中央氣象局全球數值預報之資料同化系統的初步測試

2013/05/13 天氣分析研討會

沈彥志[1] 陳登舜[1,4] 陳御群[3] 曹伶伶[1] 楊舒芝[4] 曾建翰[1]
陳雯美[1] 馮欽賜[2] 周俊宇[1] 陳建河[2] 楊明仁[4]

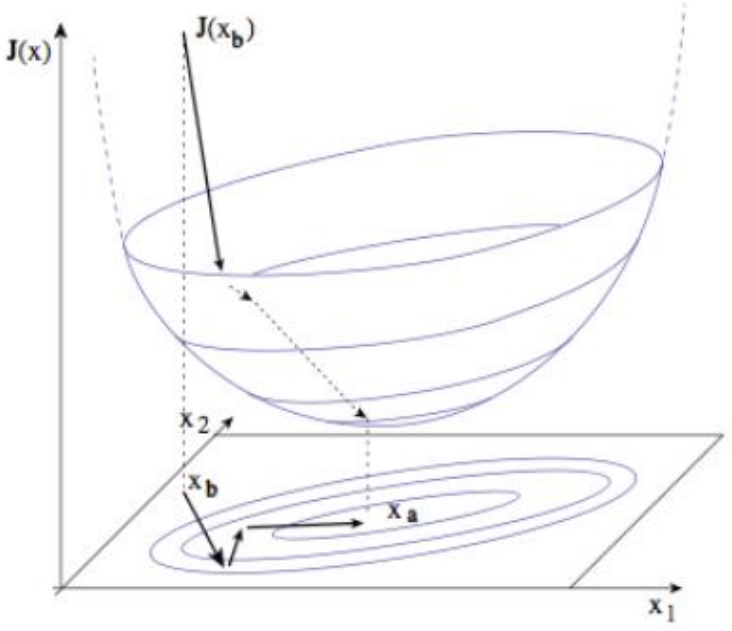
中央氣象局氣象科技研究中心[1] 中央氣象局氣象資訊中心[2]
臺灣颱風洪水研究中心[3] 國立中央大學大氣科學系[4]

3D VARIATIONAL DATA ASSIMILATION

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x_b})^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x_b}) + \frac{1}{2}(\mathbf{y_o}-\mathbf{H}(\mathbf{x}))^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y_o}-\mathbf{H}(\mathbf{x})) = \mathbf{J_b} + \mathbf{J_o}$$

$$\nabla J(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x_b}) - \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y_o}-\mathbf{H}(\mathbf{x}))$$

- Pros:
  1. Easy to add various constraints to cost function
  2. Easy to implement and maintain
- Cons:
  1. Static background error covariance
  2. All observations are treated as instantaneous

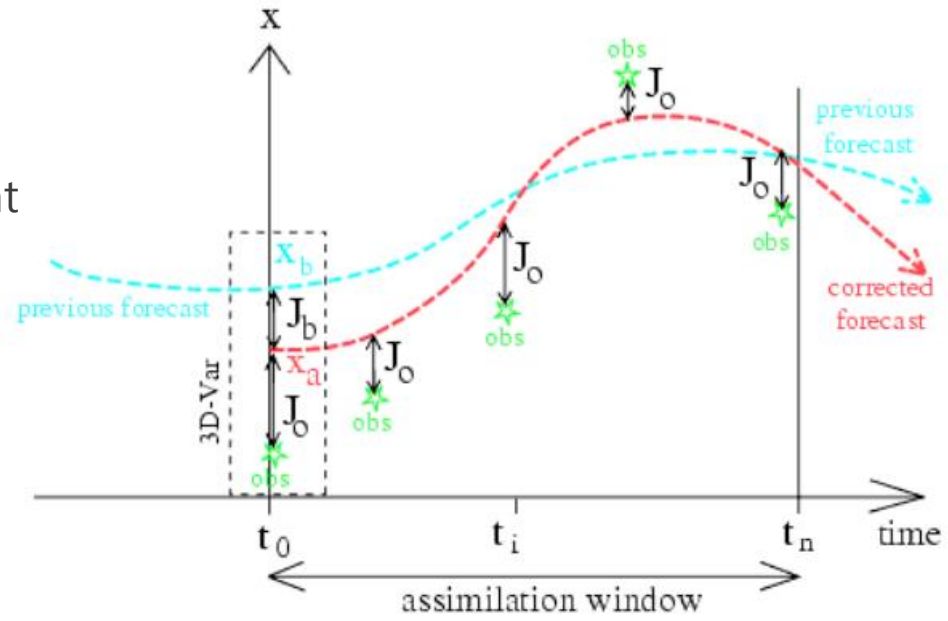


From Bouttier and Courtier (1999)

4D VARIATIONAL DATA ASSIMILATION

$$J = \frac{1}{2}\left(\mathbf{x_0} - \mathbf{x_b}\right)^{\mathrm{T}} \mathbf{B}^{-1}\left(\mathbf{x_0} - \mathbf{x_b}\right) + \frac{1}{2}\sum_{\mathbf{i=0}}^{\mathbf{n}}\left(\mathbf{H_i M_i}\left[\mathbf{x_0}\right] - \mathbf{y_i}\right)^{\mathrm{T}} \mathbf{R_i}^{-1}\left(\mathbf{H_i M_i}\left[\mathbf{x_0}\right] - \mathbf{y_i}\right)$$

- Pros:
  1. Implicit flow-dependent B
  2. Assimilate observations at right time
- Cons:
  1. Hard to develop and maintain adjoint model $\mathbf{M}^{\mathrm{T}}$
  2. The assumption of "perfect model"



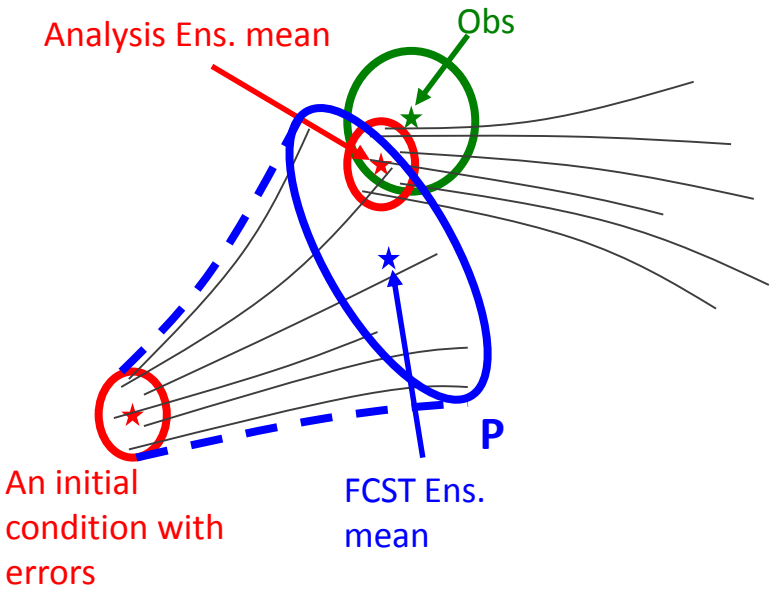
From Tom Auligne

ENSEMBLE KALMAN FILTER

$$\mathbf{x_a} = \mathbf{x_f} + \mathbf{K}\left(\mathbf{y_o} - \mathbf{H}\mathbf{x_f}\right) \qquad \mathbf{P^f H^T} = \overline{\mathbf{x'^f}\left(\mathbf{H}\mathbf{x'^f}\right)^{\mathbf{T}}} \equiv \left(1/n-1\right)\sum_{i=1}^{n} \mathbf{x_i'^f}\left(\mathbf{H}\mathbf{x_i'^f}\right)^{\mathbf{T}}$$

$$\mathbf{K} = \mathbf{P^f H^T}\left(\mathbf{H}\mathbf{P^f}\mathbf{H^T} + \mathbf{R}\right)^{-1} \qquad \mathbf{H}\mathbf{P^f}\mathbf{H^T} = \overline{\mathbf{H}\mathbf{x'^f}\left(\mathbf{H}\mathbf{x'^f}\right)^{\mathbf{T}}} \equiv \left(1/n-1\right)\sum_{i=1}^{n} \mathbf{H}\mathbf{x_i'^f}\left(\mathbf{H}\mathbf{x_i'^f}\right)^{\mathbf{T}}$$

- Pros:
  1. Flow-dependent B
  2. No need to linearize and get adjoint model of **H** and **M**
- Cons:
  - Under sampling:
    1. Filter divergence: Need to inflate the members
    2. Spurious noise: Need Localization



From S.-C.Yang

HYBRID DATA ASSIMILATION

$$J(\mathbf{x}_{\mathrm{f}}^{'},\alpha)=\beta_{\mathrm{f}}\frac{1}{2}(\mathbf{x}_{\mathrm{f}}^{'})^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}_{\mathrm{f}}^{'})+\beta_{\mathrm{e}}\frac{1}{2}(\alpha)^{\mathrm{T}}\mathbf{L}^{-1}(\alpha)+\frac{1}{2}(\mathbf{y}_{\mathrm{o}}^{'}-\mathbf{H}\mathbf{x}_{\mathrm{t}}^{'})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}_{\mathrm{o}}^{'}-\mathbf{H}\mathbf{x}_{\mathrm{t}}^{'})$$

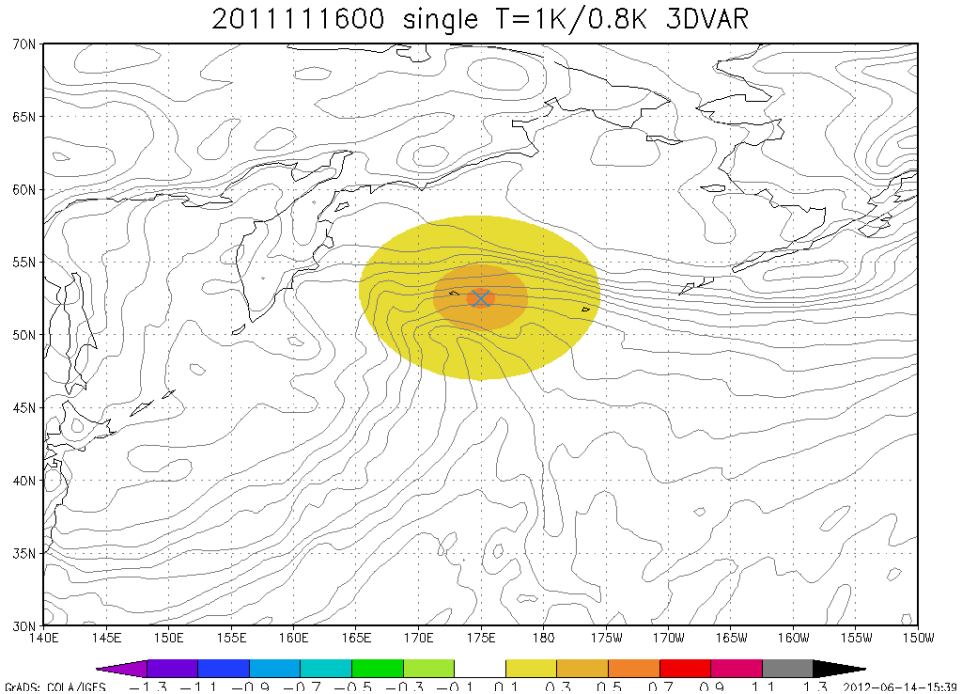
$$\mathbf{x}_{\mathrm{t}}^{'}=\mathbf{x}_{\mathrm{f}}^{'}+\sum_{k=1}^{K}(\alpha_k\circ\mathbf{x}_k^{\mathrm{e}}) \qquad \frac{1}{\beta_{\mathrm{f}}}+\frac{1}{\beta_{\mathrm{e}}}=1$$

