MTG-IRS: An Observing System Simulation Experiment (OSSE) on regional scales

Xiang-Yu Huang, Hongli Wang, Yongsheng Chen
National Center for Atmospheric Research, Boulder, Colorado, U.S.A.

Xin Zhang
University of Hawaii, Honolulu. Hawaii, U.S.A.

Stephen A. Tjemkes, Rolf Stuhlmann
EUMETSAT, Darmstadt, Germany

Contents

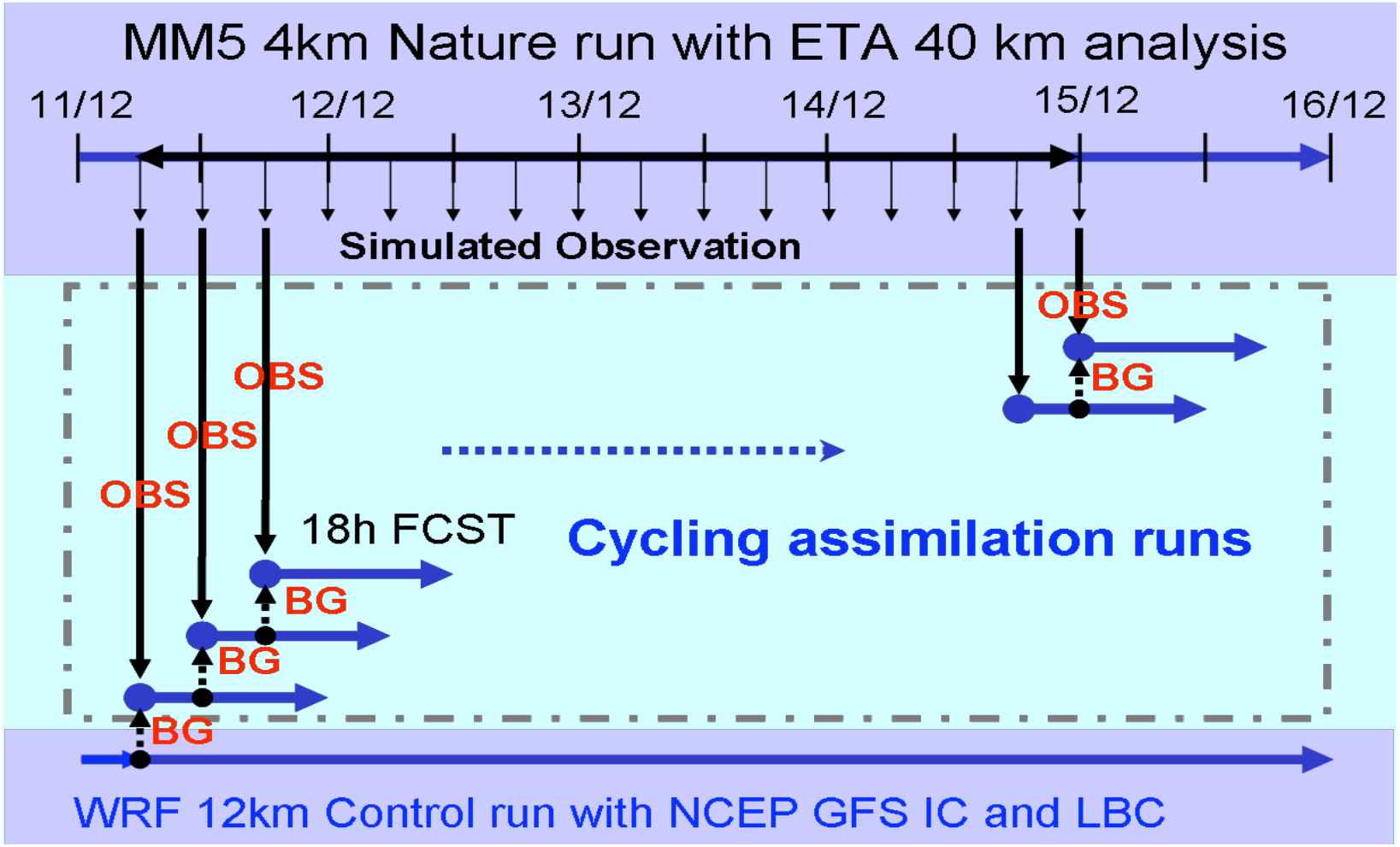
- **Background**
- **The nature run (MM5)**
- **Calibration experiments (WRF)**
- **MTG-IRS retrievals**
- **Data assimilation and forecast results (WRF)**
- **Summary**
- **Future work**
- **Identical twin experiments**

Background

- IRS sounding Mission on MTG will provide high-resolution data which includes temperature and water vapor information.
- Realistic mesoscale details in moisture are important for forecasting convective events (e.g., Koch et al. 1997; Parsons et al. 2000; Weckwerth 2000, 2004).
- **Objective**: To document the added value of water vapor observations derived from a hyperspectral infrared sounding instrument on a geostationary satellite for regional forecasting.

OSSE setup

2 models; Degraded resolution and LBCs

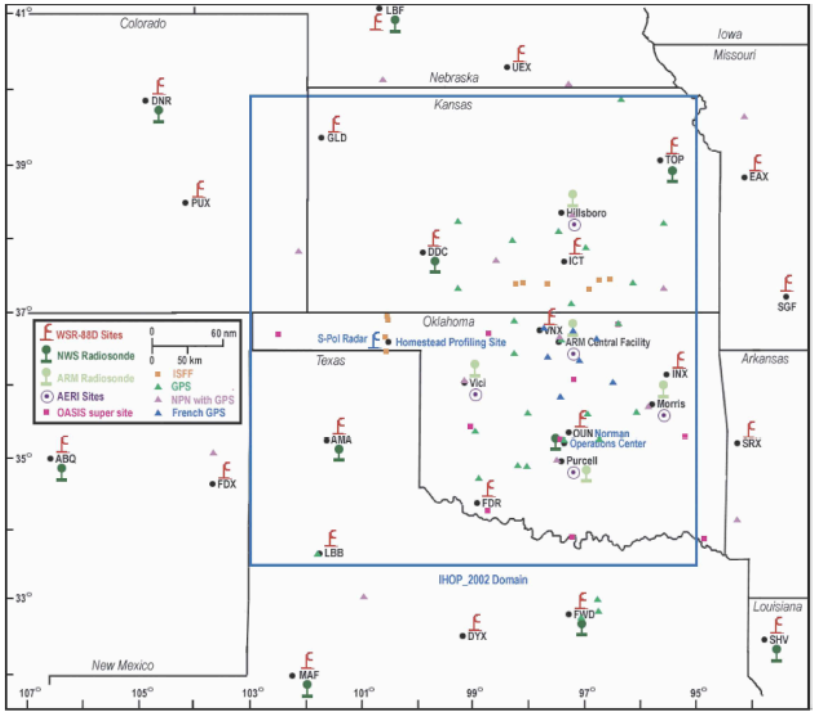


(Tom Schlatter:) In a true OSSE,

1. The nature run should be a long, uninterrupted forecast.
   - Here, NR is a 5-day “free” run.
2. The nature run should exhibit the same statistical behavior as the real atmosphere but be completely independent of it.
   - We need ideas on how to do this properly.
   - We just made a comparison between NR and real obs.
3. The assimilation period runs sufficiently long that the statistics comparing control and experimental forecasts are stable.
   - Data assimilation experiments are run over a 5-day period.
4. The lateral boundary conditions should vary with the experiment being performed in the inner domain.
   - We cannot run global experiment with new data. More coordinated effort is needed.
   - We tried to make the LBC in DA different to that in NR: Use ETA for the nature run and GFS for the assimilation run. We also add perturbations to lateral boundaries.
5. All major operational observing systems should be simulated.
   - We have ADP data, but could be better.
6. Errors are added to the hypothetical observations extracted from the nature run.
   - We add "realistic" errors to the truth.
7. True OSSEs are calibrated.
   - We made calibration runs.

- Modern variational assimilation systems deal with radiances directly because the error characteristics are easier to track.
  - It is still difficult to use radiance data over land.
  - Data thinning is used to account for the correlated observation errors.
  - This is also a pilot (small) project.

Nature Run: IHOP Case (11-16 June 2002)

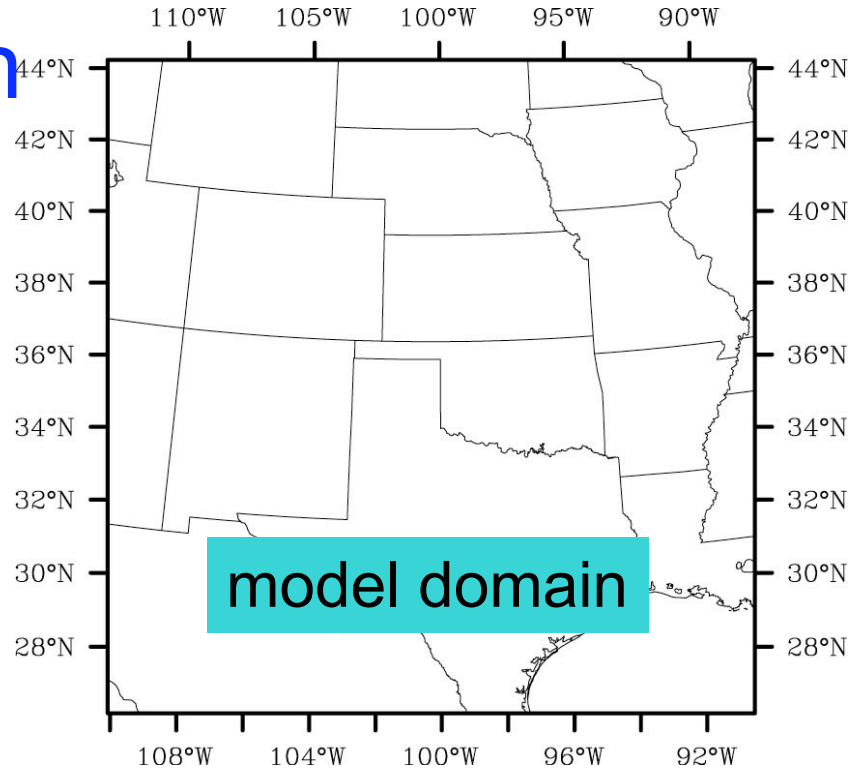


Map illustrating the operational instrumentation within the IHOP_2002 domain. (From Weckwerth et al. 2004.)

There are three convection cases in the selected period:

