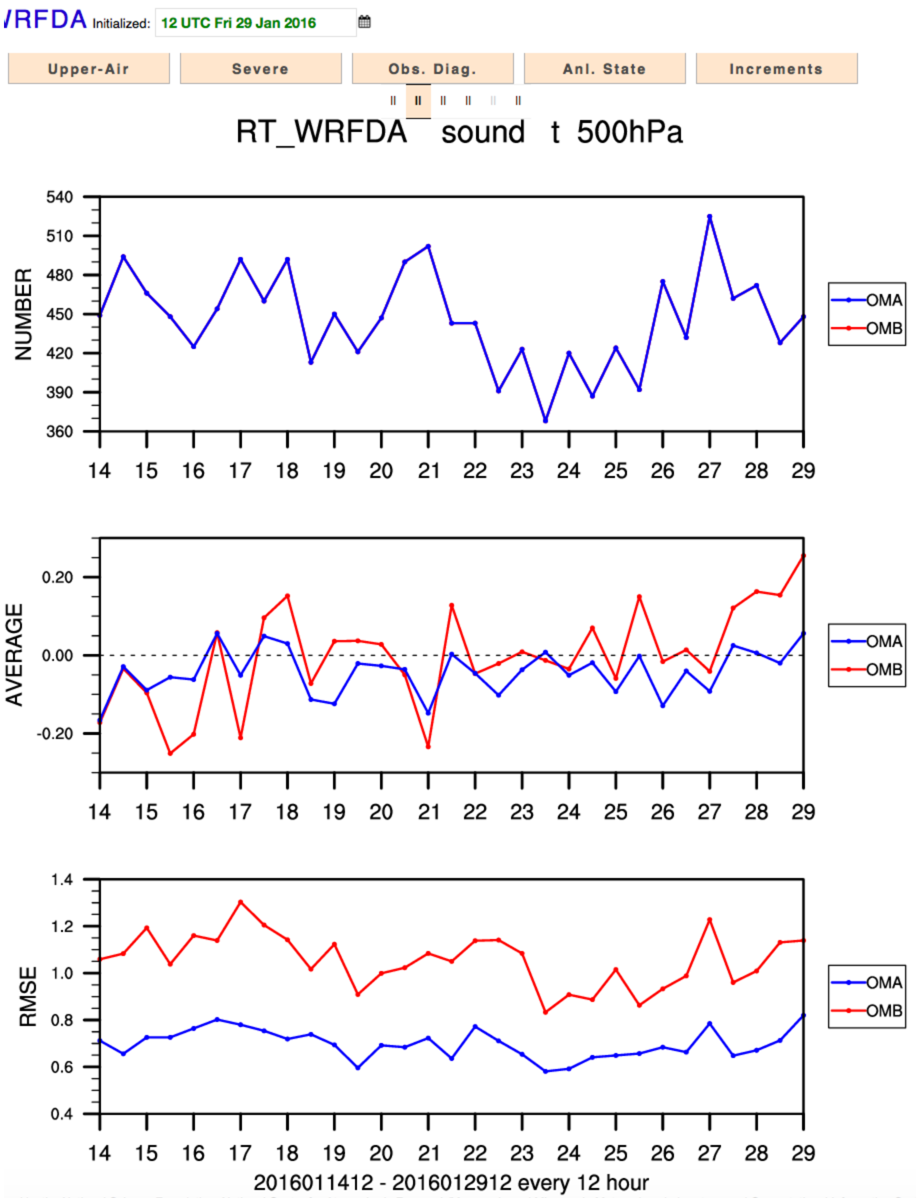
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 **24** 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48