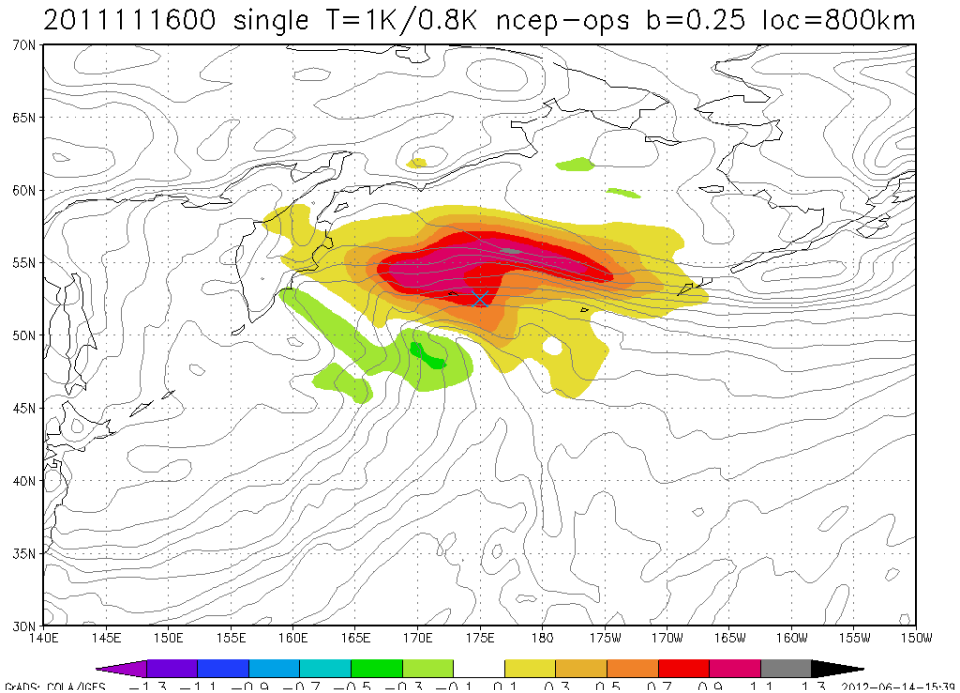
- Benefit from use of flow dependent ensemble covariance instead of static B
- Avoid "filter divergence" problem in EnKF
- Less expensive then EnKF (small ensemble)
- Easy framework to add various constraints (make balance to initial conditions)
- Use of various existing capabilities in VAR
- Now hybrid systems are developed or operated …
  - NCEP, ECMWF, NCAR, CMC, UK … CWB

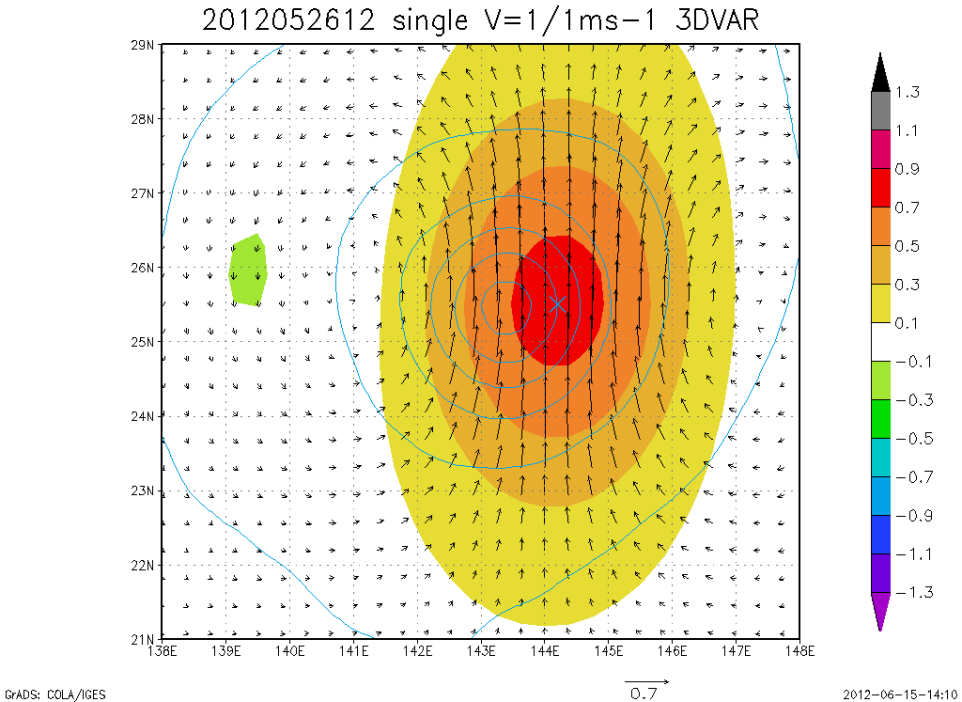
Lorenc, 2003
Buehner, 2005
Wang et al., 2007, 2008
Kleist et al., 2011
etc.

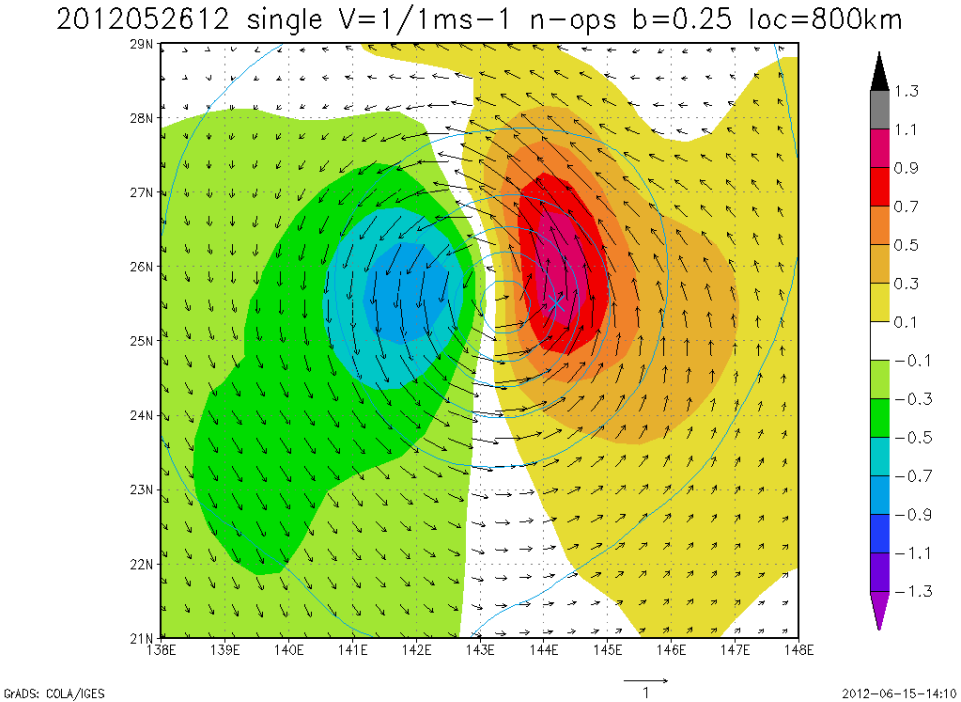
EXPERIMENT DESIGN: SINGLE OBS. TEST

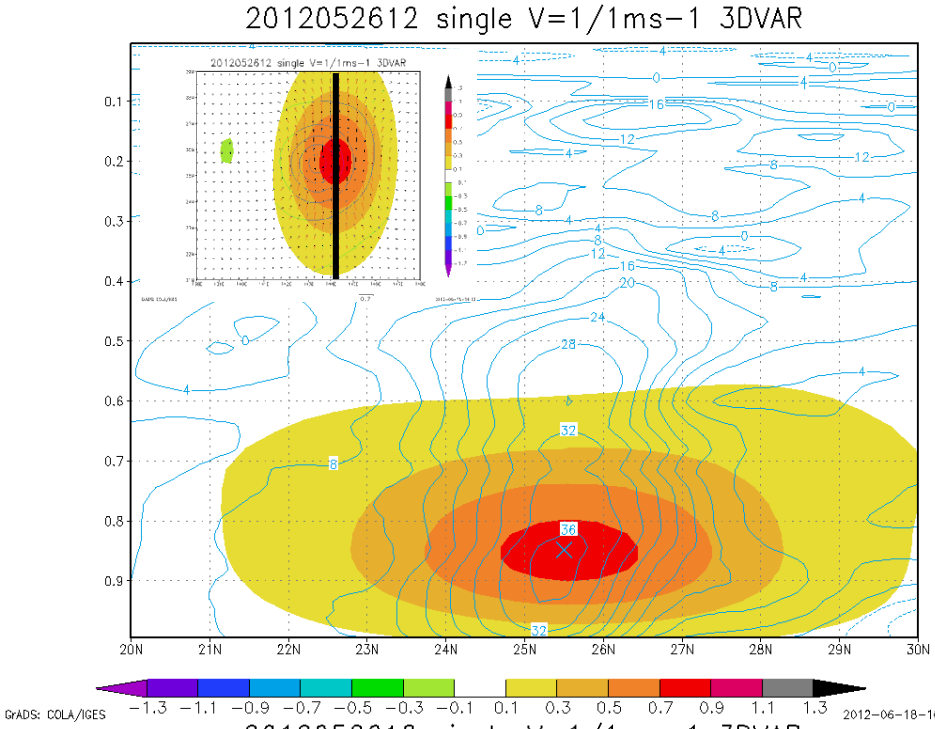
- Data assimilation system: NCEP GSI operational version, may 2012
- Due to lack of EnKF system, the hybrid ensemble members are time lagged members: the forecasts of 6~72HR, per 6hr, total are 12 members
  - 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66, 72HR forecasts
- Single observation test 1 on 2011/11/16,00
  - Innovation: 1(K); observation error: 0.8(K)
  - Location: 175W, 52.5N at 850 hPa
  - β_1=0.25 and β_2=0.75, localization: 800 km
- Single observation test 2 on 2012/05/26,12
  - Innovation (V): 1(m/s); observation error: 1.0(m/s)
  - Location: 144.2E, 25.5N at 850 hPa
  - β_1=0.25 and β_2=0.75 , localization: 800 km

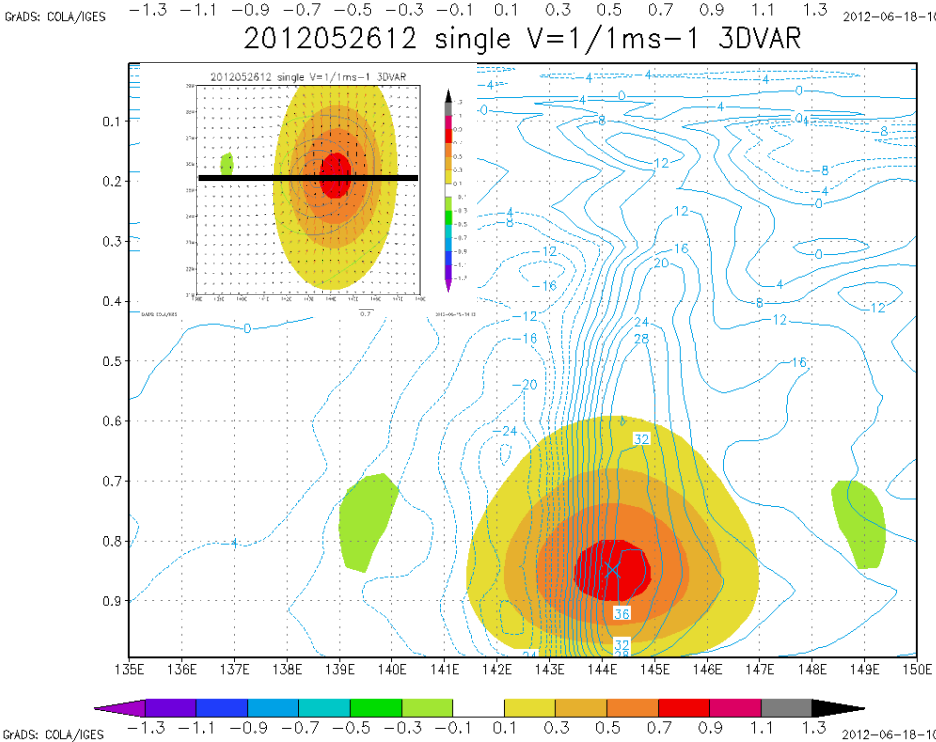

2011111600 single T=1K/0.8K 3DVAR
70N
65N
60N
55N
50N
45N
40N
35N
30N
140E 145E 150E 155E 160E 165E 170E 175E 180 175W 170W 165W 160W 155W 150W
−1.3 −1.1 −0.9 −0.7 −0.5 −0.3 −0.1 0.1 0.3 0.5 0.7 0.9 1.1 1.3
GrADS: COLA/IGES
2012-06-14-15:39