- 11 June: Dryline and Storm
- 12 June: Dryline and Storm
- 15 June: Severe MCS

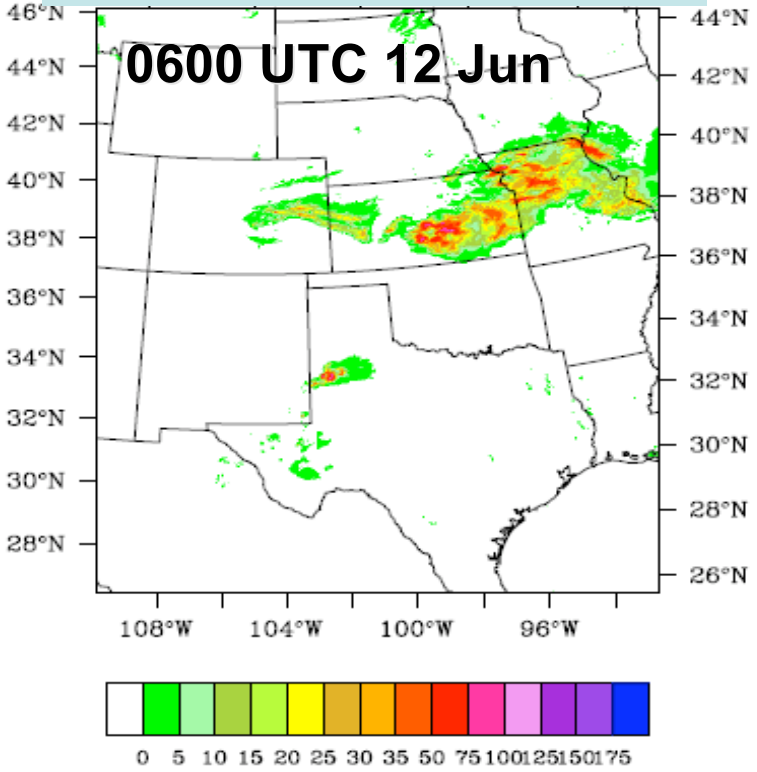
Nature Run Design



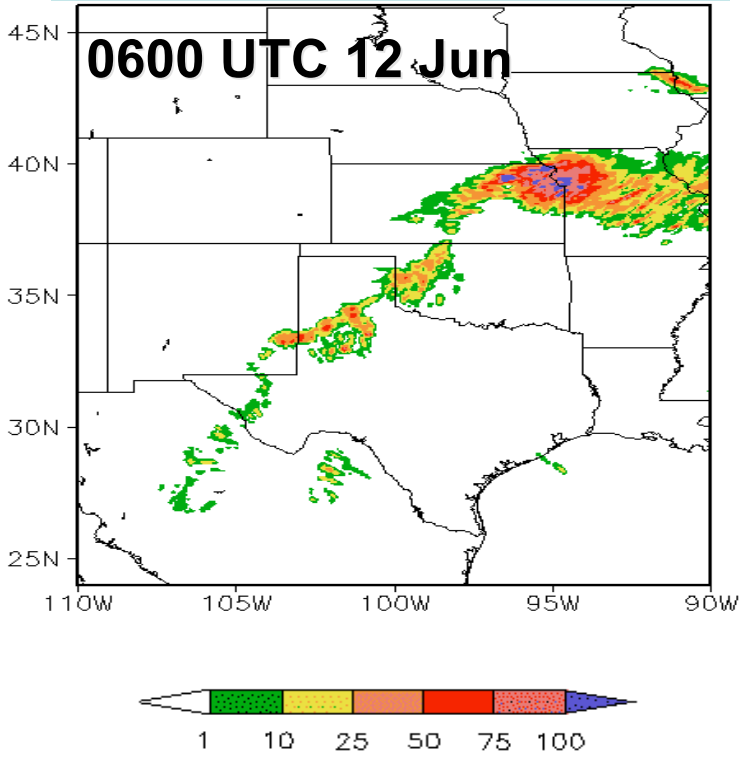
- Nature model: MM5
- Grid points: 505X505X35
- Horizontal resolution: 4Km
- Time step: 20s
- Physics parameterizations:
  - Reisner 2 microphysics
  - No cumulus parameterization
  - MRF boundary layer
- Initial and Lateral boundary condition:
  - 6-hourly ETA model 40-km analyses
- ~ 220 minutes with 256 CPUs

Case A: 11 June Case

observed 6-h rainfall

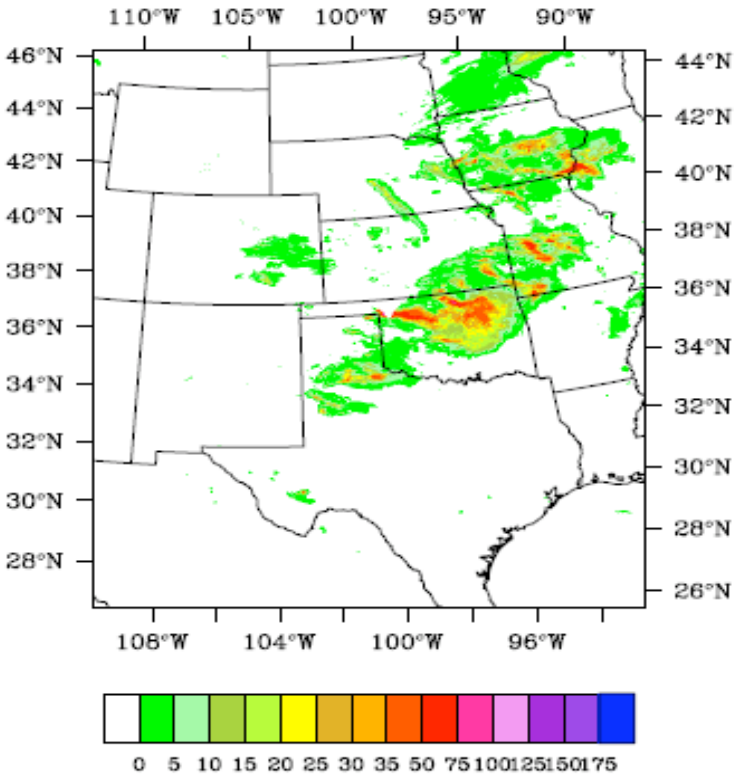


simulated 6-h rainfall



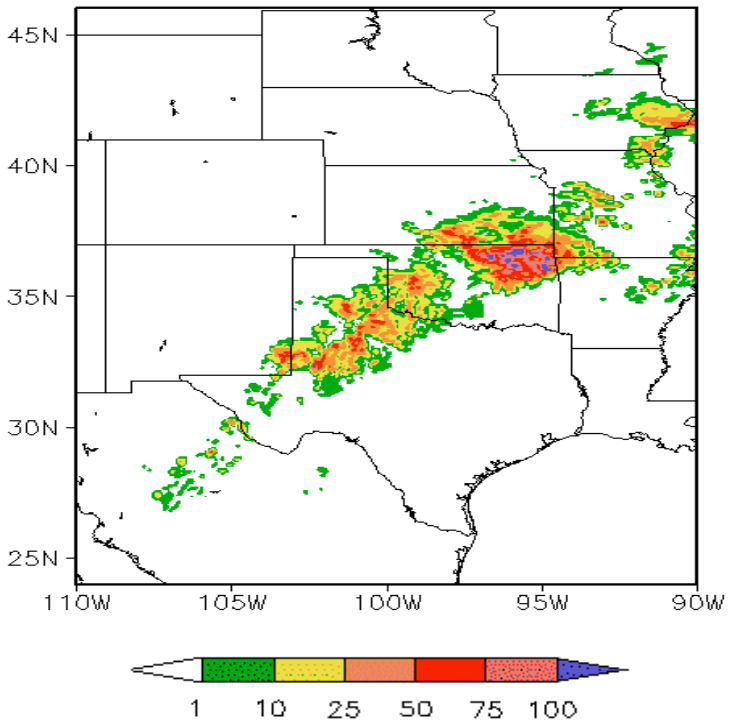
The observation is on Polar Stereographic Projection Grid.
The simulated rainfall is on Lambert Projection Grid.
The color scales are different.

Case B: 12 June Case



0600 UTC 13 Jun

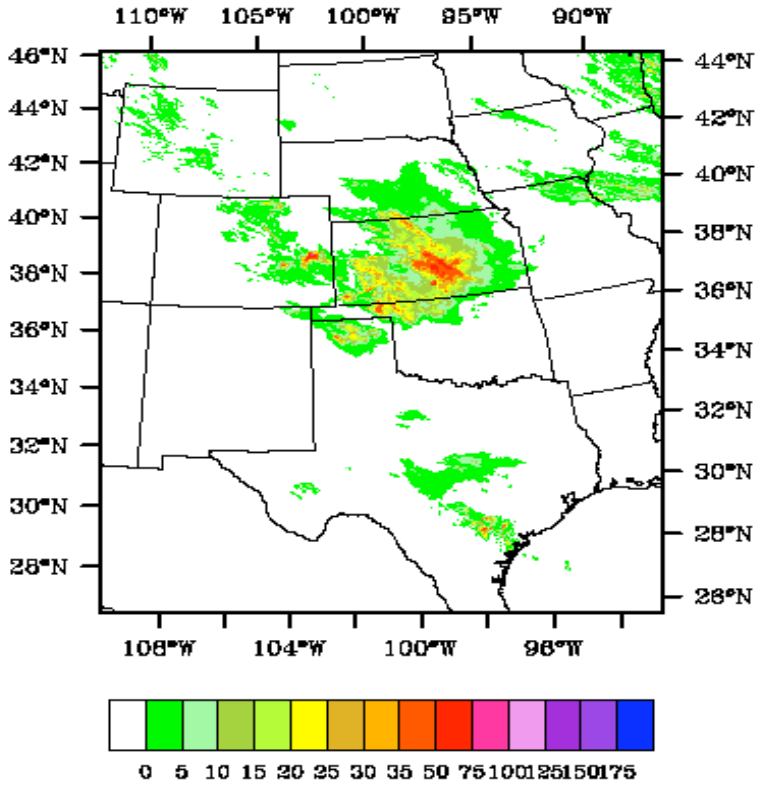
observed 6h-rainfall



0600 UTC 13 Jun

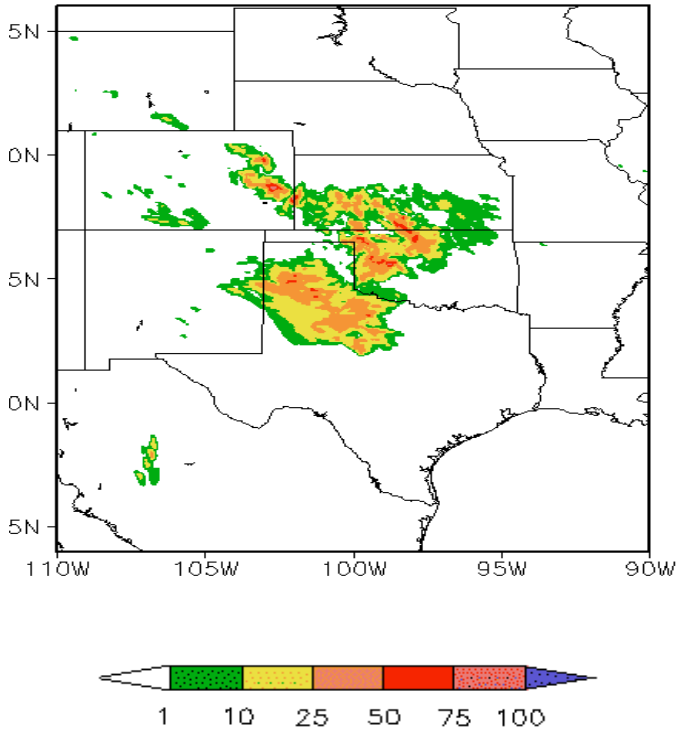
simulated 6h-rainfall

Case C: 15 June Case



0000 UTC 16 Jun

observed 6h-rainfall



0000 UTC 16 Jun

simulated 6h-rainfall

Calibration runs

- Generate pseudo (ADP) observations from the nature run.
- Exp 1. No-obs.
- Exp 2. Pseudo-obs. Data assimilation experiment using the pseudo observations.
- Exp 3. Real-obs. Data assimilation experiment using real (ADP) observations.

Simulated Dataset

WRF-Var is employed to produce simulated conventional observations

(NCEP ADP Upper Air sounding and Surface Observation)

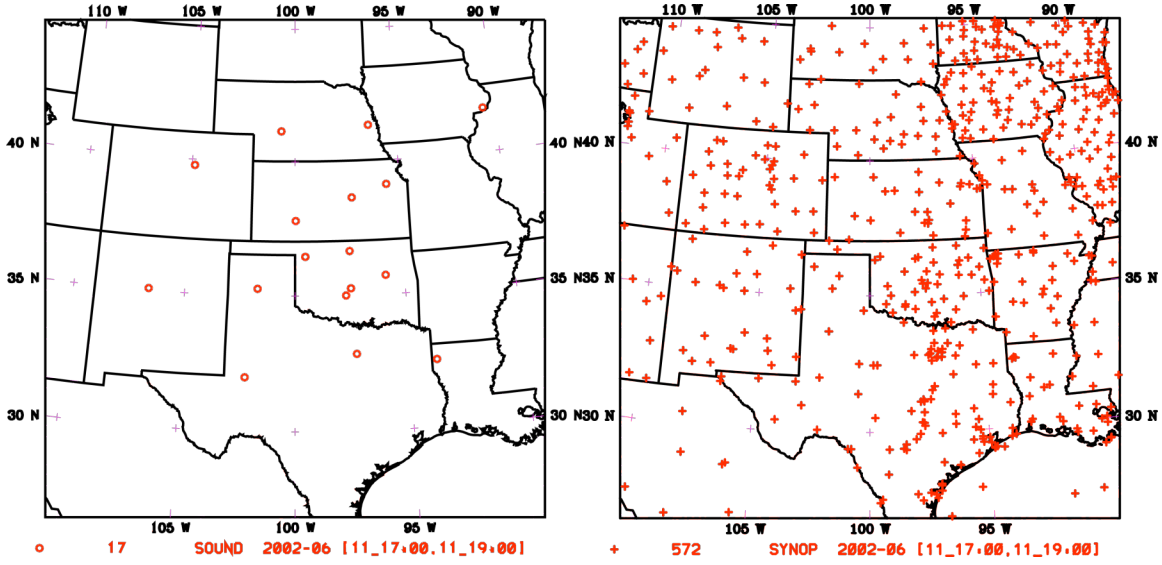
- Simulated conventional observations use the actual locations and times
- Add realistic observation errors