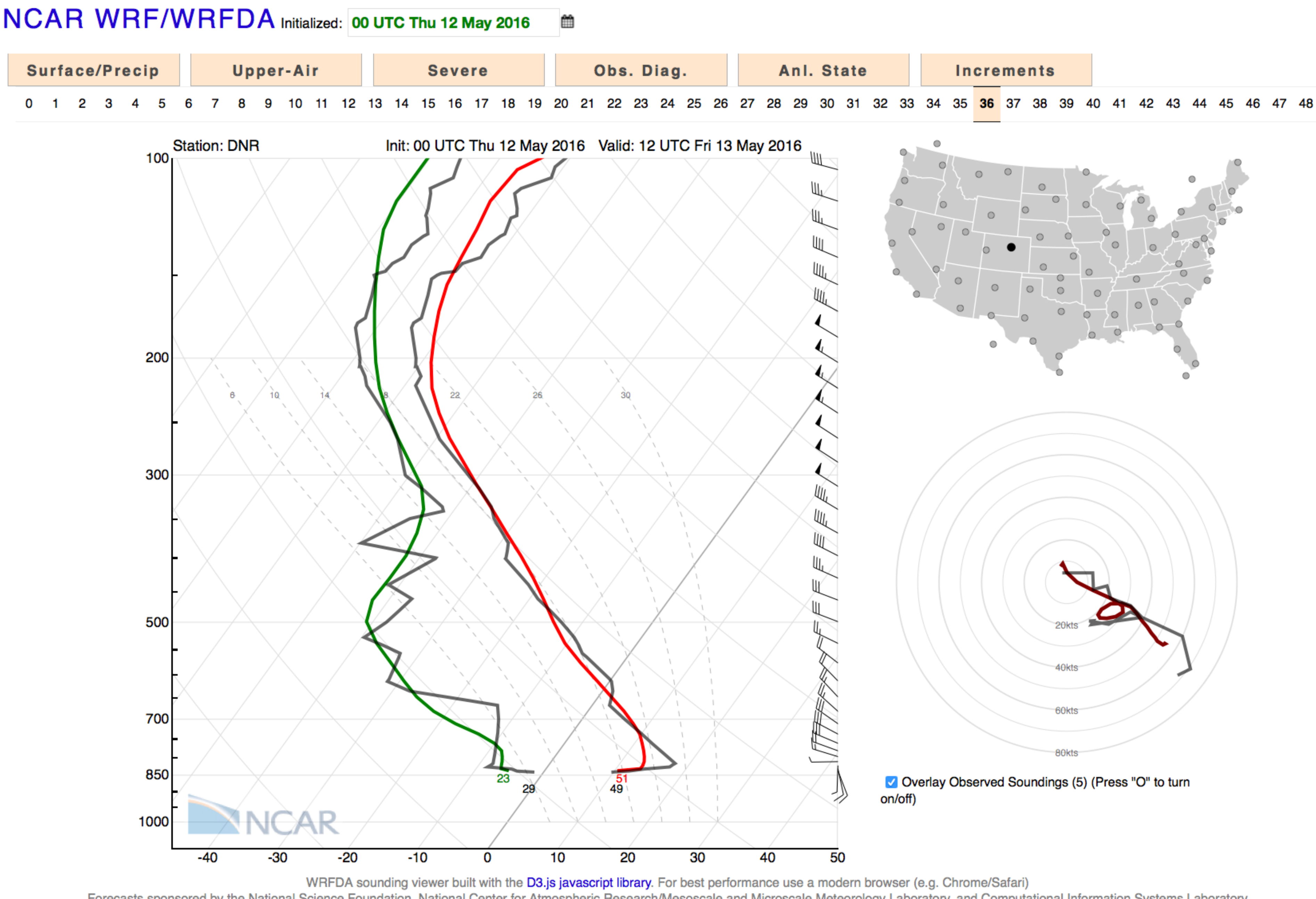
NCAR

Forecasts sponsored by the National Science Foundation, National Center for Atmospheric Research/Mesoscale and Microscale Meteorology Laboratory, and Computational Information Systems Laboratory


FDA Initialized: 12 UTC Fri 29 Jan 2016
Upper-Air
Severe
Obs. Diag.
Anl. State
Increments
sound t
DATE: 2016012912
OMA_RMSE
OMB_RMSE
OMA_BIAS
OMB_BIAS


WRFDA Initialized: 12 UTC Fri 29 Jan 2016
Upper-Air
Severe
Obs. Diag.
Anl. State
Increments
RT_WRFDA sound t 500hPa
NUMBER
AVERAGE
RMSE
OMA
OMB
2016011412 - 2016012912 every 12 hour