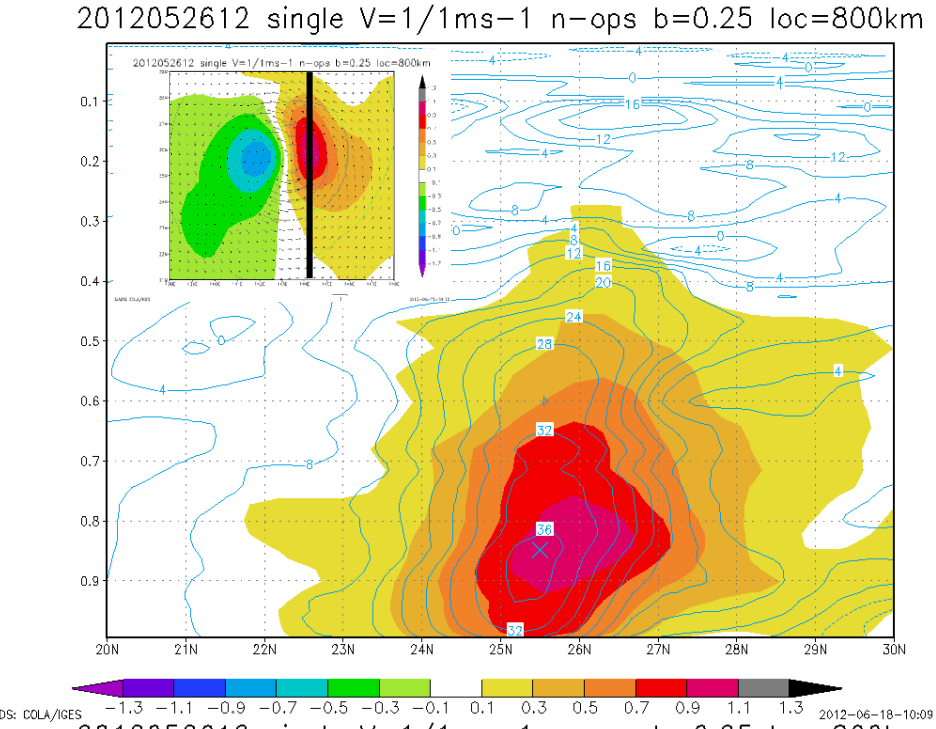

2011111600 single T=1K/0.8K ncep−ops b=0.25 loc=800km
70N
65N
60N
55N
50N
45N
40N
35N
30N
140E 145E 150E 155E 160E 165E 170E 175E 180 175W 170W 165W 160W 155W 150W
−1.3 −1.1 −0.9 −0.7 −0.5 −0.3 −0.1 0.1 0.3 0.5 0.7 0.9 1.1 1.3
GrADS: COLA/IGES
2012-06-14-15:39

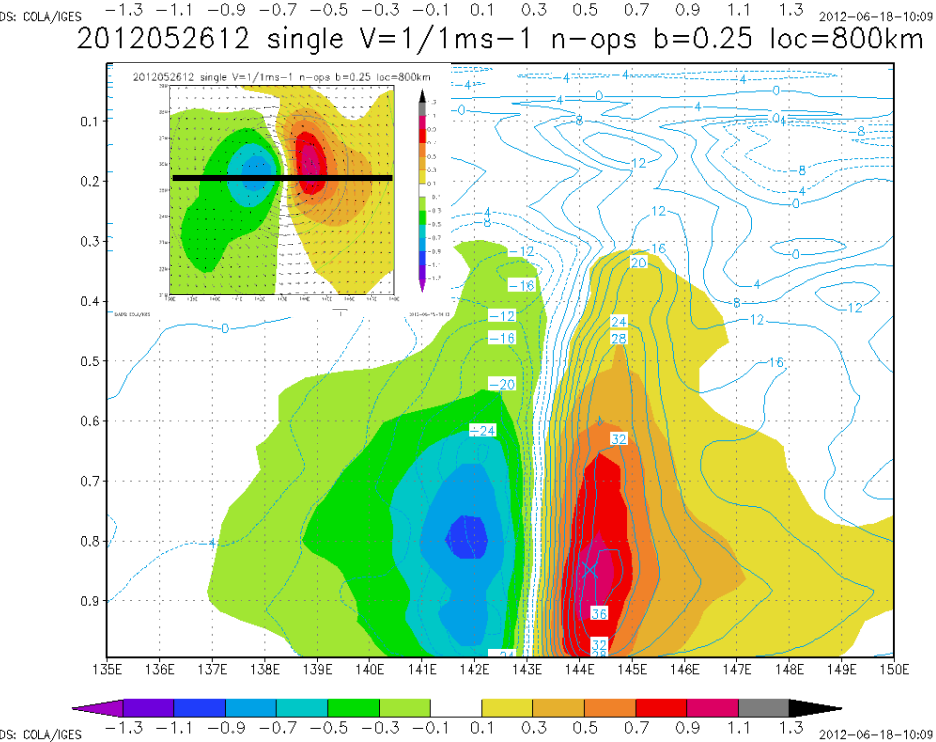

2012052612 single V=1/1ms−1 3DVAR
29N
28N
27N
26N
25N
24N
23N
22N
21N
138E 139E 140E 141E 142E 143E 144E 145E 146E 147E 148E
1.3 1.1 0.9 0.7 0.5 0.3 0.1 −0.1 −0.3 −0.5 −0.7 −0.9 −1.1 −1.3
0.7
GrADS: COLA/IGES
2012-06-15-14:10


2012052612 single V=1/1ms−1 n−ops b=0.25 loc=800km
29N
28N
27N
26N
25N
24N
23N
22N
21N
138E 139E 140E 141E 142E 143E 144E 145E 146E 147E 148E
1.3 1.1 0.9 0.7 0.5 0.3 0.1 −0.1 −0.3 −0.5 −0.7 −0.9 −1.1 −1.3
1
GrADS: COLA/IGES
2012-06-15-14:10


2012052612 single V=1/1ms-1 3DVAR
20N 21N 22N 23N 24N 25N 26N 27N 28N 29N 30N
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
−1.3 −1.1 −0.9 −0.7 −0.5 −0.3 −0.1 0.1 0.3 0.5 0.7 0.9 1.1 1.3
GrADS: COLA/IGES
2012-06-18-10:10

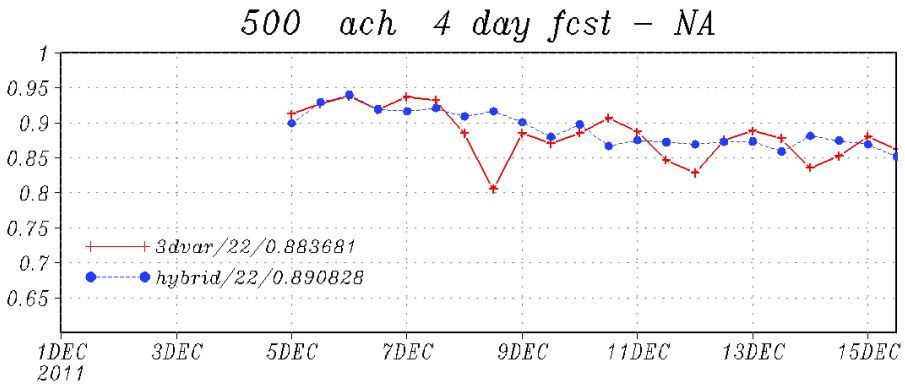

2012052612 single V=1/1ms-1 3DVAR
135E 136E 137E 138E 139E 140E 141E 142E 143E 144E 145E 146E 147E 148E 149E 150E
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
−1.3 −1.1 −0.9 −0.7 −0.5 −0.3 −0.1 0.1 0.3 0.5 0.7 0.9 1.1 1.3
GrADS: COLA/IGES
2012-06-18-10:10

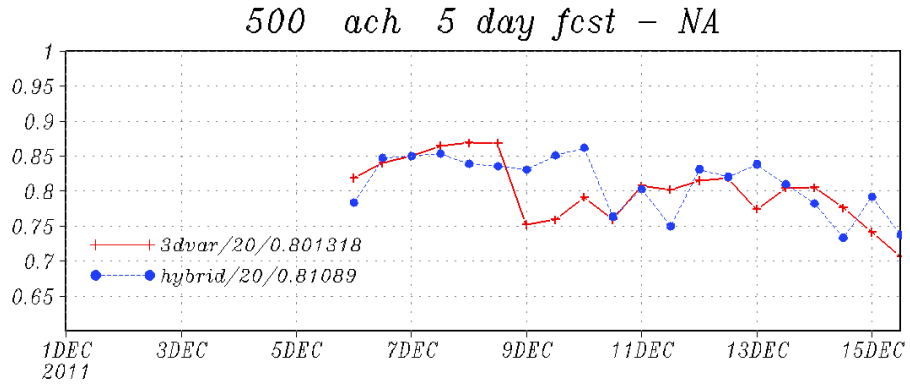

2012052612 single V=1/1ms-1 n-ops b=0.25 loc=800km
20N 21N 22N 23N 24N 25N 26N 27N 28N 29N 30N
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
−1.3 −1.1 −0.9 −0.7 −0.5 −0.3 −0.1 0.1 0.3 0.5 0.7 0.9 1.1 1.3
GrADS: COLA/IGES
2012-06-18-10:09

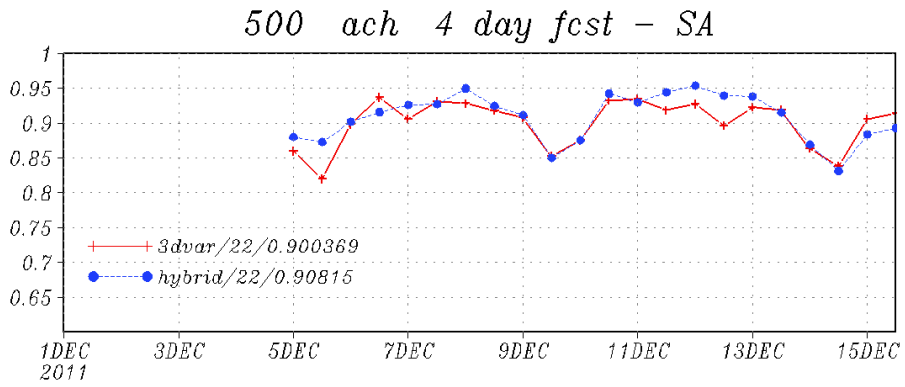

2012052612 single V=1/1ms-1 n-ops b=0.25 loc=800km
135E 136E 137E 138E 139E 140E 141E 142E 143E 144E 145E 146E 147E 148E 149E 150E
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
−1.3 −1.1 −0.9 −0.7 −0.5 −0.3 −0.1 0.1 0.3 0.5 0.7 0.9 1.1 1.3
GrADS: COLA/IGES
2012-06-18-10:09