Simulated Dataset

- NCEP ADP Upper Air sounding
- NCEP ADP Surface Observation

Example of Simulated Data distribution

within the time window: 1700 UTC to 1900 UTC 12 June 2002

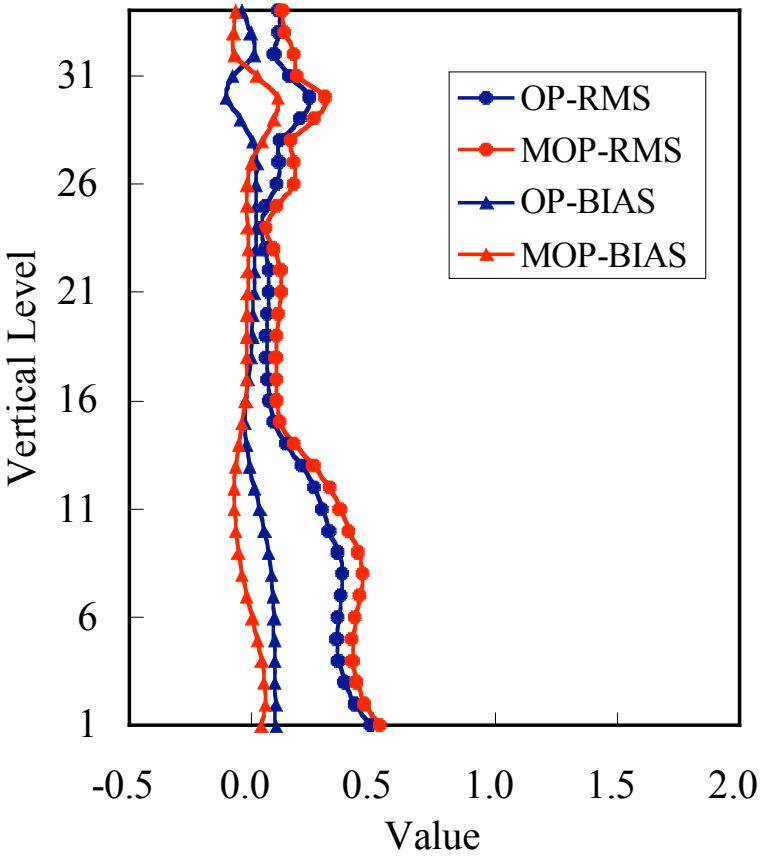


17 SOUND **572 Surface**

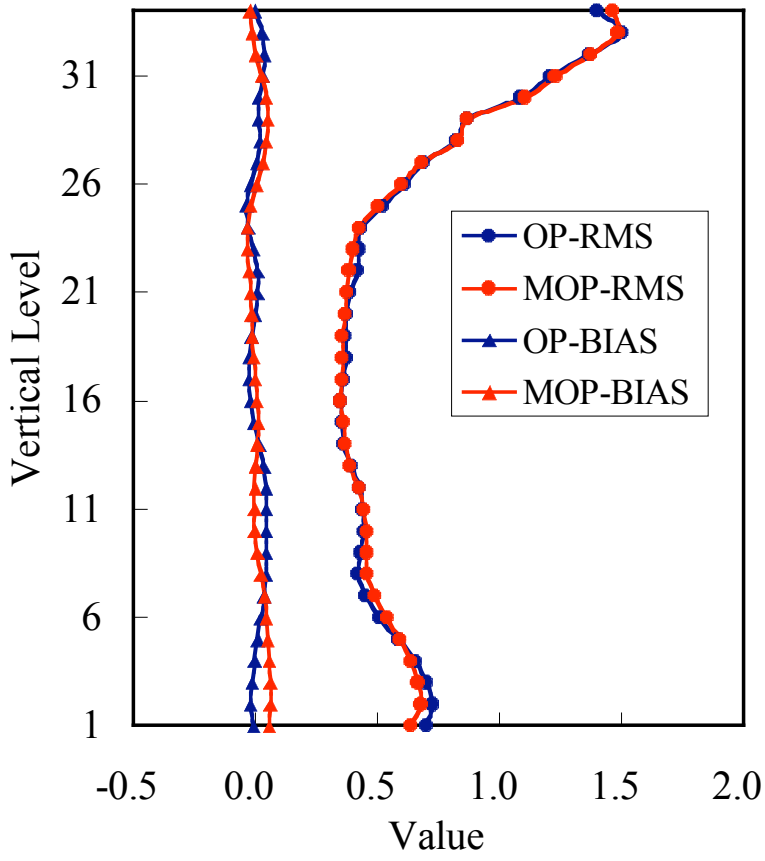
Observation errors in the simulated dataset

Statistics	Observation	u (m/s)	v (m/s)	T (K)	q (g/Kg)
Bias	Surface	-0.0018	-0.020388	-0.001976	-0.138
	Sounding	0.021482	0.035018	0.020335	-0.031
RMSE	Surface	1.10728	1.12405	1.00531	2.8376
	Sounding	2.36775	2.40612	0.83287	1.1964

Difference in T (K), 4 km results, averaged over 1800 UTC 11 to 1200 UTC 15 June 2002



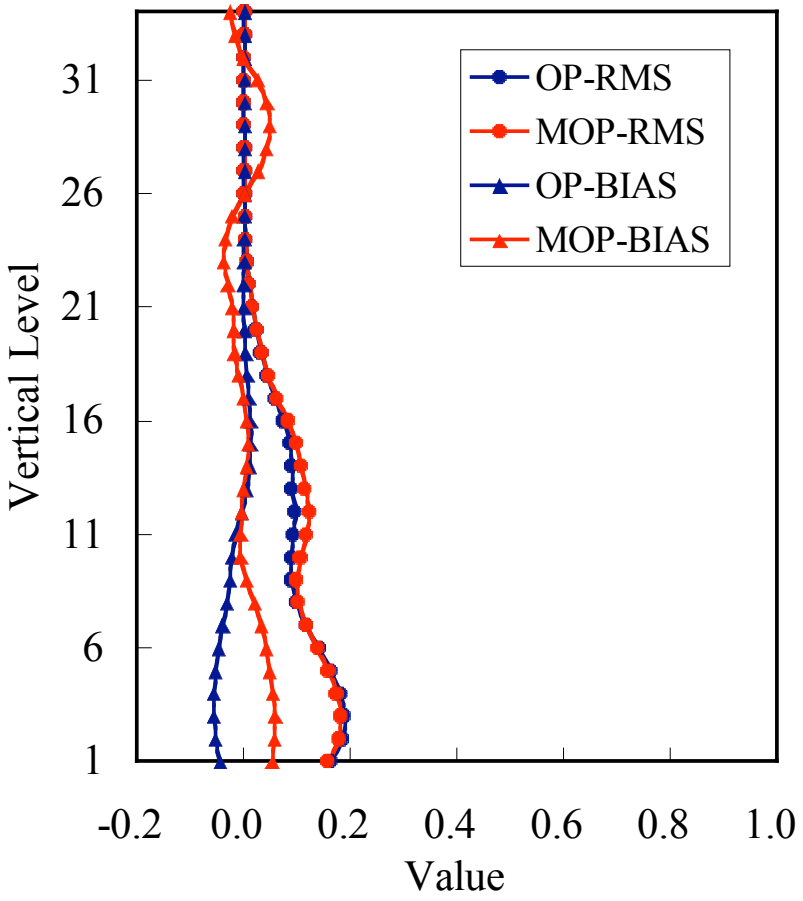
At analysis time



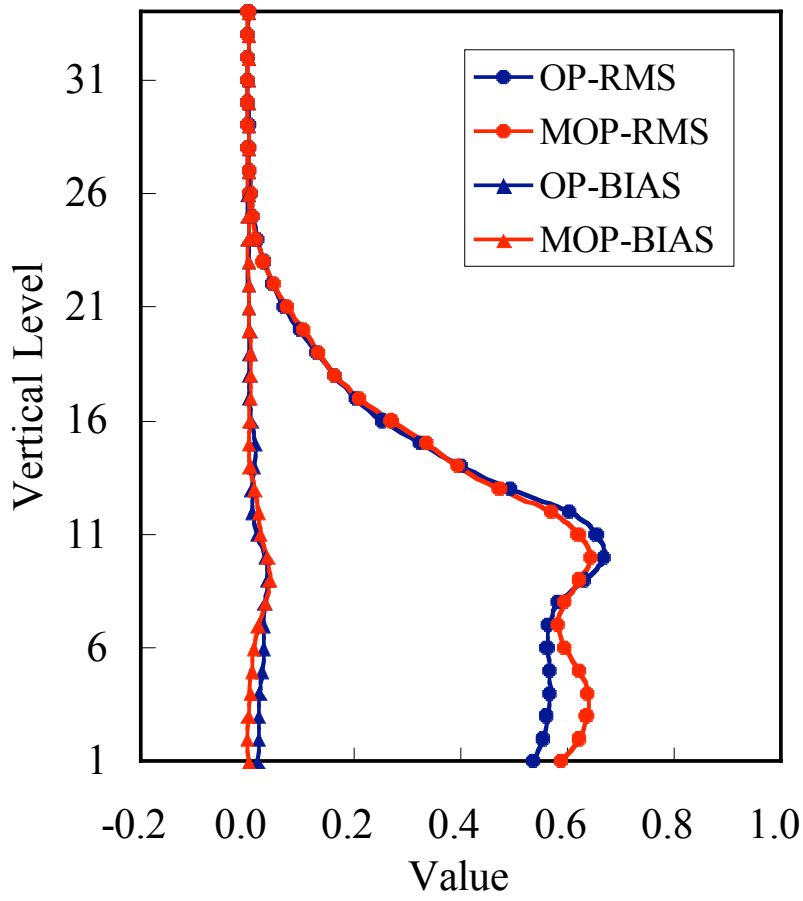
At 18 h FCST

MOP: Modeled Observation Profiles; OP: (real) Observation Profiles

Difference in q (g/kg), 4 km results, averaged over 1800 UTC 11 to 1200 UTC 15 June 2002