Real-time WRF/WRFDA forecast



Ongoing work: Variational Bias Correction of Aircraft T

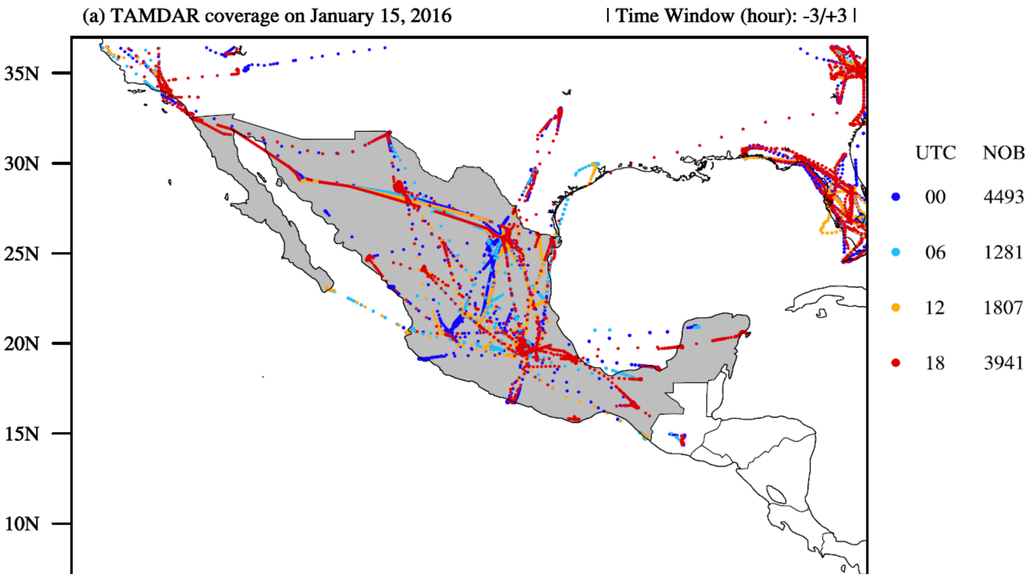
$$J(\mathbf{x},\boldsymbol{\beta}) = (\mathbf{x}-\mathbf{x}_\mathrm{b})^\mathrm{T}\mathbf{B}_\mathrm{x}^{-1}(\mathbf{x}-\mathbf{x}_\mathrm{b}) + (\boldsymbol{\beta}-\boldsymbol{\beta}_\mathrm{b})^\mathrm{T}\mathbf{B}_\beta^{-1}(\boldsymbol{\beta}-\boldsymbol{\beta}_\mathrm{b}) + (\mathbf{y}-H[\mathbf{x},\mathbf{y},\boldsymbol{\beta}])^\mathrm{T}\mathbf{R}^{-1}(\mathbf{y}-H[\mathbf{x},\mathbf{y},\boldsymbol{\beta}])$$

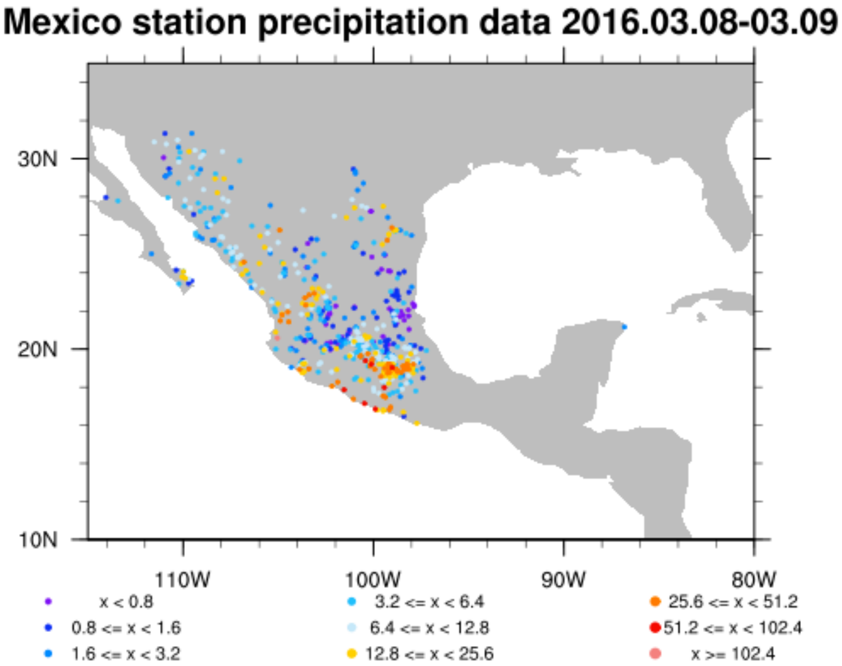
$$\tilde{H}(\mathbf{x},\mathbf{y},\boldsymbol{\beta}) = H(\mathbf{x}) - b(\mathbf{y},\boldsymbol{\beta})$$