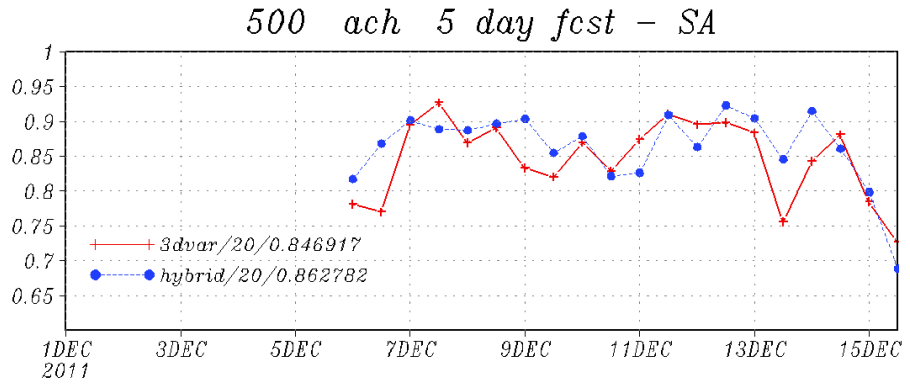
EXPERIMENT DESIGN: PARALLEL

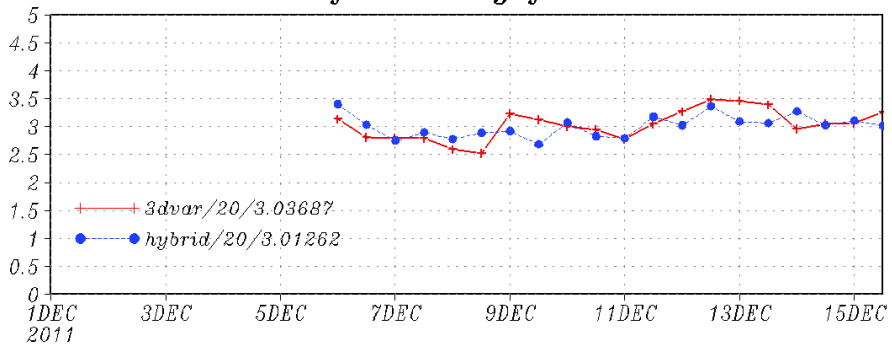
- Data assimilation system: NCEP GSI operational version, may 2012
- Due to lack of EnKF system, the hybrid ensemble members are time lagged members: the forecasts of 6~72HR, per 6hr, total are 12 members
  - 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66, 72HR forecasts
- Experiment period: 2011120100 ~ 2011121512
- Control (3dvar): Use NCEP operational GSI and do only 3dvar analysis
- Experiment (hybrid): Same as control exp., but turn on the hybrid capability. The members are time lagged members.
- β_1=0.5 and β_2=0.5, localization: 800 km

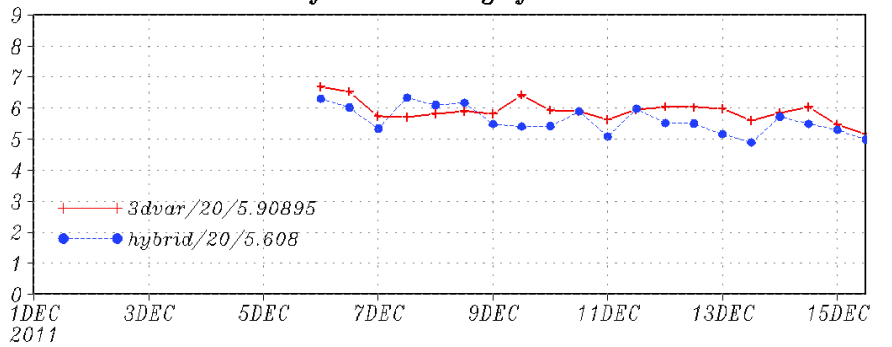

500 ach 4 day fcst - NA
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
3dvar/22/0.883681
hybrid/22/0.890828
1DEC
2011
3DEC
5DEC
7DEC
9DEC
11DEC
13DEC
15DEC

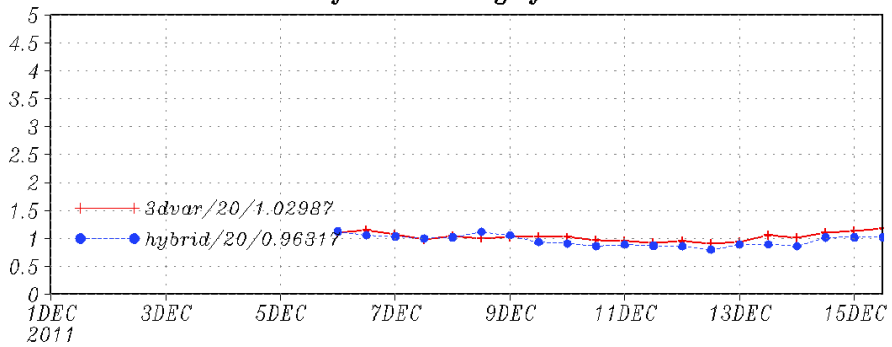

500 ach 5 day fcst - NA
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
3dvar/20/0.801318
hybrid/20/0.81089
1DEC
2011
3DEC
5DEC
7DEC
9DEC
11DEC
13DEC
15DEC

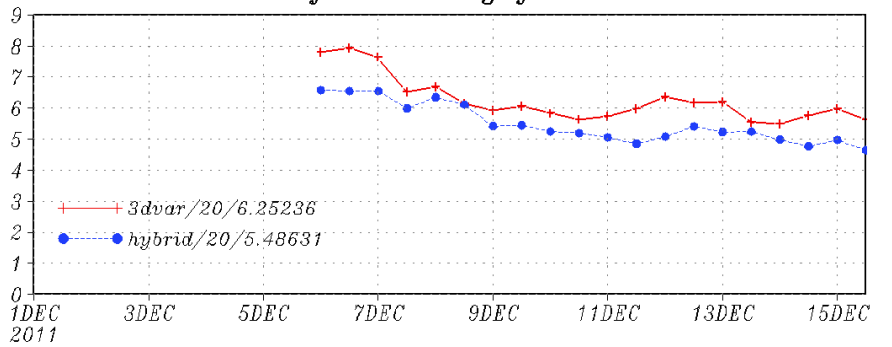

500 ach 4 day fcst - SA
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
3dvar/22/0.900369
hybrid/22/0.90815
1DEC
2011
3DEC
5DEC
7DEC
9DEC
11DEC
13DEC
15DEC

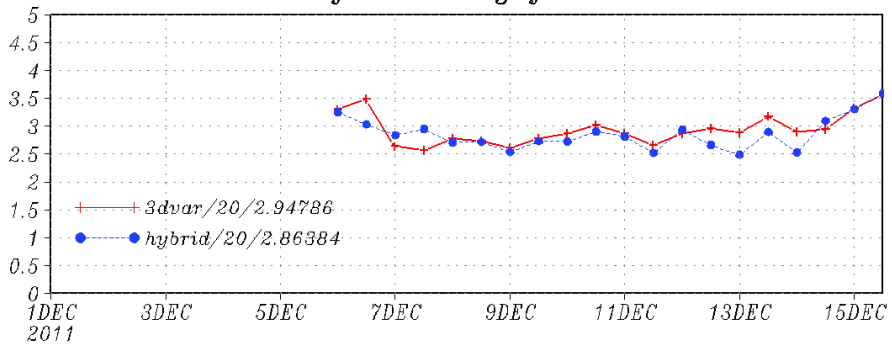

500 ach 5 day fcst - SA
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
3dvar/20/0.846917
hybrid/20/0.862782
1DEC
2011
3DEC
5DEC
7DEC
9DEC
11DEC
13DEC
15DEC

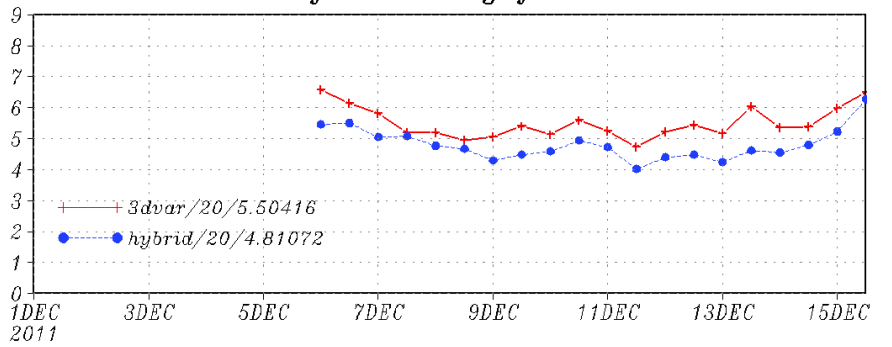

500 frt 5 day fcst - NA
3dvar/20/3.03687
hybrid/20/3.01262
1DEC 2011
3DEC
5DEC
7DEC
9DEC
11DEC
13DEC
15DEC


100 fru 5 day fcst - NA
3dvar/20/5.90895
hybrid/20/5.608
1DEC 2011
3DEC
5DEC
7DEC
9DEC
11DEC
13DEC
15DEC

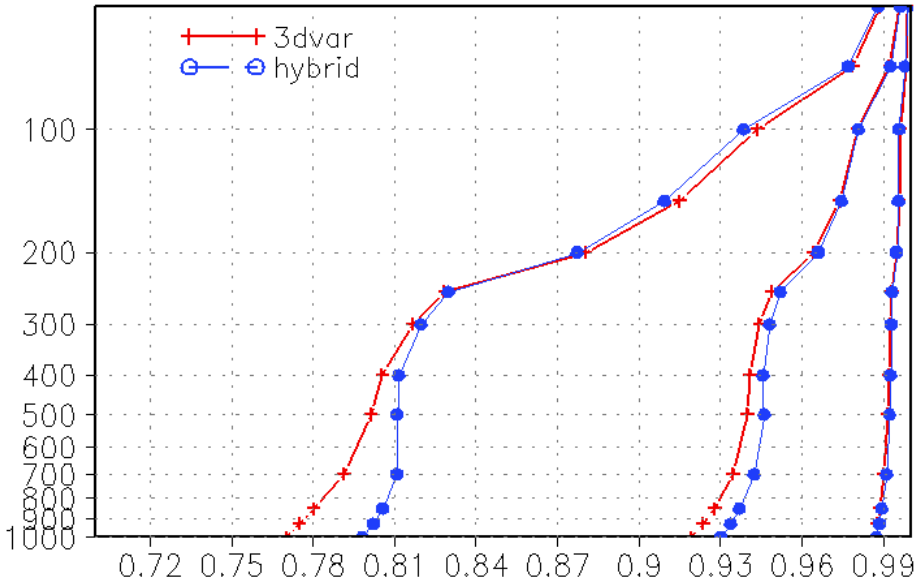

500 frt 5 day fcst - TP
3dvar/20/1.02987
hybrid/20/0.96317
1DEC 2011
3DEC
5DEC
7DEC
9DEC
11DEC
13DEC
15DEC


100 fru 5 day fcst - TP
3dvar/20/6.25236
hybrid/20/5.48631
1DEC 2011
3DEC
5DEC
7DEC
9DEC
11DEC
13DEC
15DEC

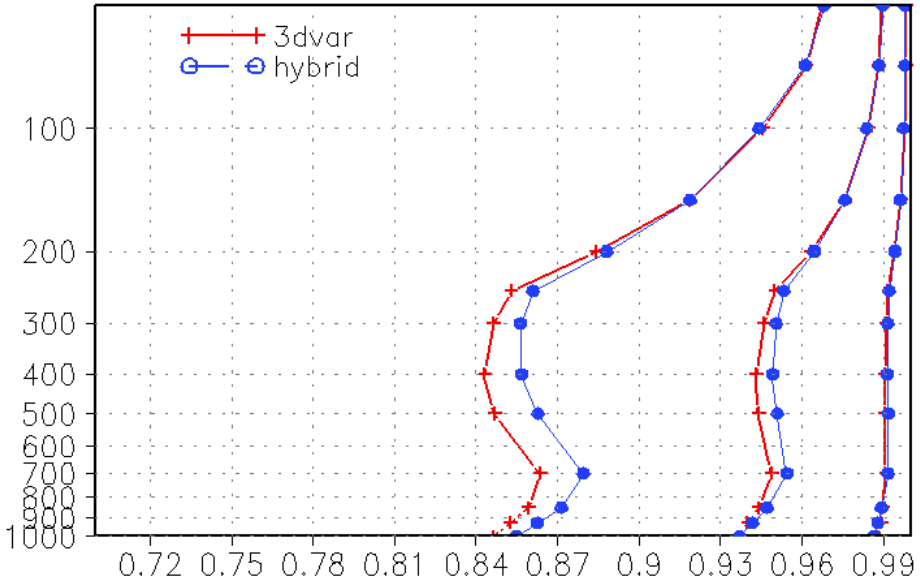

500 frt 5 day fcst - SA
3dvar/20/2.94786
hybrid/20/2.86384
1DEC 2011
3DEC
5DEC
7DEC
9DEC
11DEC
13DEC
15DEC