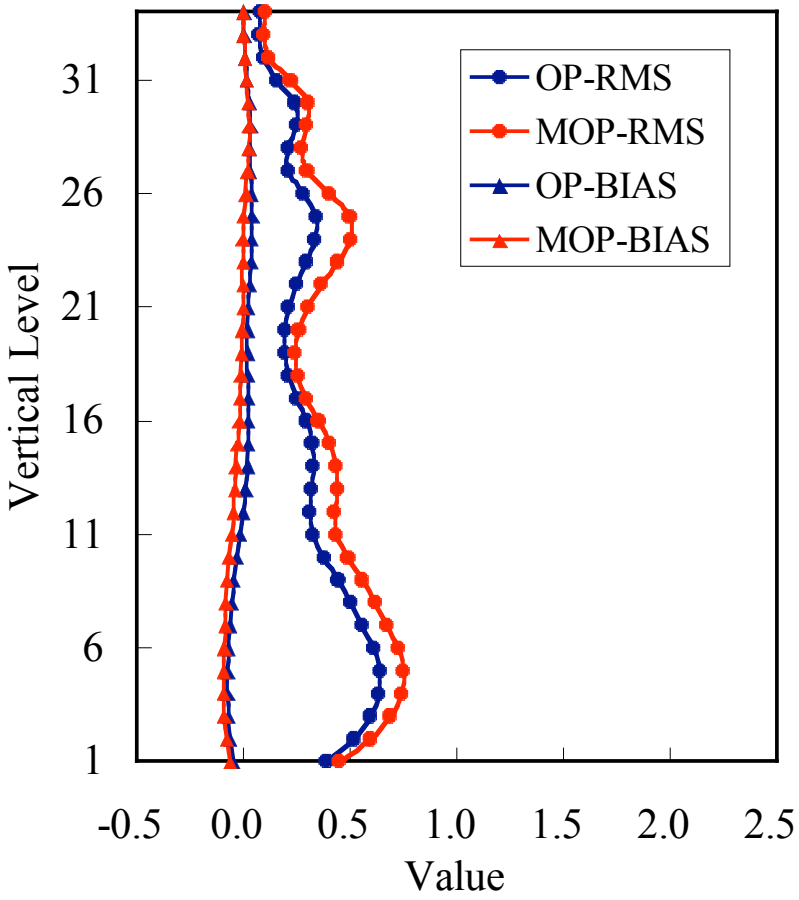


At analysis time

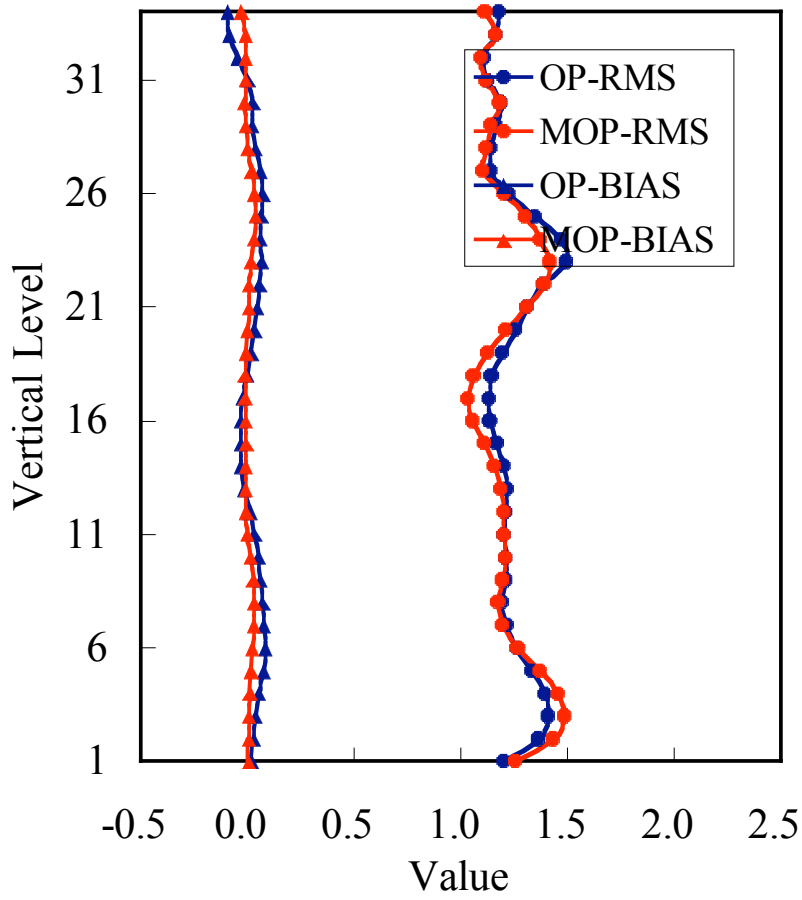


At 18 h FCST

Difference in u (m/s), 4 km results, averaged over 1800 UTC 11 to 1200 UTC 15 June 2002



At analysis time



At 18 h FCST

MTG-IRS Retrieval (I)
Forward calculations

- Profile information for the forward calculations are combination of climatology (above 50 hPa) and MM5 results (below 50 hPa), Ozone information is extracted from climatology. For each hour for five days 505 x 505 profiles (= one “data cube”).

- RTM adopted is same code as used for HES/GIFTS trade-off studies by SSEC, which is a statistical model. Only clear sky calculations, accuracy is not known.

- CPU: To generate R(toa) for one “data cube” takes about 20 hours CPU.

MTG-IRS Retrieval (II)
Inverse Calculations

- Results are based on EOF retrievals

- Four datasets:

  S^t : Training dataset: $T^t(p)$, $q^t(p)$ and $R^t(toa)$,
  S^o: Synthetic observational dataset: $R^o(toa)$
  S^r : Retrieval dataset: $T^r(p)$, $q^r(p)$
  S^n : Nature (here taken from MM5): $T^n(p)$, $q^n(p)$

  Objective of retrieval is to generate a S^r from S^o, which is equal to S^n

- Flowchart of EOF retrieval:

  Step 1: Truncate R^t (toa) through an EOF decomposition
  Step 2: Correlate the truncated R^t (toa) with $T^t(p)$, $q^t(p)$ to generate regression coefficients
  Step 3: Project $R^o(toa)$ onto EOF space of R^t (toa)
  Step 4: Generate $T^r(p)$, $q^r(p)$ using regression coefficients from 3) and EOF from 2)

Two EOF Training methods

Two different training methods applied:

Global Training: generated a “global dataset” by random selection of profiles from a number data cubes covering dynamical range of the diurnal cycle. About 100000 profiles, a single training dataset

As this global dataset had different properties than an individual data-cube; assimilation generated not satisfactory results (mainly because of bias)

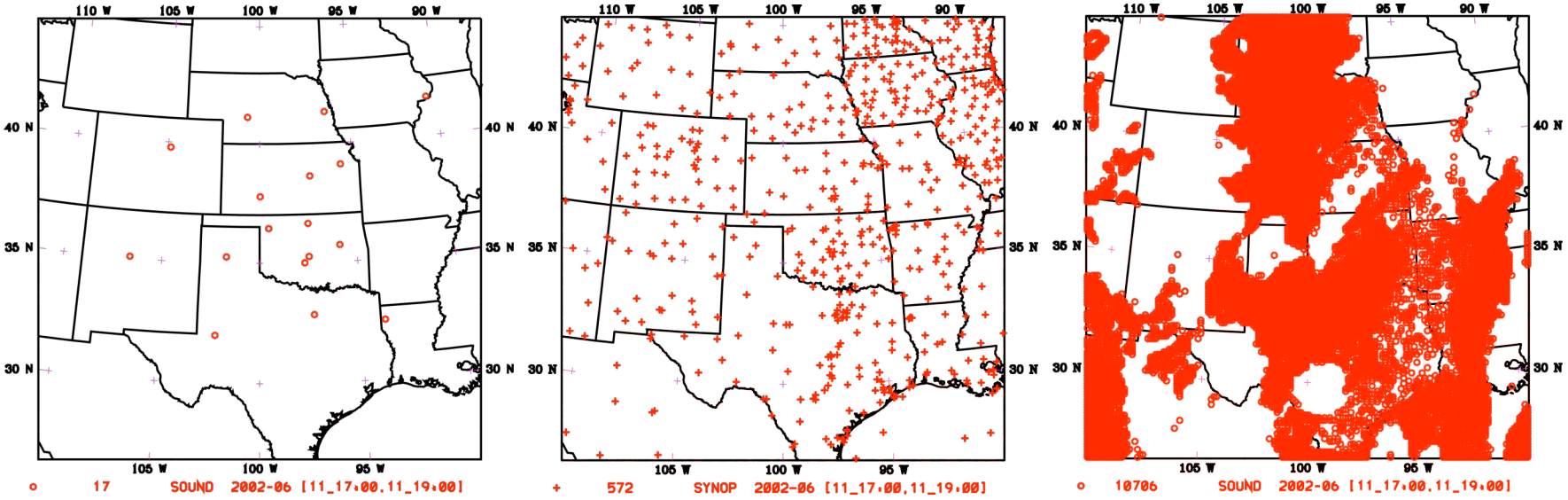
“Bias free Training”: For each datacube a separate training dataset consisting of 10% of the data in the particular datacube.

Simulated Dataset

- NCEP ADP Upper Air sounding
- NCEP ADP Surface Observation
- MTG-IRS retrieved profiles

Example of Simulated Data distribution

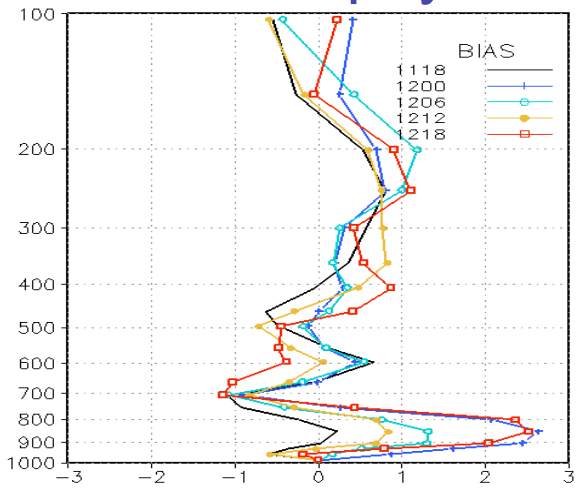
within the time window: 1700 UTC to 1900 UTC 12 June 2002