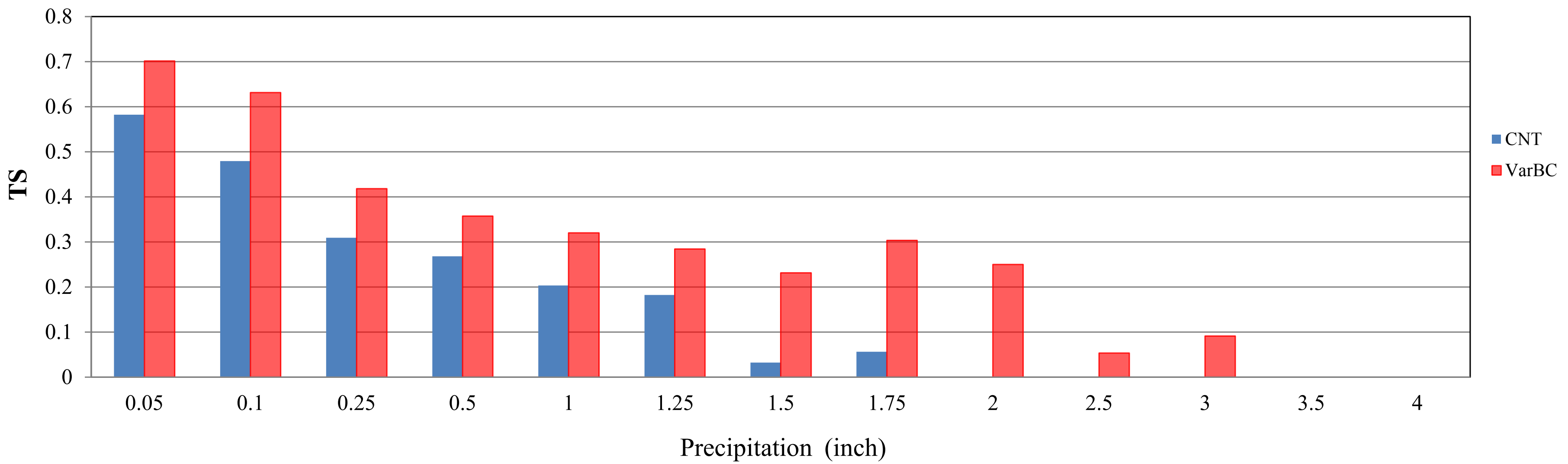
$$b(\mathbf{y},\boldsymbol{\beta}) = \sum_{n=0}^{N}\beta_n p_n(\mathbf{y}) = \begin{cases} \beta_{0+} + \beta_{1+} w & \text{if} \quad w > 0 \\ \beta_{0-} + \beta_{1-} w & \text{if} \quad w < 0 \end{cases}$$

$w = \dfrac{\mathrm{dp}}{\mathrm{dt}}$, β **is updated in cost function each cycle and written in parameter table.**

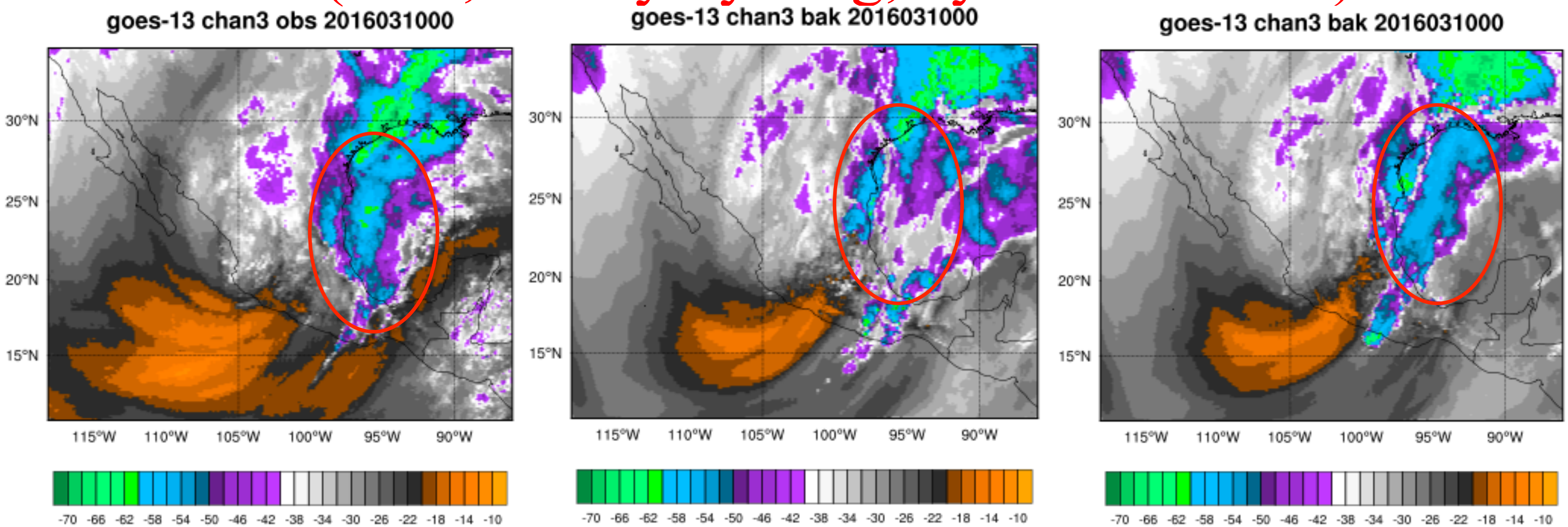
Impact of Aircraft T VarBC on rainfall forecast