100 fru 5 day fcst - SA
3dvar/20/5.50416
hybrid/20/4.81072
1DEC 2011
3DEC
5DEC
7DEC
9DEC
11DEC
13DEC
15DEC

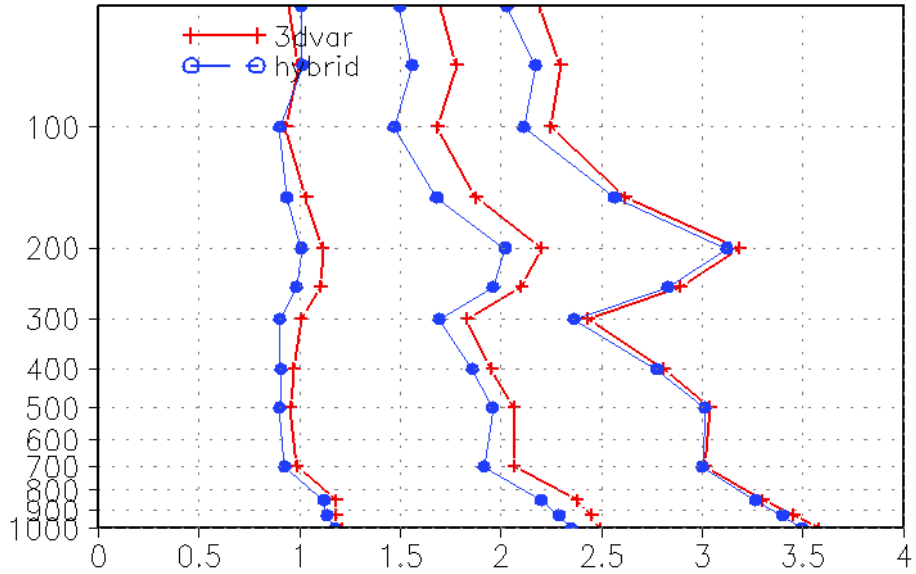
2011 Dec. AC score of Height, NH



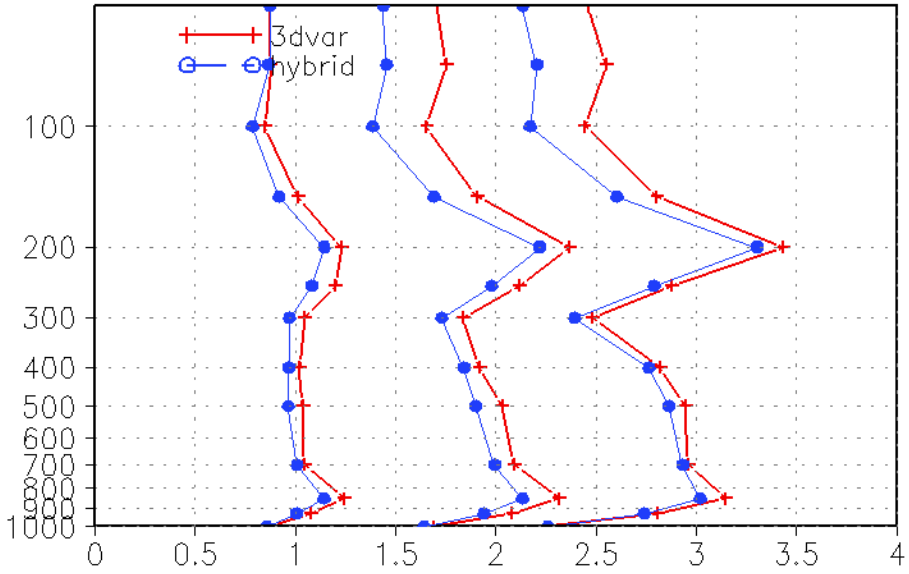
2011 Dec. AC score of Height, SH



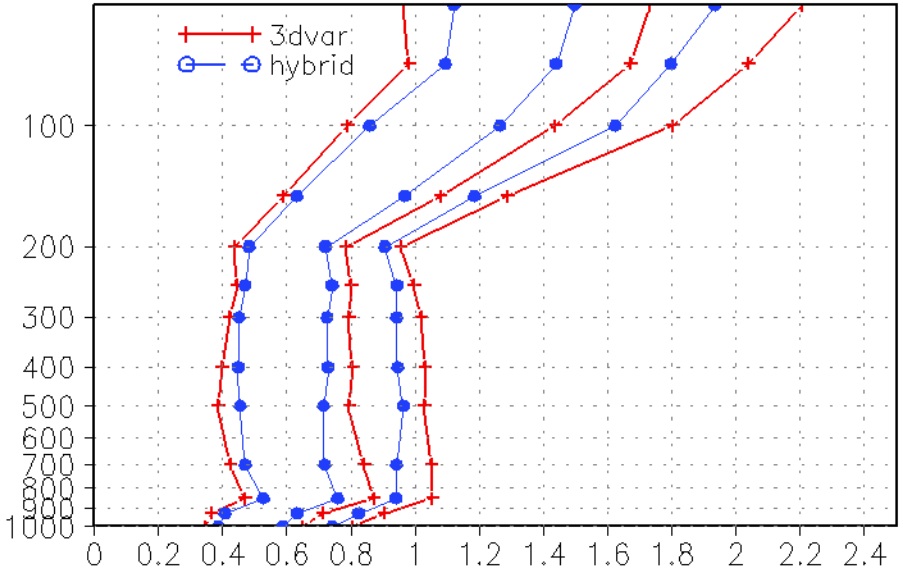
2011 Dec. RMSE of temperature, NH



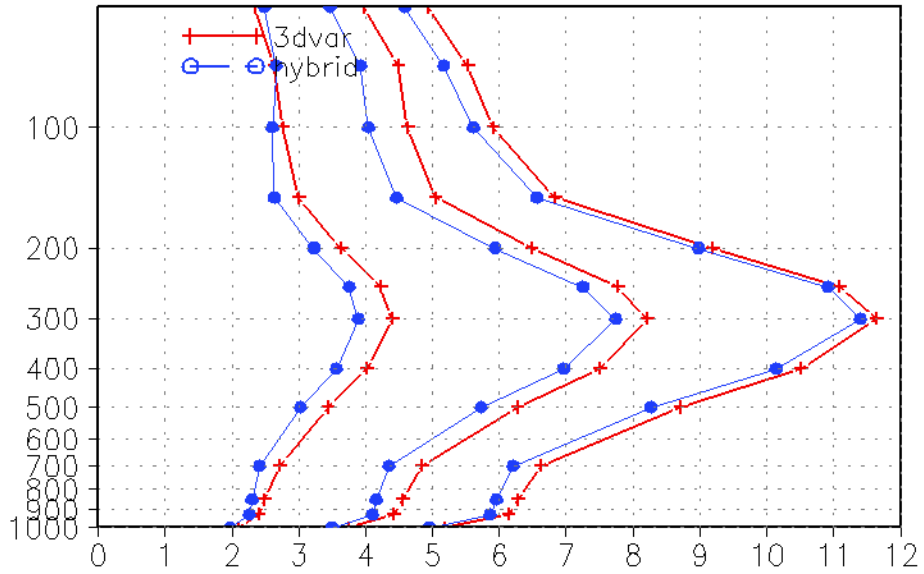
2011 Dec. RMSE of temperature, SH



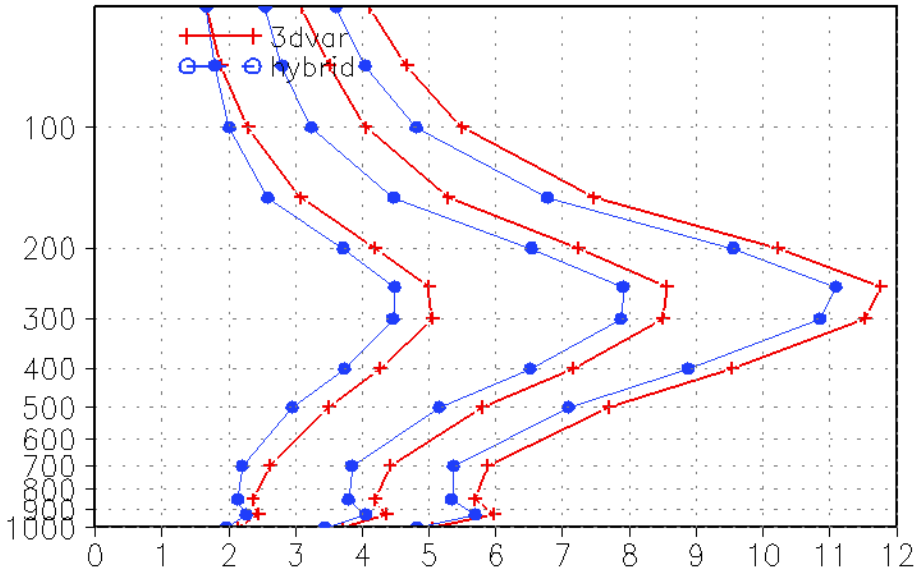
2011 Dec. RMSE of temperature, TR



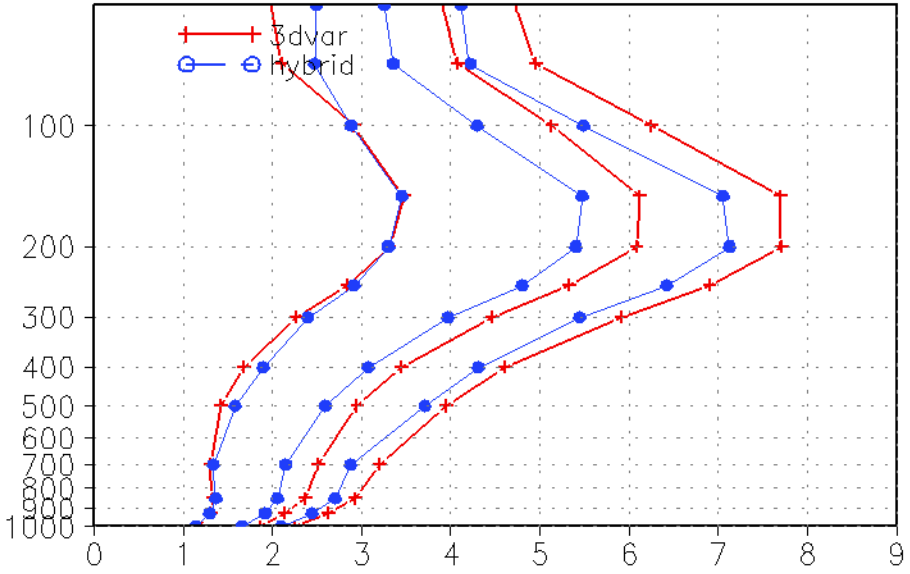
2011 Dec. RMSE of u-wind, NH



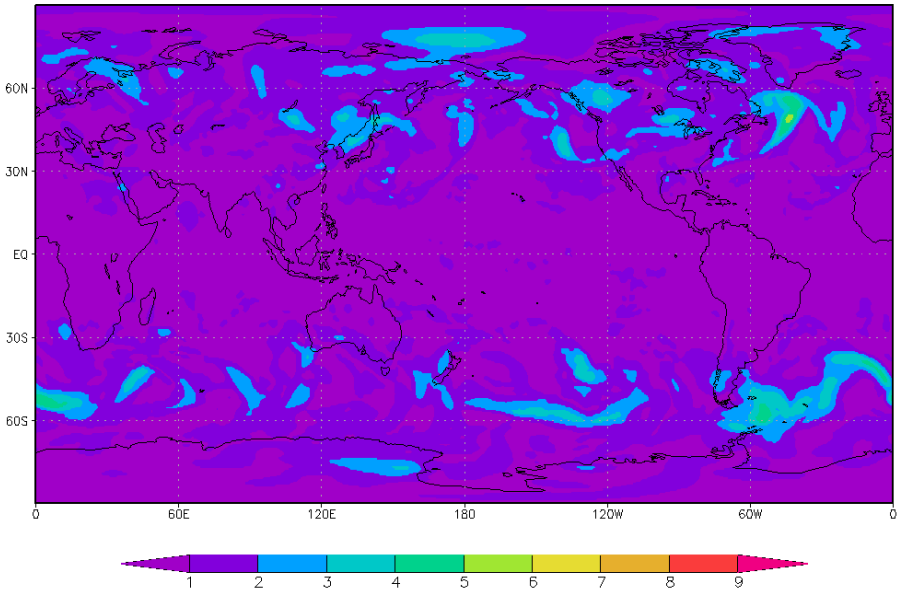