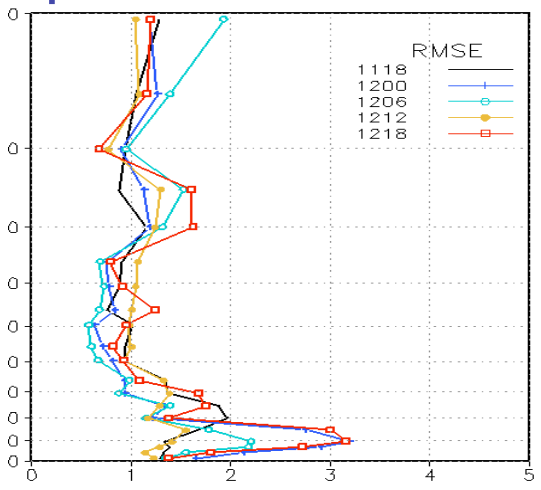


17 SOUND **572 Surface** **10706 MTG-IRS RP**

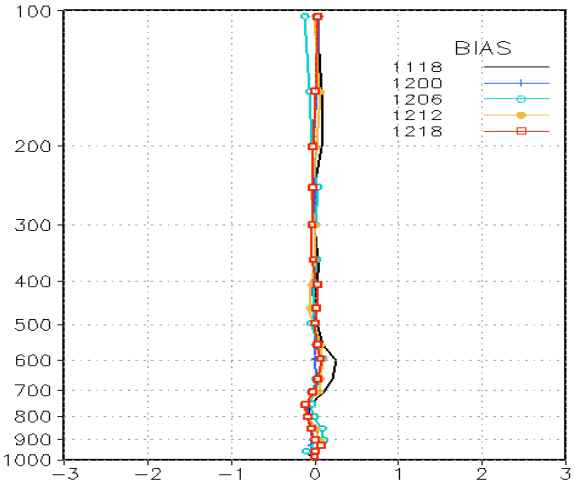
Temperature error statistics for RP

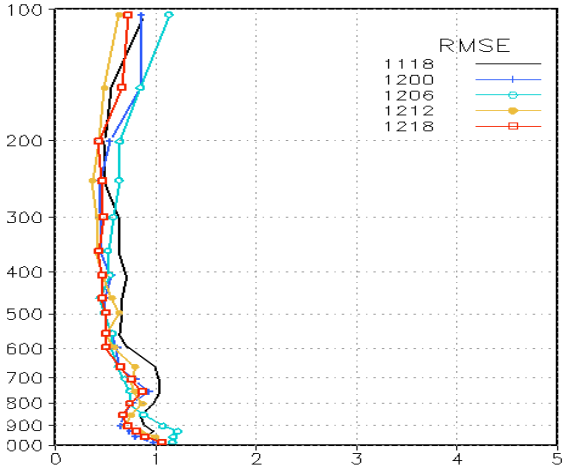
Old physical retrieval profiles





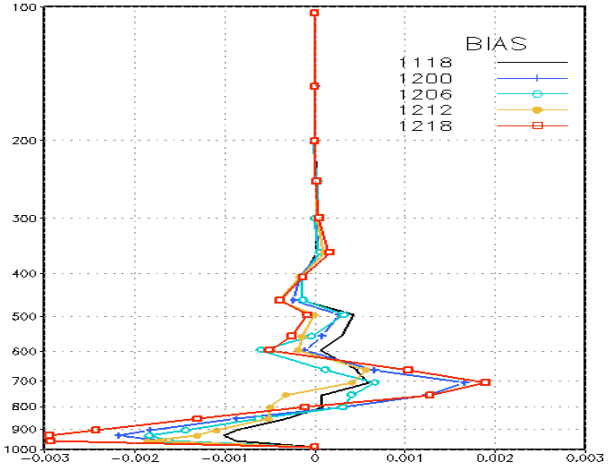
New EOF retrieval profiles

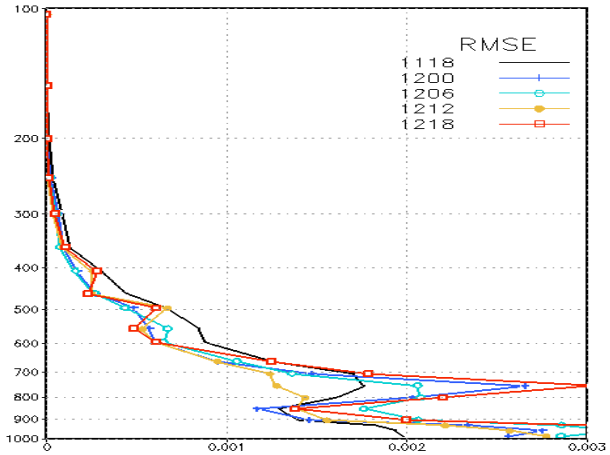




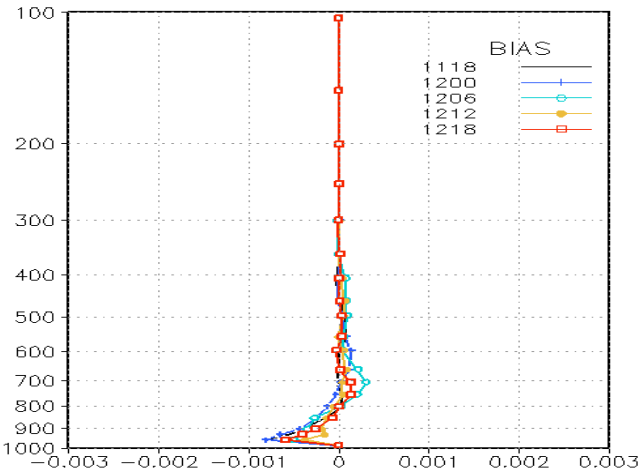
Humidity error statistics

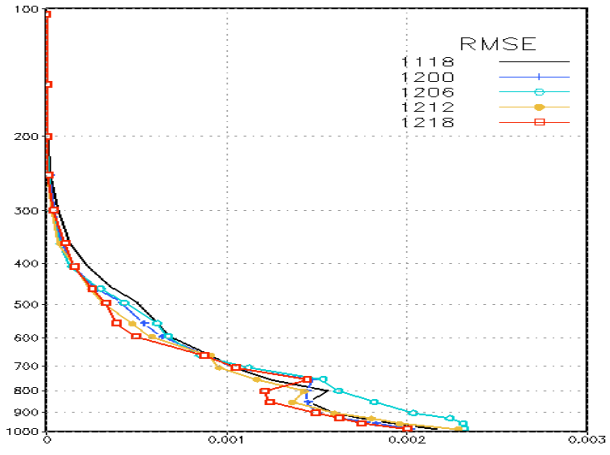
Old physical retrieval profiles



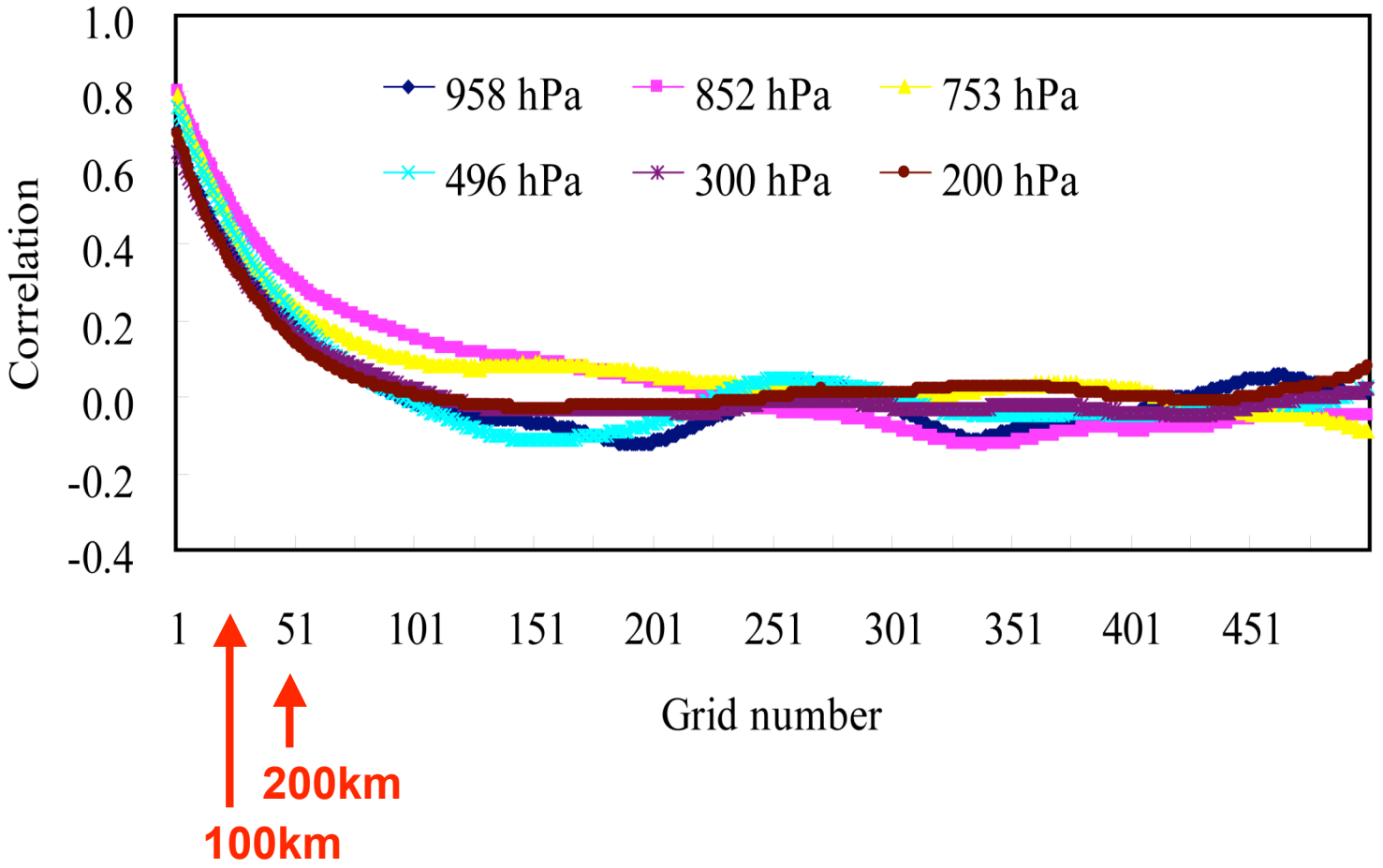


New EOF retrieval profiles

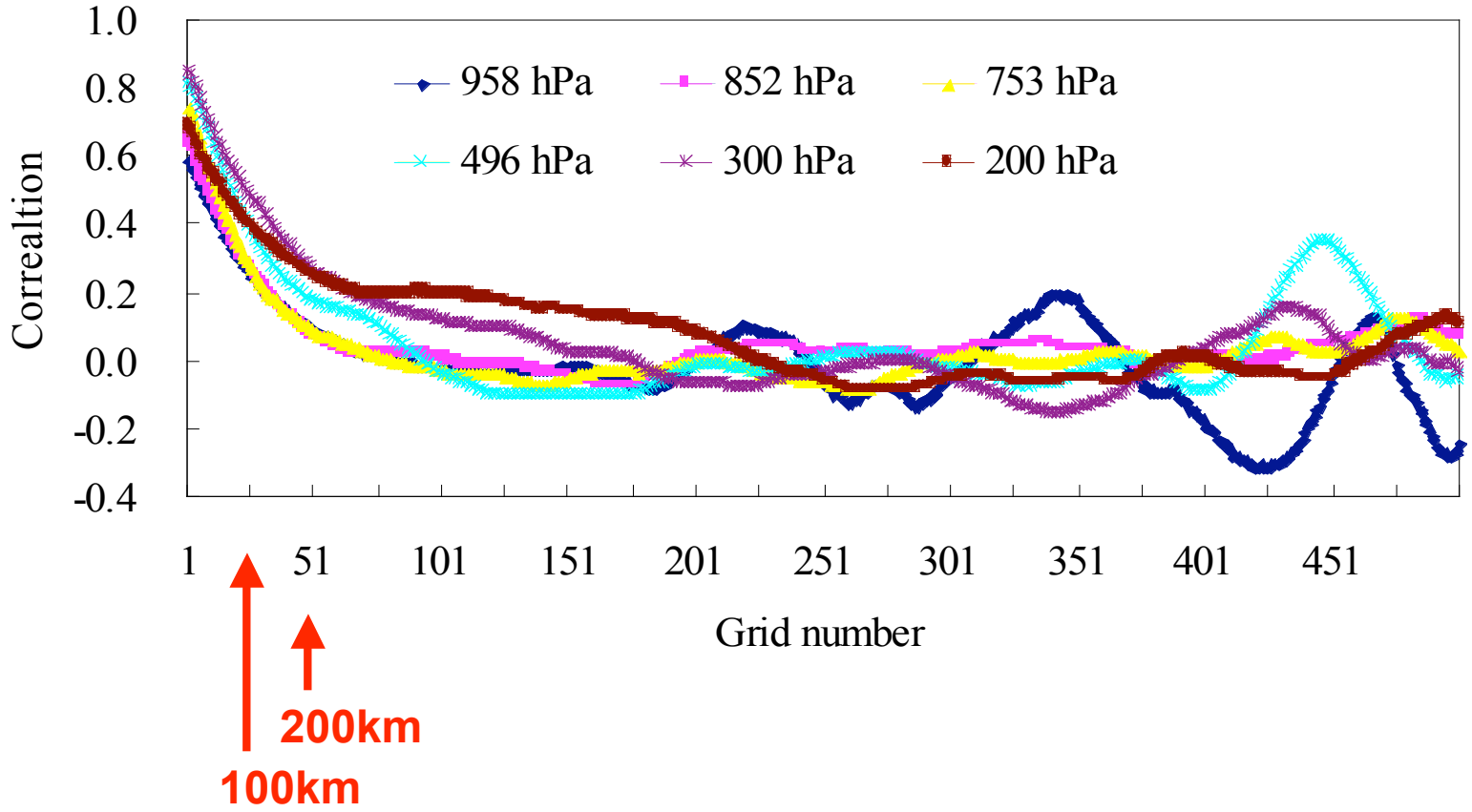




Temperature error correlation



Humidity error correlation



Vertical temperature error correlation at 18Z 11 June

103.02																			1
151.27																		1	0.3
200.99																	1	0.46	0.29
247.41																1	0.53	0.37	0.3
300.00															1	0.29	0.19	0.15	0.13
358.97														1	0.26	0.17	0.14	0.12	0.11
407.47													1	0.27	0.16	0.12	0.1	0.09	0.09
459.71												1	0.32	0.21	0.17	0.15	0.13	0.13	0.12
496.63											1	0.34	0.23	0.19	0.17	0.16	0.15	0.14	0.14
555.17										1	0.27	0.16	0.13	0.11	0.1	0.09	0.08	0.08	0.08
596.31									1	0.26	0.16	0.13	0.11	0.1	0.09	0.08	0.08	0.08	0.08
661.19								1	0.32	0.2	0.16	0.14	0.13	0.12	0.11	0.11	0.11	0.1	0.1
706.57							1	0.29	0.17	0.12	0.1	0.09	0.08	0.07	0.07	0.06	0.06	0.06	0.05
753.63						1	0.38	0.25	0.21	0.18	0.17	0.16	0.15	0.15	0.14	0.14	0.14	0.14	0.13
802.37						1	0.44	0.32	0.28	0.25	0.24	0.23	0.22	0.22	0.22	0.21	0.21	0.21	0.2
852.79					1	0.33	0.22	0.18	0.17	0.15	0.15	0.14	0.14	0.13	0.13	0.13	0.13	0.12	0.12	0.12
904.87				1	0.32	0.18	0.12	0.1	0.08	0.07	0.07	0.06	0.06	0.05	0.05	0.05	0.04	0.04	0.04	0.04
931.52			1	0.36	0.19	0.13	0.1	0.08	0.07	0.06	0.05	0.05	0.04	0.04	0.04	0.04	0.03	0.03	0.03	0.03
958.59		1	0.26	0.07	-0	-0.04	-0.06	-0.07	-0.08	-0.09	-0.1	-0.1	-0.1	-0.11	-0.11	-0.11	-0.11	-0.12	-0.12	-0.12
986.07	1	0.36	0.07	-0.03	-0.09	-0.12	-0.14	-0.15	-0.16	-0.17	-0.18	-0.18	-0.19	-0.19	-0.19	-0.2	-0.2	-0.2	-0.2	-0.21