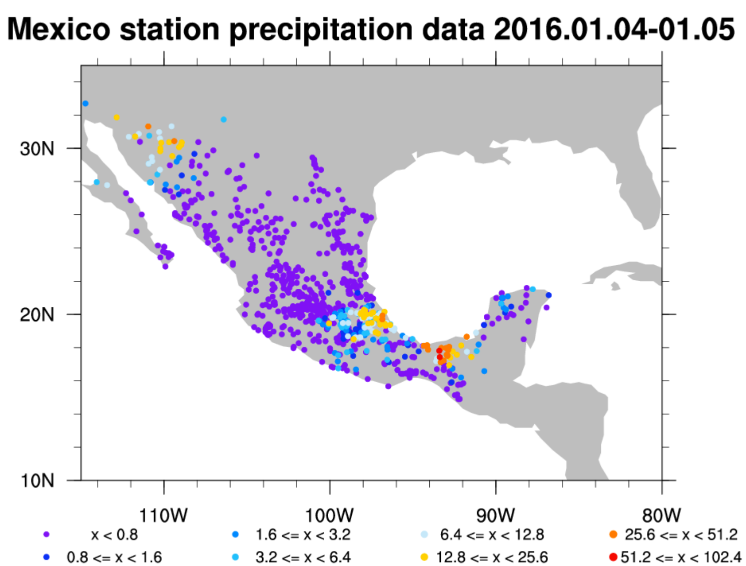


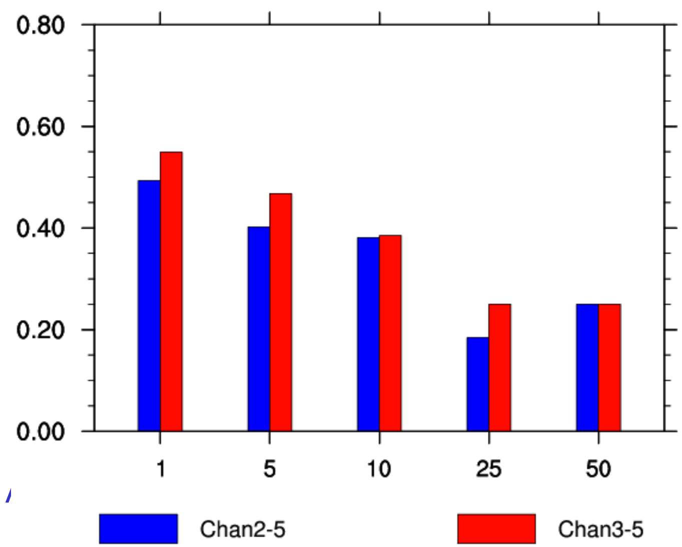


GEOS imager radiance DA at convection-permitting scale (4km, hourly-cycling, hybrid-3DVAR)



obs **No-GOES DA** **GOES-DA**





Other ongoing work

- Implemented Hybrid-4DEnVAR
  - Improving computing efficiency
- Continue developing Multi-Resolution Incremental 4DVAR (MRI-4DVAR)
- Continue developing cloudy radiance/product DA
- Improving surface data assimilation
- Improving radar DA
  - Adding divergence constraint
  - Assimilation of non-rain data
- WRFPlus-Chem & WRFDA-Chem
  - CU Boulder.

Last Remarks

- We welcome contributions from external users/ developers.
  - Contact wrfhelp@ucar.edu or directly email to me liuz@ucar.edu for contributing back your code

- We maintain a WRFDA-related publications list, please inform us your papers to be included
  - http://www2.mmm.ucar.edu/wrf/users/wrfda/publications.html