2011 Dec. RMSE of u-wind, SH



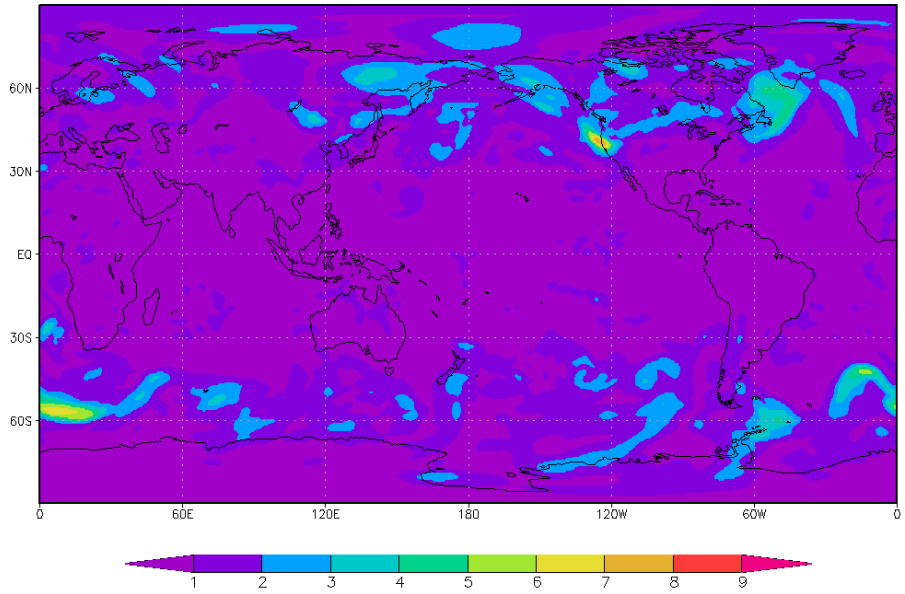
2011 Dec. RMSE of u-wind, TR



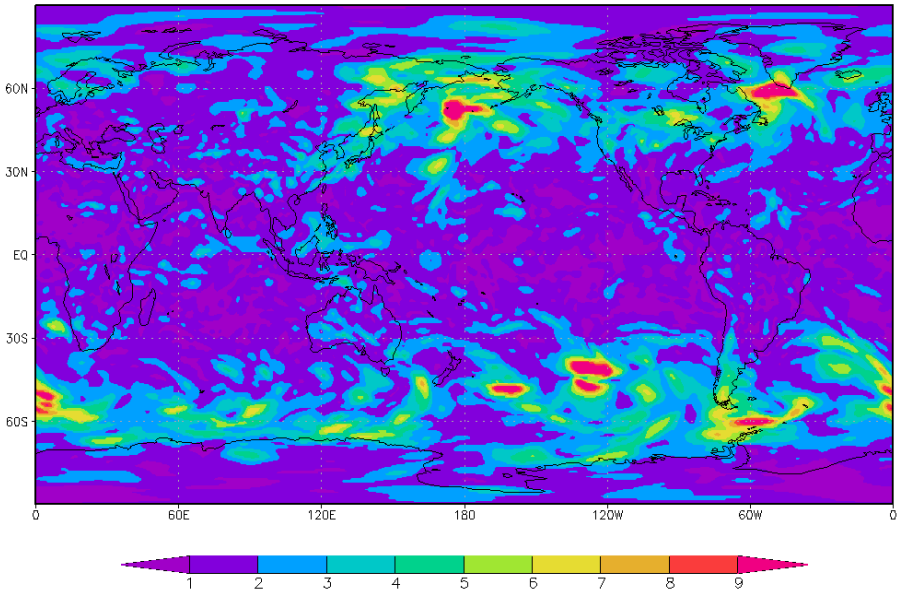
2011111900 T sig=11 spread timelag



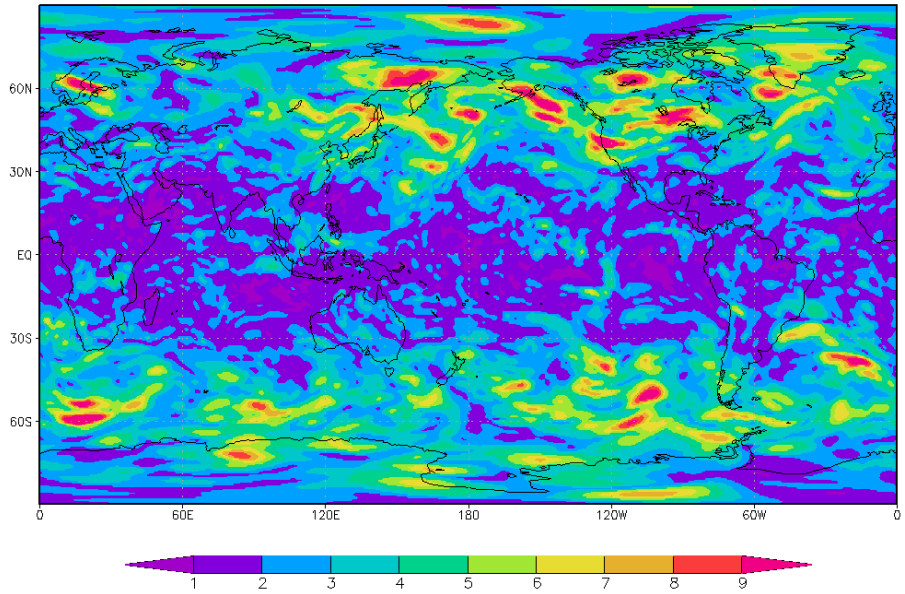
2011111900 T sig=21 spread timelag

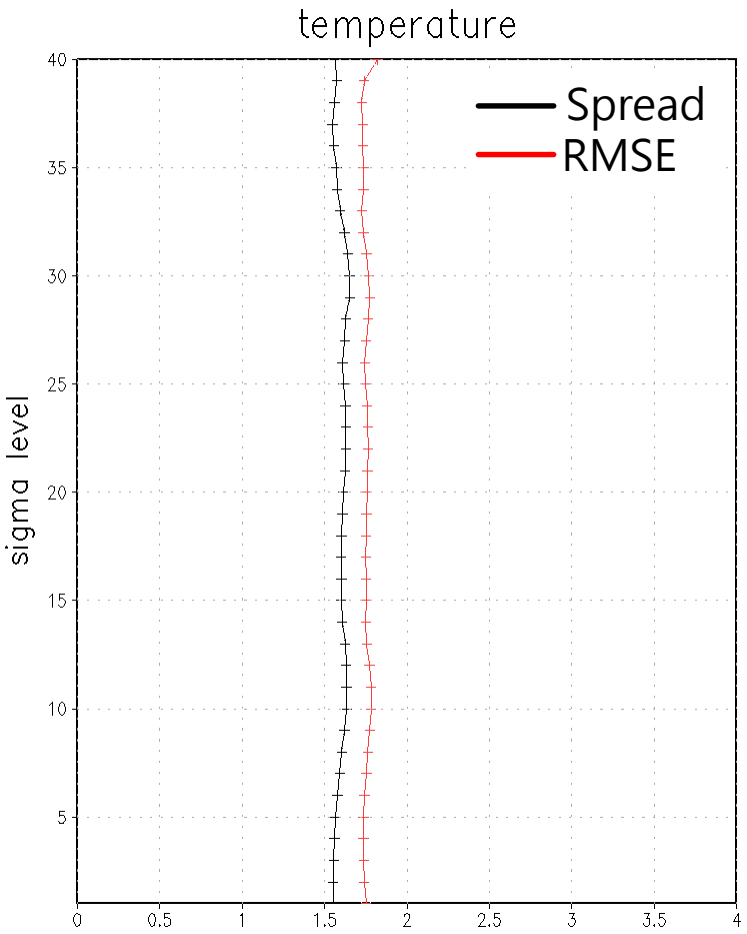


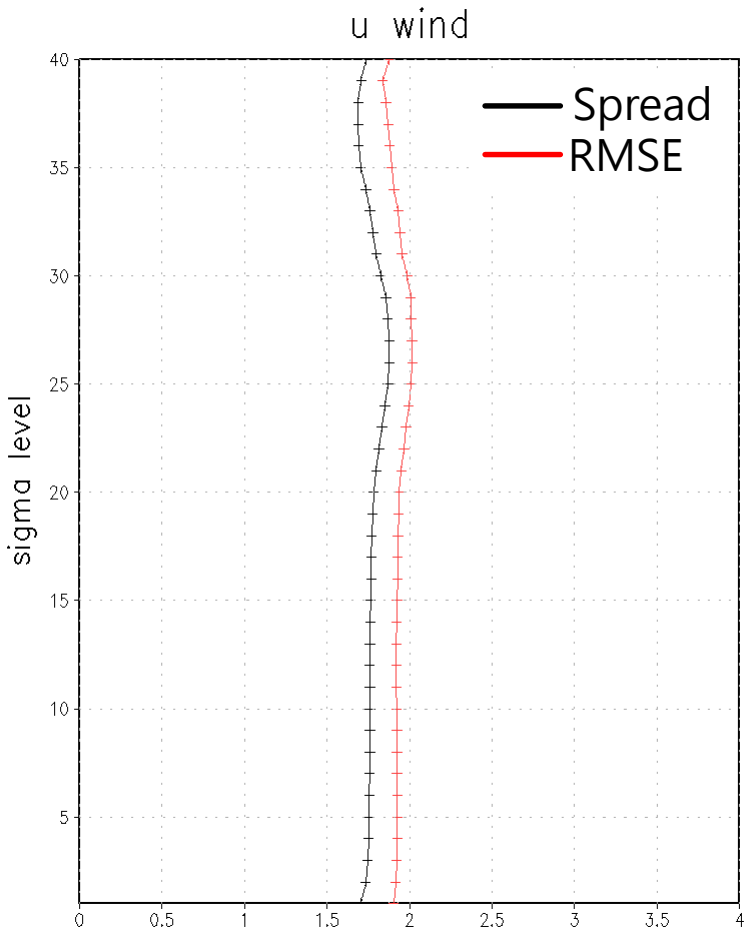
2011111900 U sig=11 spread timelag



2011111900 U sig=21 spread timelag

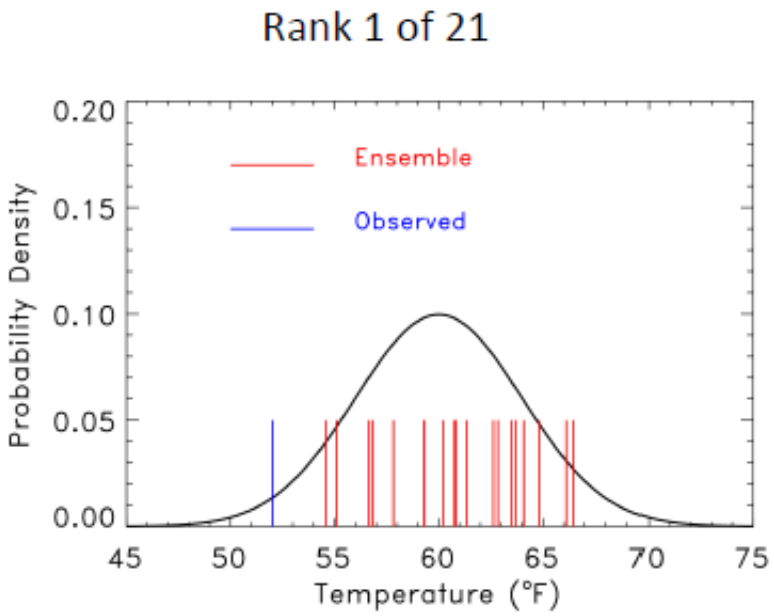


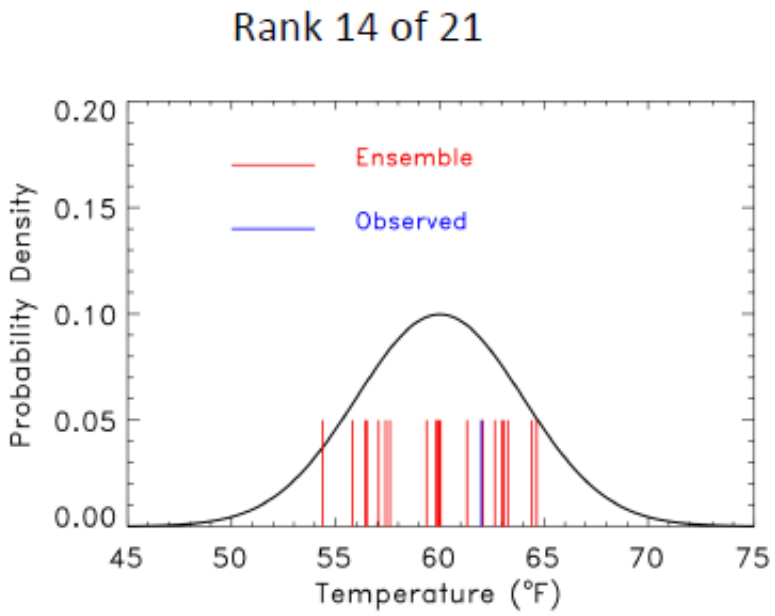

temperature
Spread
RMSE
sigma level
40
35
30
25
20
15
10
5
0
0.5
1
1.5
2
2.5
3
3.5
4