Vertical humidity error correlation at 18Z 11 June

103.02																			1
151.27																		1	0.58
200.99																	1	0.51	0.28
247.41																1	0.42	0.21	0.13
300															1	0.29	0.06	-0.02	-0.06
358.97														1	0.41	0.19	0.11	0.07	0.04
407.47													1	0.56	0.37	0.29	0.25	0.22	0.21
459.71												1	0.67	0.53	0.47	0.44	0.42	0.41	0.4
496.63											1	0.62	0.47	0.41	0.37	0.35	0.34	0.33	0.32
555.17										1	0.49	0.3	0.22	0.18	0.16	0.14	0.13	0.12	0.12
596.31									1	0.45	0.23	0.15	0.1	0.08	0.06	0.05	0.04	0.03	0.03
661.19								1	0.52	0.29	0.2	0.15	0.12	0.1	0.08	0.07	0.06	0.06	0.05
706.57							1	0.54	0.32	0.24	0.19	0.17	0.15	0.13	0.12	0.11	0.11	0.1	0.1
753.63						1	0.56	0.38	0.31	0.27	0.25	0.23	0.22	0.21	0.21	0.2	0.2	0.19	0.19
802.37					1	0.58	0.44	0.38	0.35	0.33	0.32	0.31	0.3	0.3	0.29	0.29	0.29	0.28	0.28
852.79				1	0.55	0.43	0.38	0.36	0.35	0.34	0.33	0.32	0.32	0.31	0.31	0.31	0.31	0.31	0.3
904.87			1	0.66	0.53	0.47	0.45	0.43	0.41	0.41	0.4	0.39	0.39	0.39	0.38	0.38	0.38	0.38	0.37
931.52		1	0.71	0.56	0.49	0.46	0.43	0.42	0.41	0.4	0.39	0.38	0.38	0.38	0.37	0.37	0.37	0.36	0.36
958.59	1	0.86	0.74	0.67	0.63	0.6	0.58	0.56	0.55	0.54	0.53	0.53	0.52	0.51	0.51	0.51	0.5	0.5	0.5
986.07	0.92	0.83	0.78	0.75	0.72	0.7	0.69	0.68	0.67	0.66	0.66	0.65	0.65	0.64	0.64	0.64	0.63	0.63	0.63

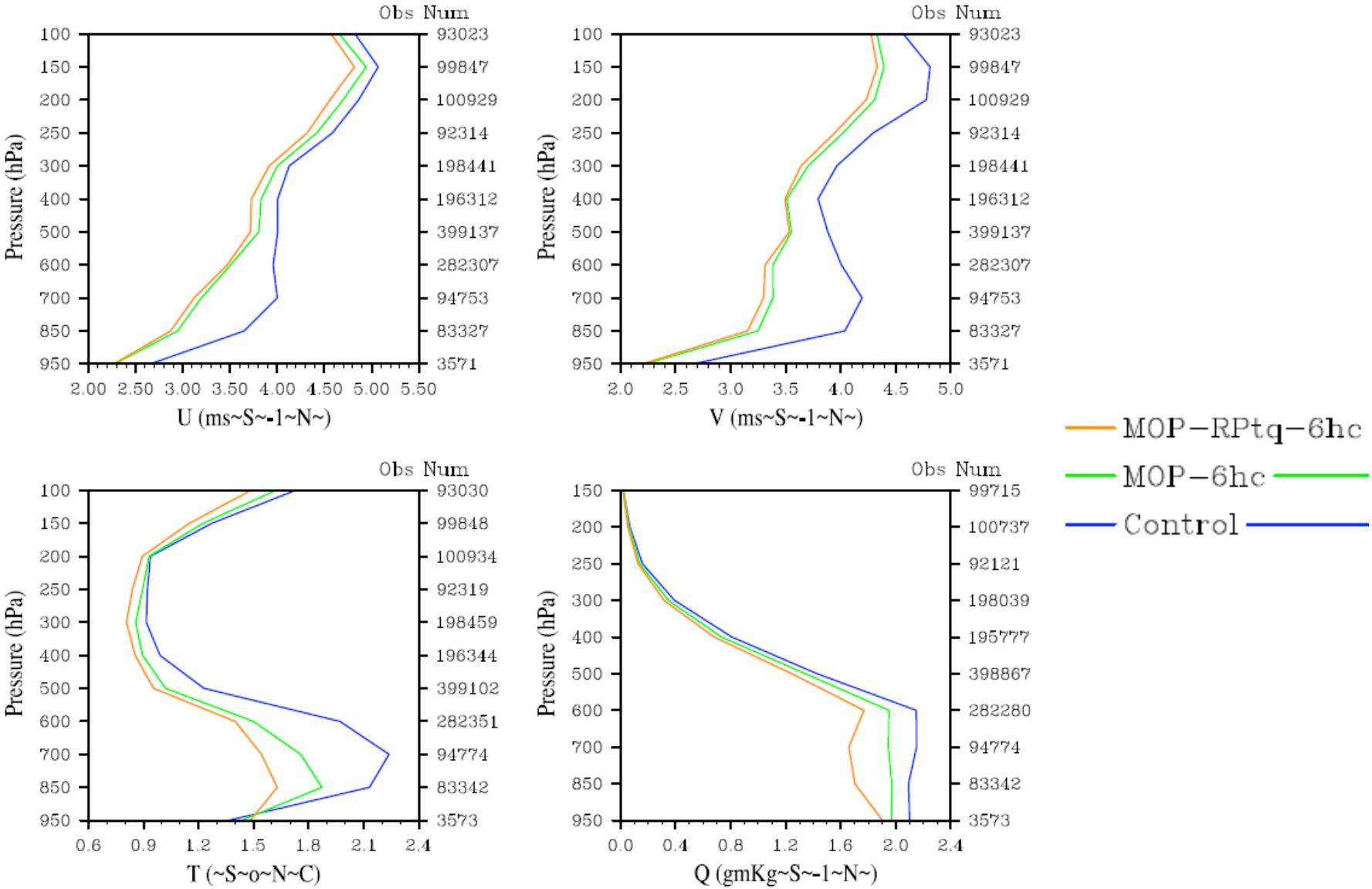
Experiments design

- Forecast model: WRF
- Data assimilation system: WRF 3D-Var
- Grid points: 169X169X35
- Horizontal resolution: 12Km
- Time step: 60s
- Physics parameterizations:
  - Lin microphysics
  - Grell cumulus parameterization
  - MRF boundary layer
- Cases: 2002-06-11 12Z to 2002-06-16 12Z
- Data:
  - MOP
  - EOF retrieved profiles (18 levels with 100 km resolution)
- Verification against truth

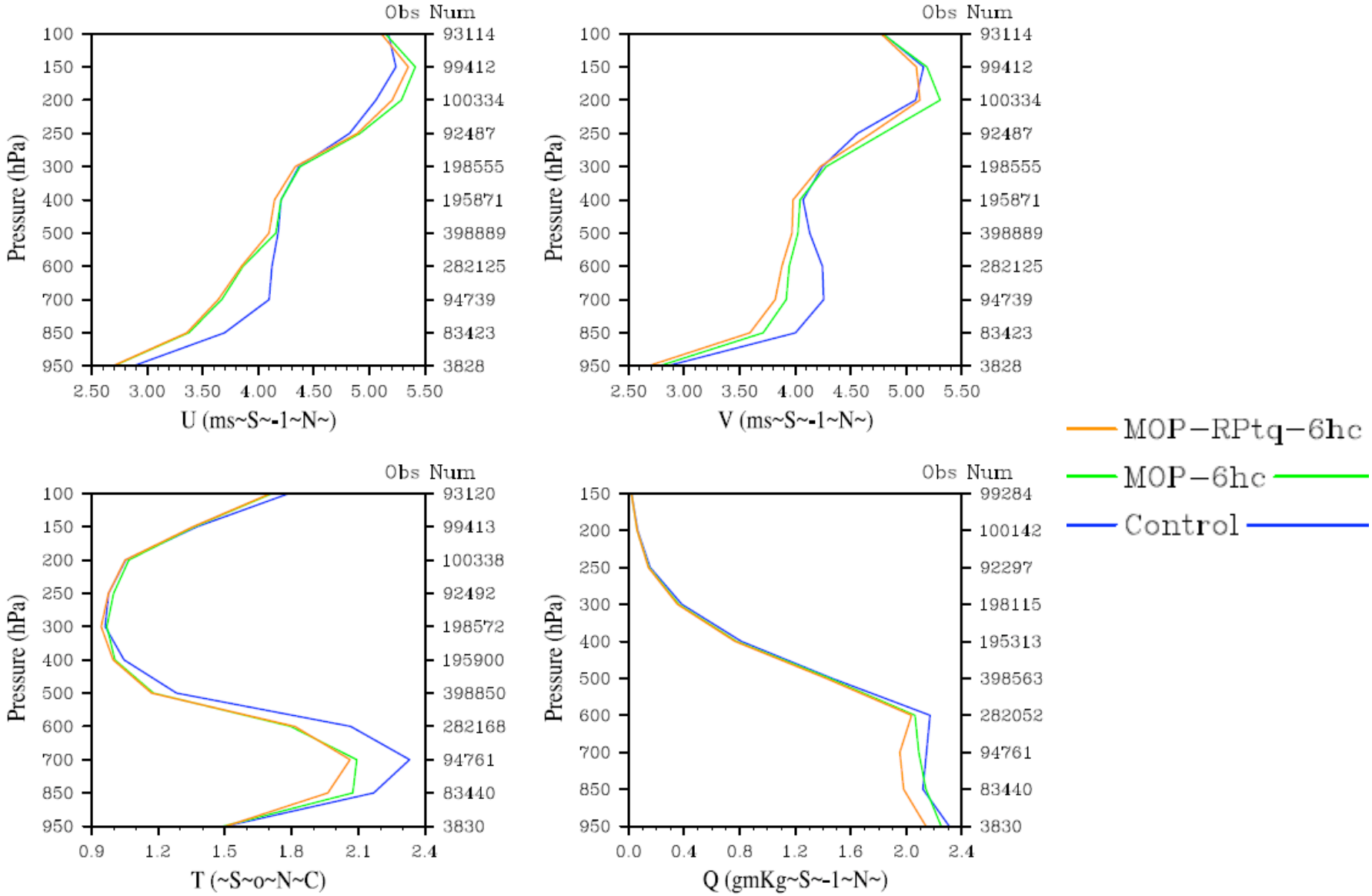
Lists of Experiments

Experiment name	Cycling period	Initial condition and assimilated data
Control	No	GFS analysis + perturbed lateral boundary conditions
MOP	6 h	Background (BG)+ Modeled Observation Profiles
MOP-RPtq-6hc	6 h	Background (BG)+ MOP +Retrieved Profiles(T,q)