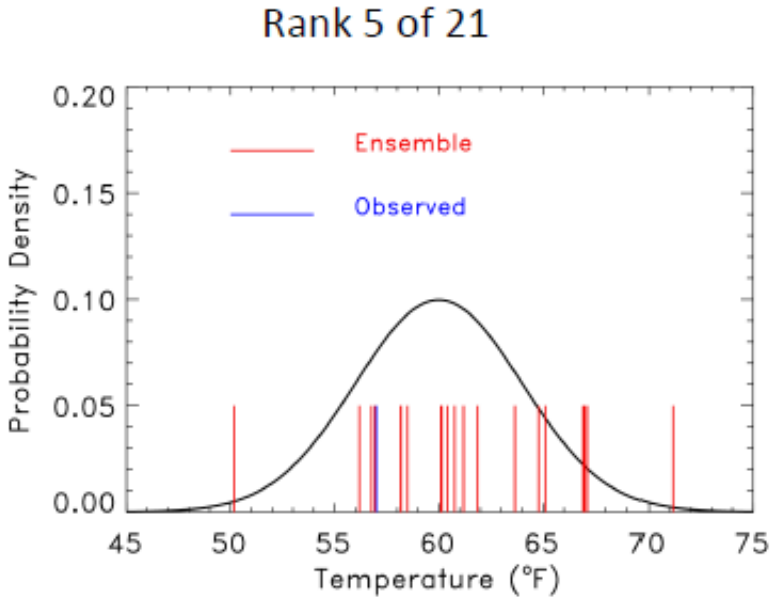

u wind
Spread
RMSE
sigma level
40
35
30
25
20
15
10
5
0
0.5
1
1.5
2
2.5
3
3.5
4

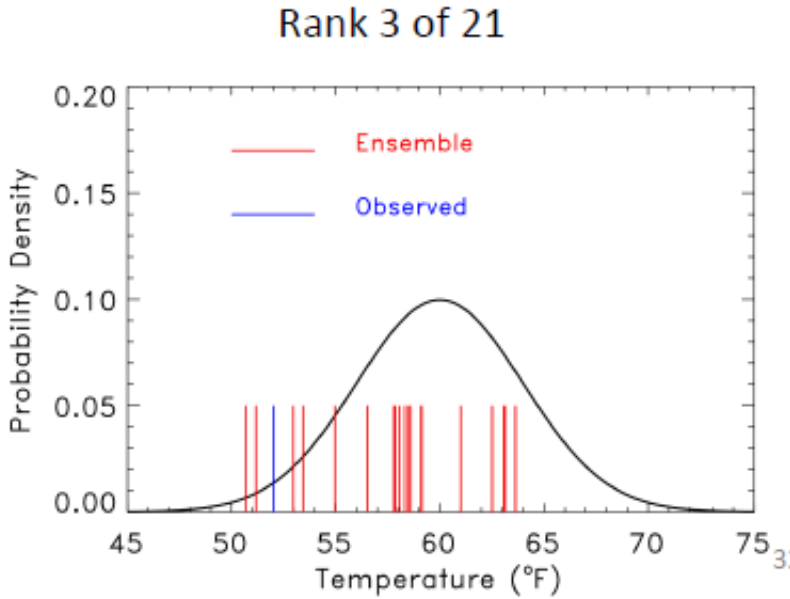
Rank histograms

Tom Hamill, 2011



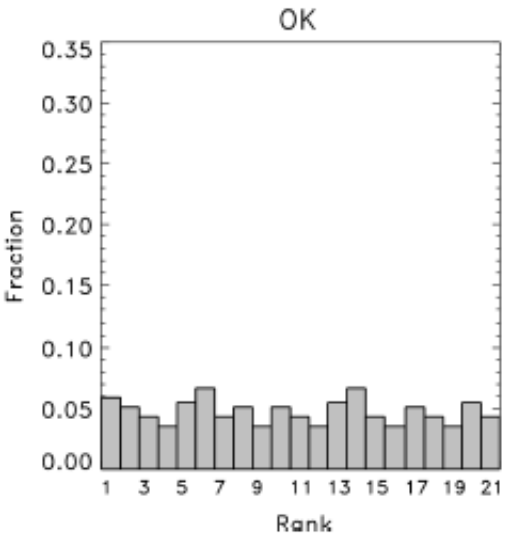




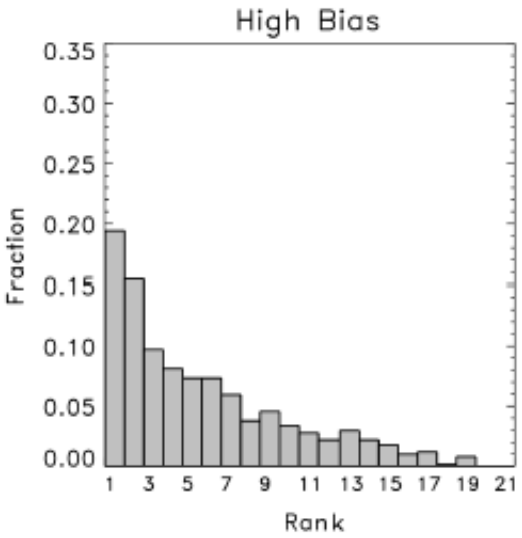


Rank histograms

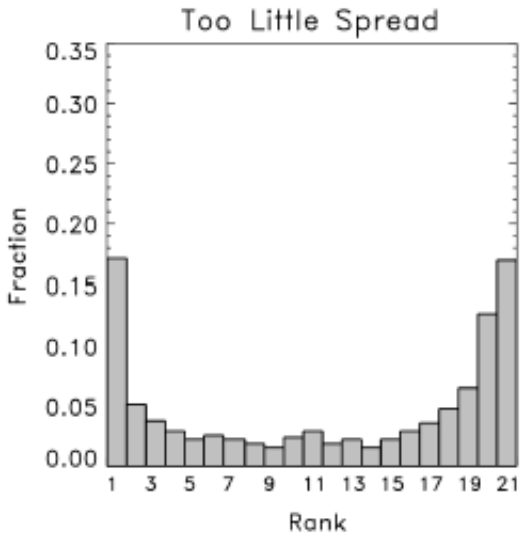
With **lots of samples** from many situations, can evaluate the characteristics of the ensemble.



Happens when observed is indistinguishable from any other member of the ensemble. Ensemble hopefully is reliable.



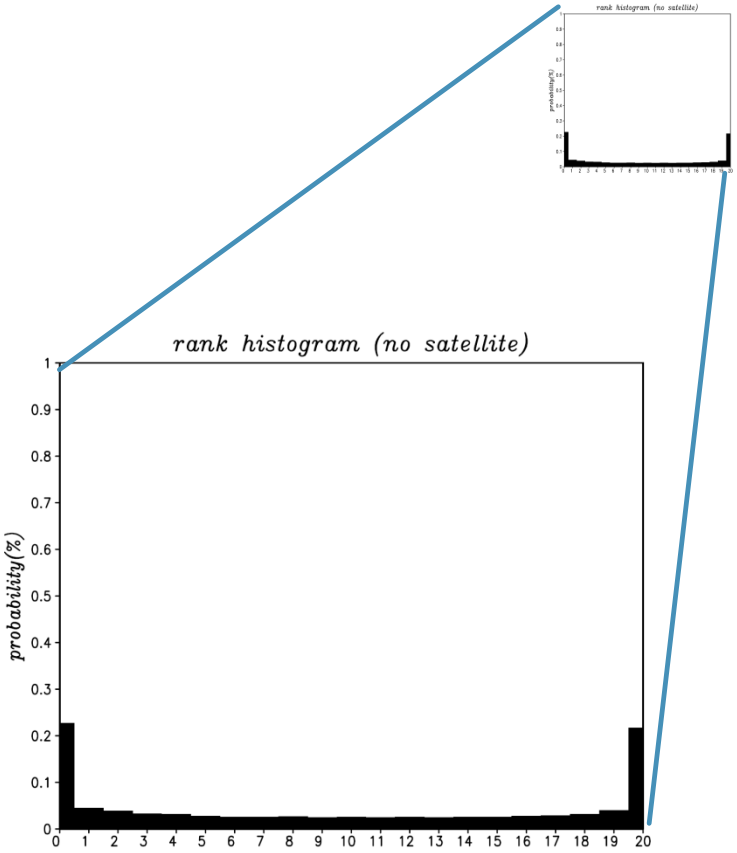
Happens when observed too commonly is lower than the ensemble members.



Happens when there are either some low and some high biases, or when the ensemble doesn't spread out enough.

ref: Hamill, *MWR*, March 2001

Indicate small spread but no bias



Time lag

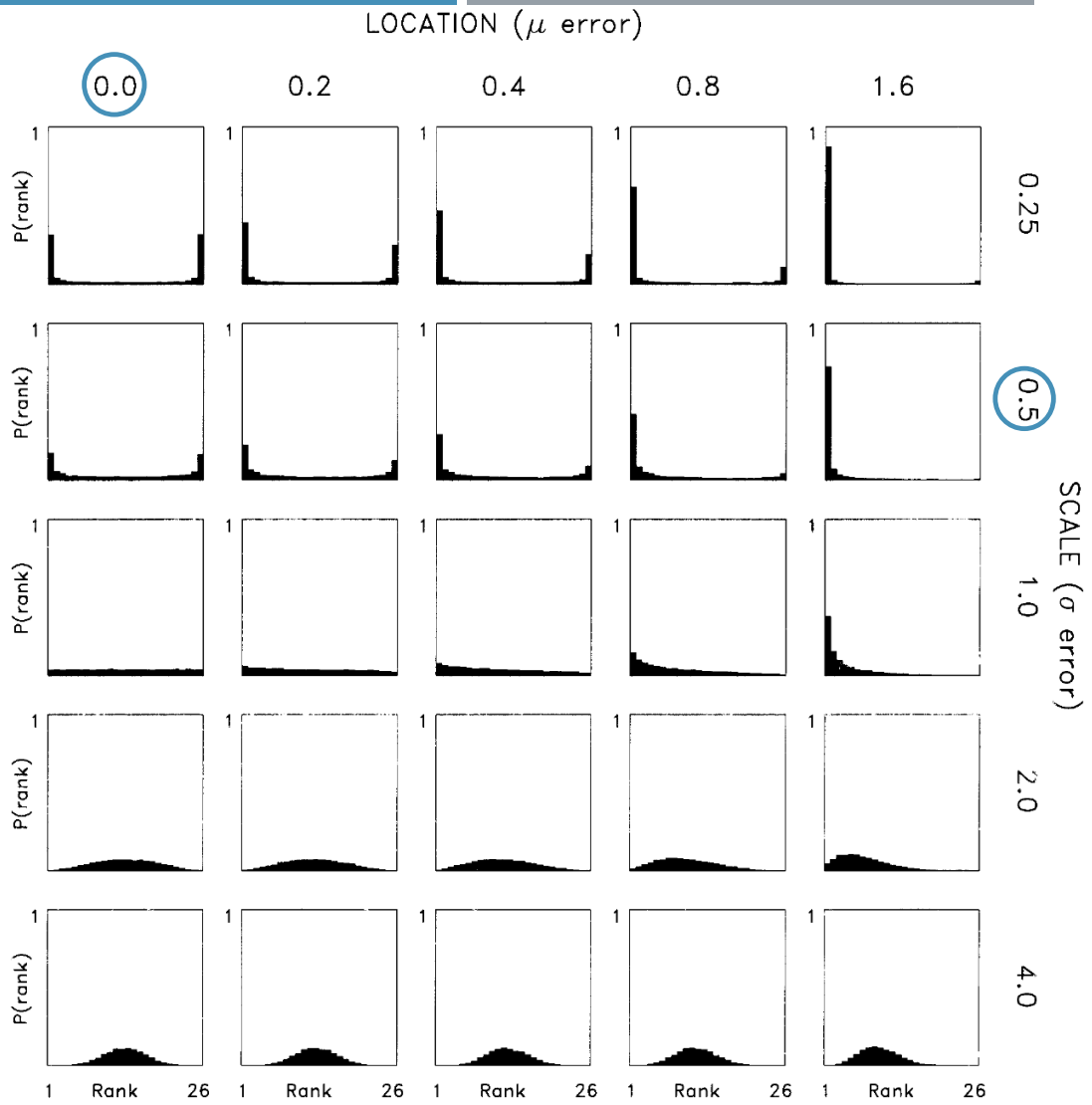


FIG. 1. Rank histograms where verification is sampled from a $N(0, 1)$ distribution and the ensemble ($n = 25$ members) is sampled from a $N(\mu, \sigma)$ distribution. The rank of the verification is tallied 10 000 times in each panel.

Hamill (2001)

CONCLUSIONS

- When providing only time lagged members, the single observation tests works well and we can see the flow-dependent increment at front or typhoon cases.
- The performance of hybrid DA system with time lagged members are better than the 3DVAR experiment.
- The spread of time lagged members are large at mid-high latitude area and small at low latitude and tropics area.
- The time lagged members during experiment period have no bias but indicates small spread.

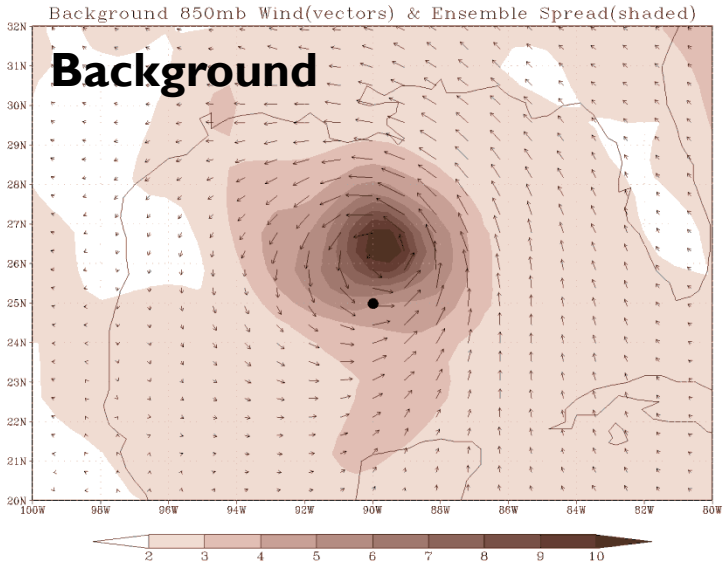
- NCEP operational GSI and EnKF systems are ported to CWB HPC system (Fujitsu FX10) and just works well last week. We will test the hybrid with EnKF system and evaluate the performance soon.

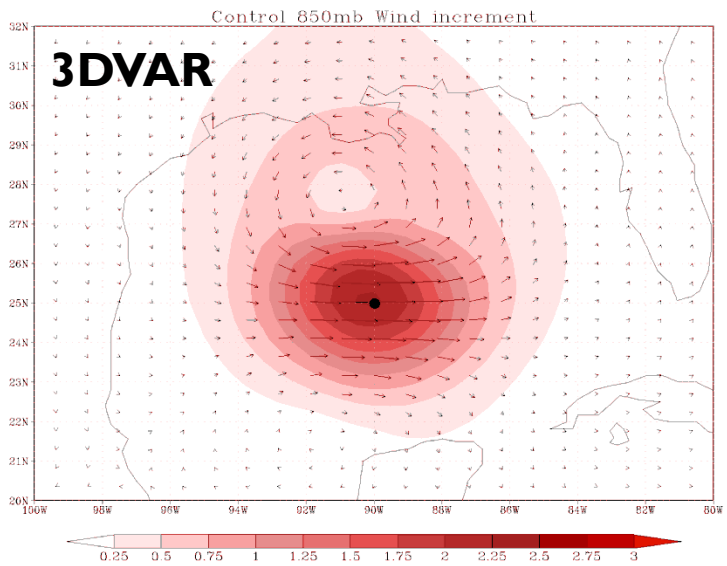
THE END

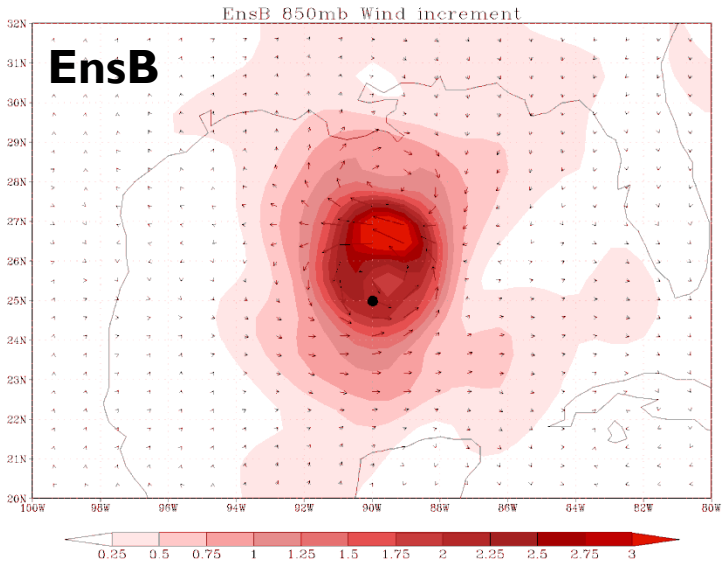
THANKS FOR YOUR ATTENTION.

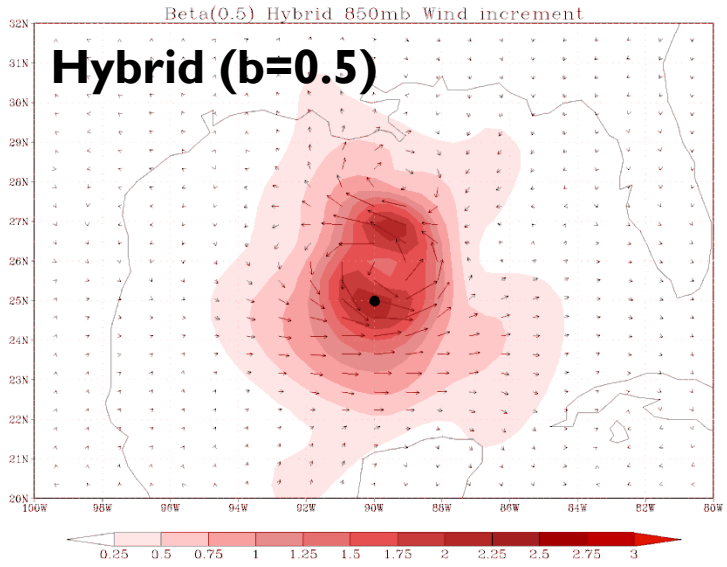
SINGLE OBSERVATION

from D. Kleist



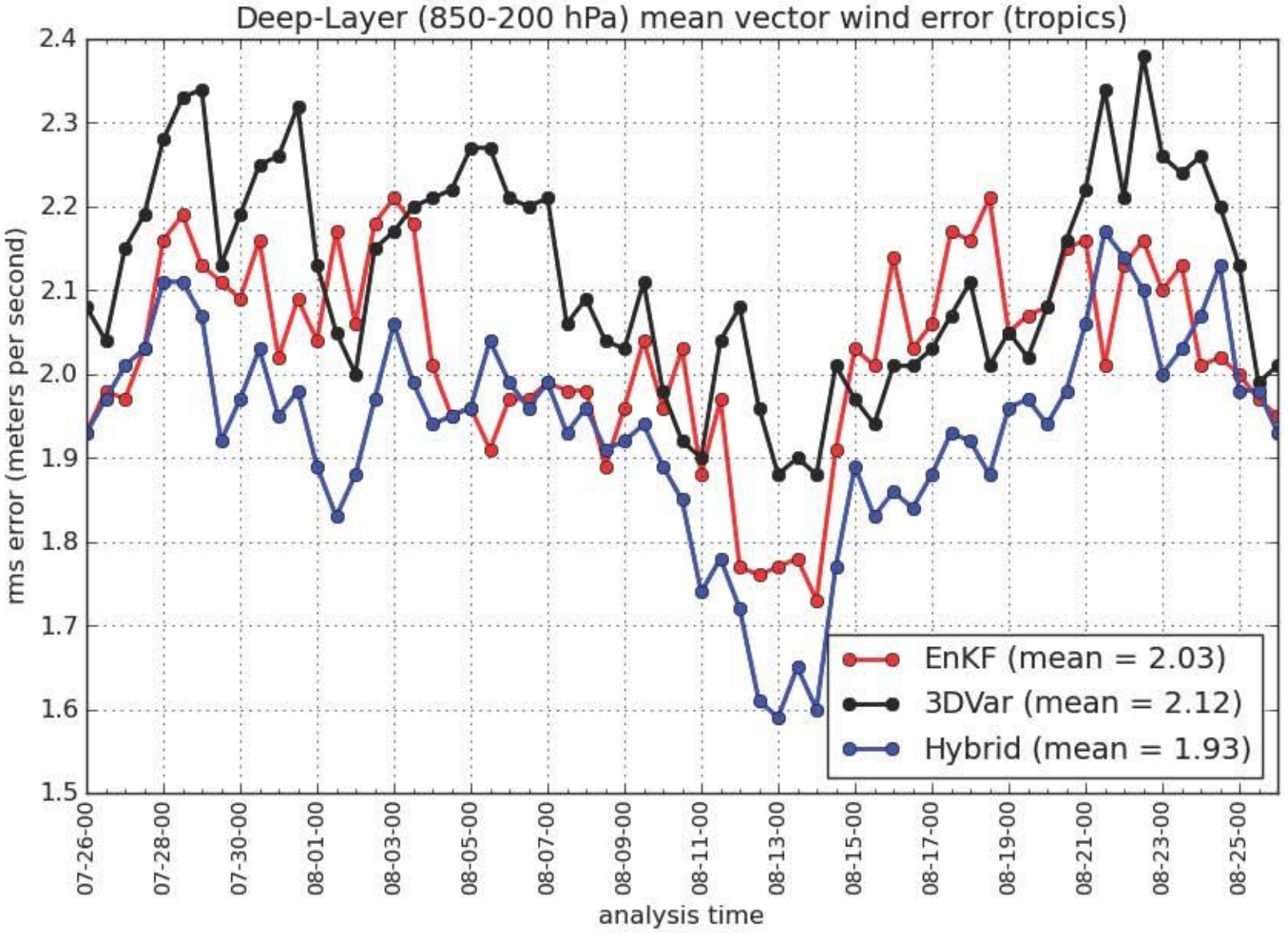






Single 850mb zonal wind observation (3 m/s O-F, 1m/s error) in Hurricane Ike circulation

Tropical Wind Errors (72-h) vs "consensus" analysis



DLM == mean of vector wind at 850,700,500,400,300,200 hPa.

"Consensus analysis' = 1/3(ECMWF + UKMET + NCEP)

Kleist et al., 2011