Averaged RMS error profiles at analysis time

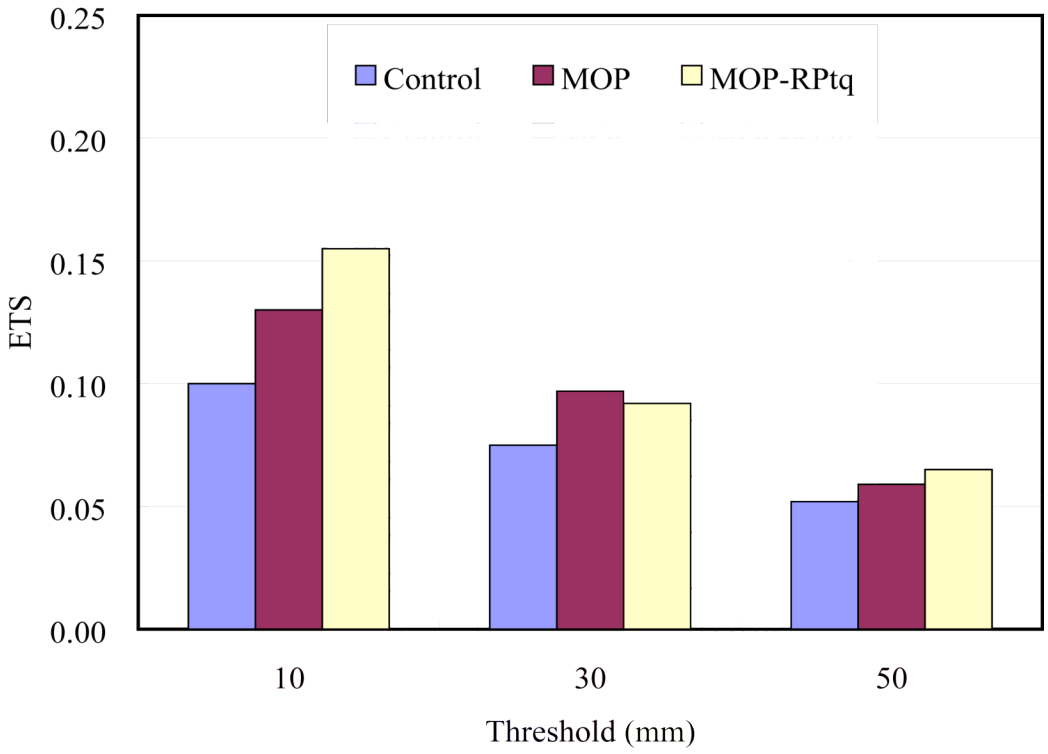


Averaged RMS error profiles at 12h FCST

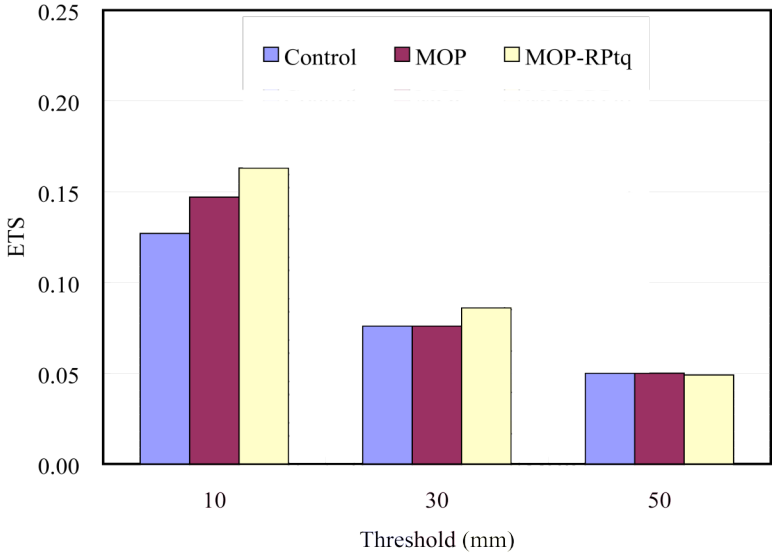


Averaged ETS

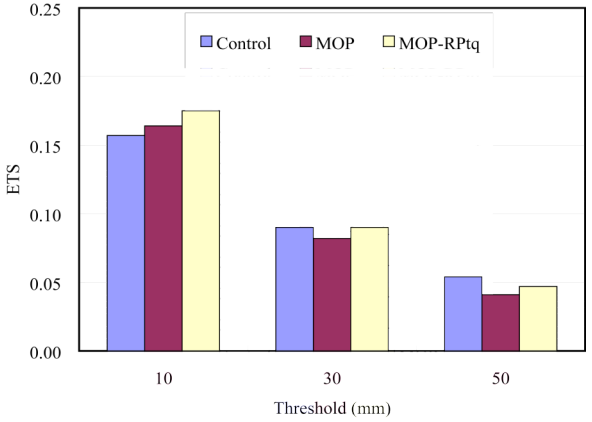
12h FCST



18h FCST

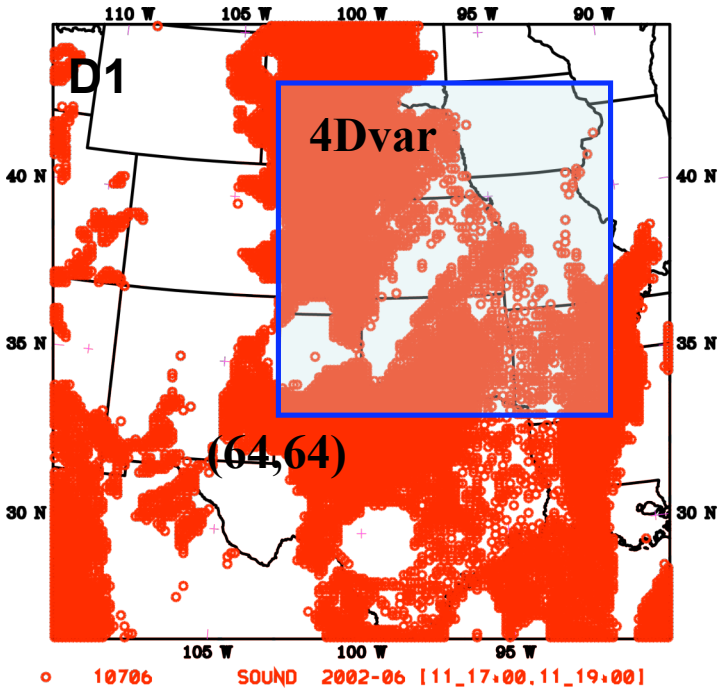


24h FCST

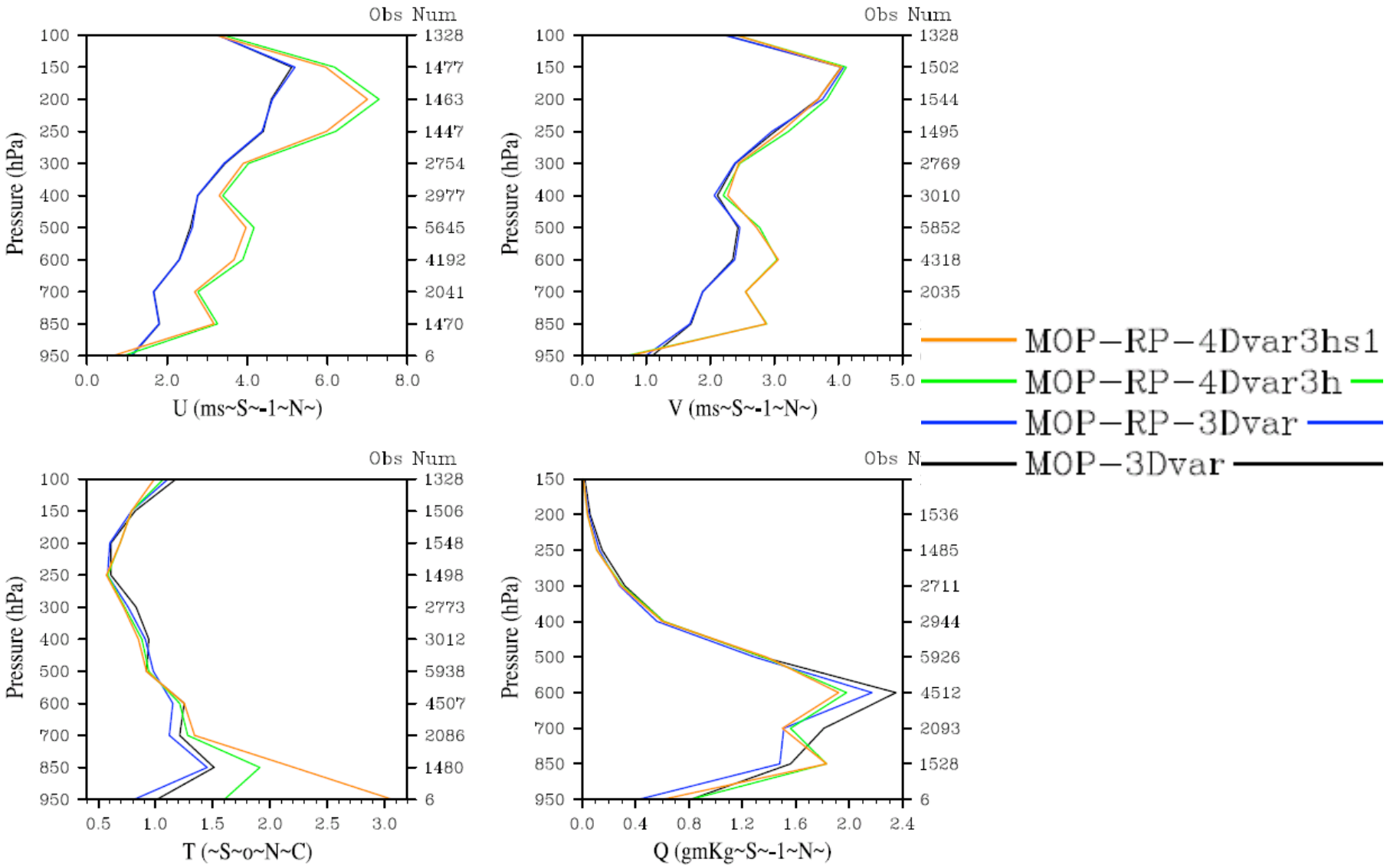


4D-Var experiments design

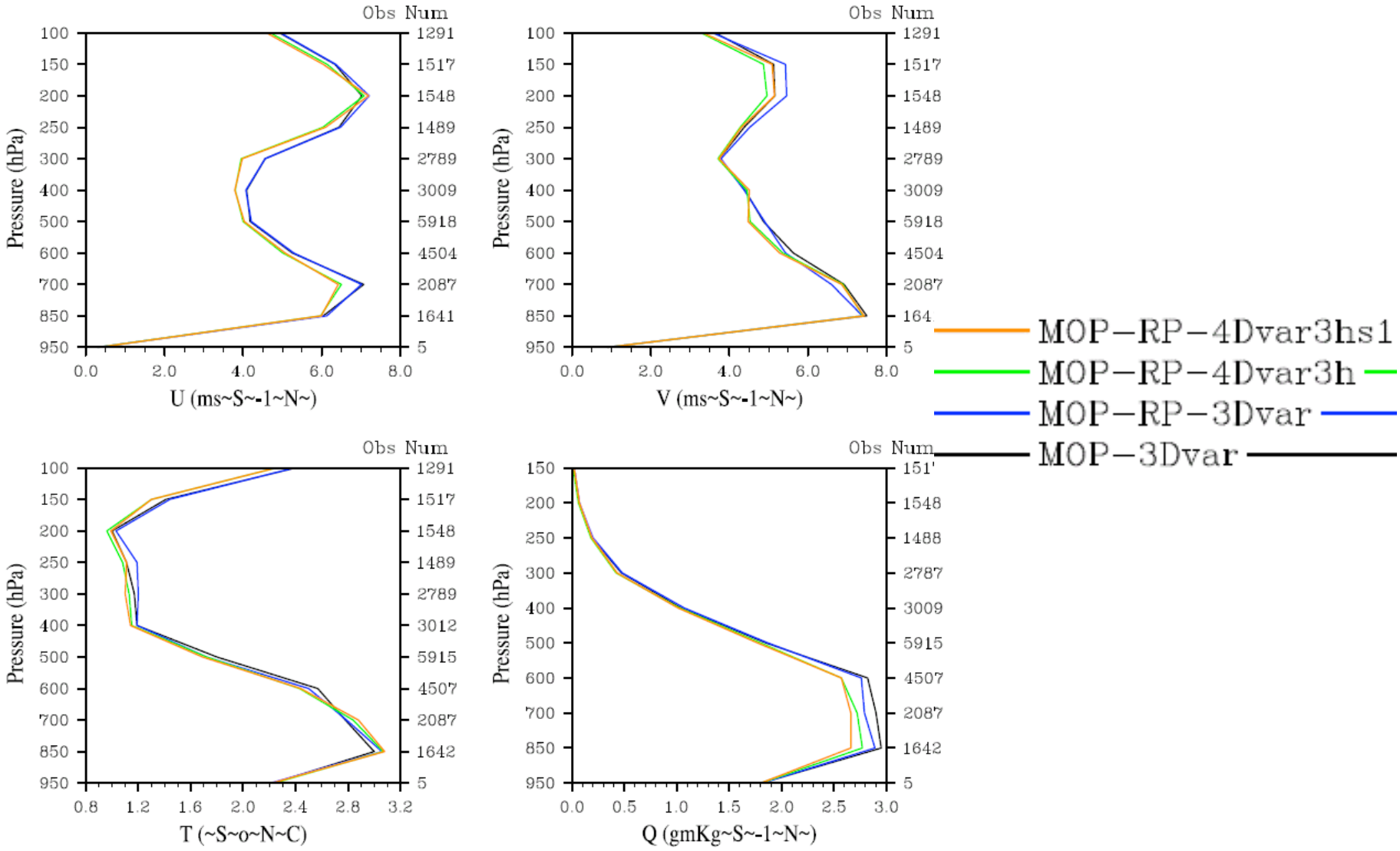
- Forecast model: WRF
- Data assimilation system: WRF-4Dvar
- Grid points: **85X85**X35
- Horizontal resolution: 12Km
- Time step: 60s
- Physics parameterizations:
  - Lin microphysics
  - Grell cumulus paramerization
  - MRF boundary layer
- Cases: 2002-06-11 12Z to 2002-06-12 12Z
- Background: Extracted from 18 h Control FCST at D1 (EC)
- Data:
  - MOP (simulated conventional data)
  - EOF retrieved profiles (18 levels, 100 km)
- Verification against truth



RMS error profiles at analysis



RMS error profiles at 12 h



Summary

- Three storms are well reproduced in the 5 day nature run.
- The calibration experiment shows that the real and simulated observations have the similar impacts on the analyses increments and forecasts differences.
- The quality of the retrievals has been improved significantly.
- The forecast skill is improved when MTG-IRS T and q retrieved profiles are assimilated.

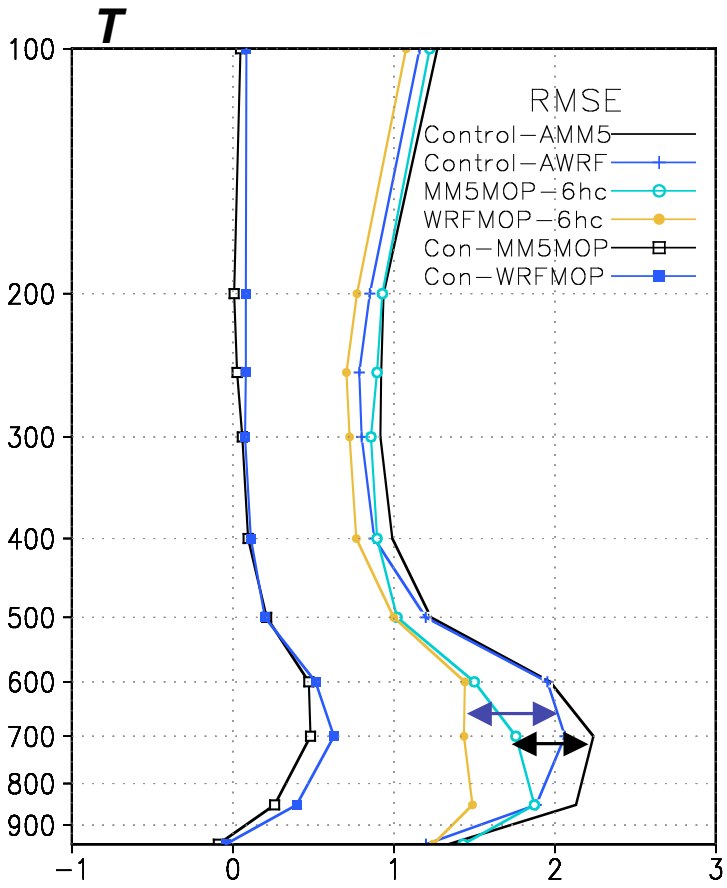
Future work

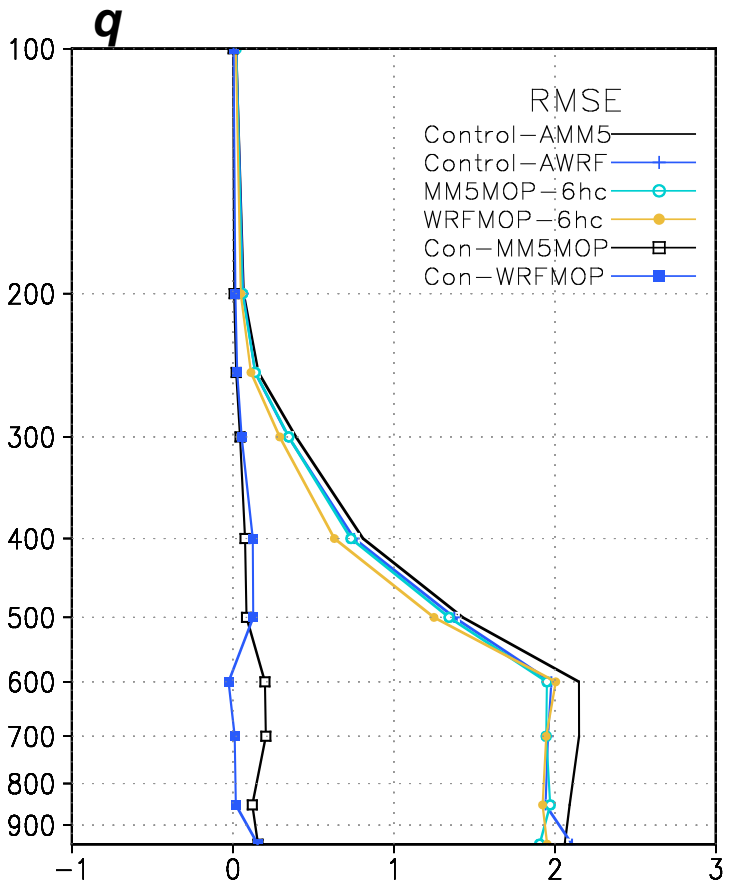
- WRF-4DVAR experiments
  - Cycling assimilation and forecast experiments
  - Time and computer source permitted
- Reduction of error correlations in MTG-IRS T(p) and q(p)
- Assimilating modeled wind observations from other platforms (such as wind profilers or radars)
- OSSE for European cases.
  - Two nature runs have been carried out

Identical Twin Experiments

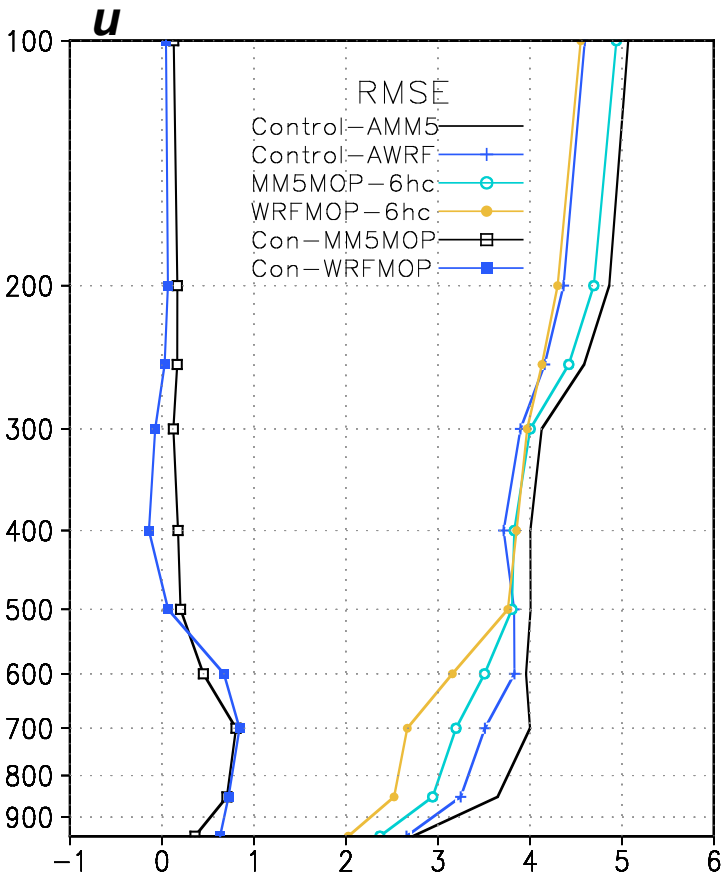
Experiment name	Observation data	Initial condition and assimilated data
Control		GFS analysis + perturbed lateral boundary conditions
MM5MOP-6hc	MM5-NR	Background (BG)+ Modeled Observation Profiles from MM5 nature run
WRFMOP-6hc	WRF-NR	Background (BG)+ Modeled Observation Profiles from WRF nature run

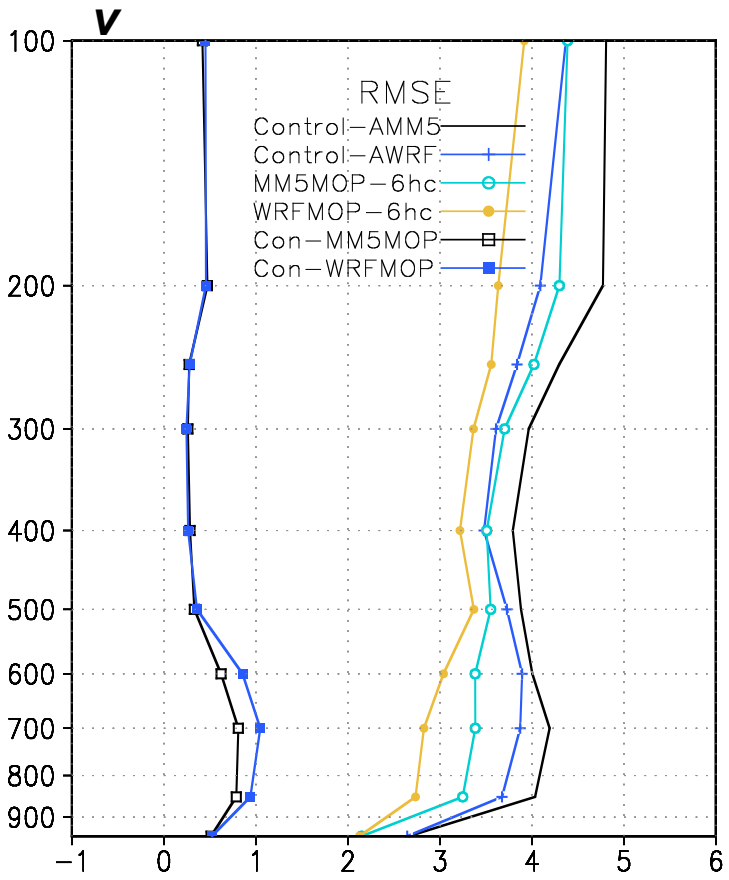
Averaged RMS error profiles at analysis time



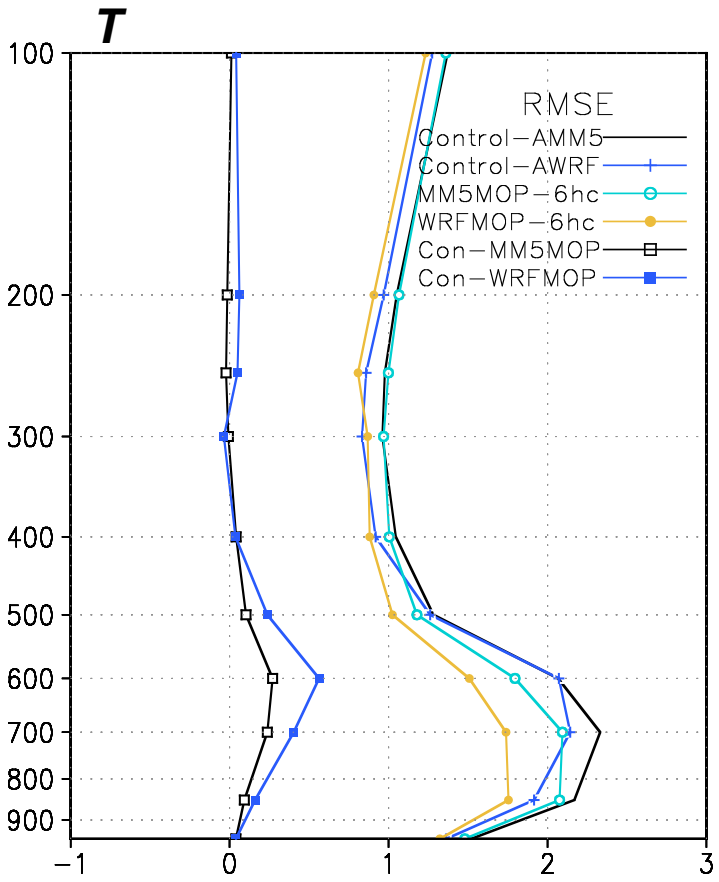


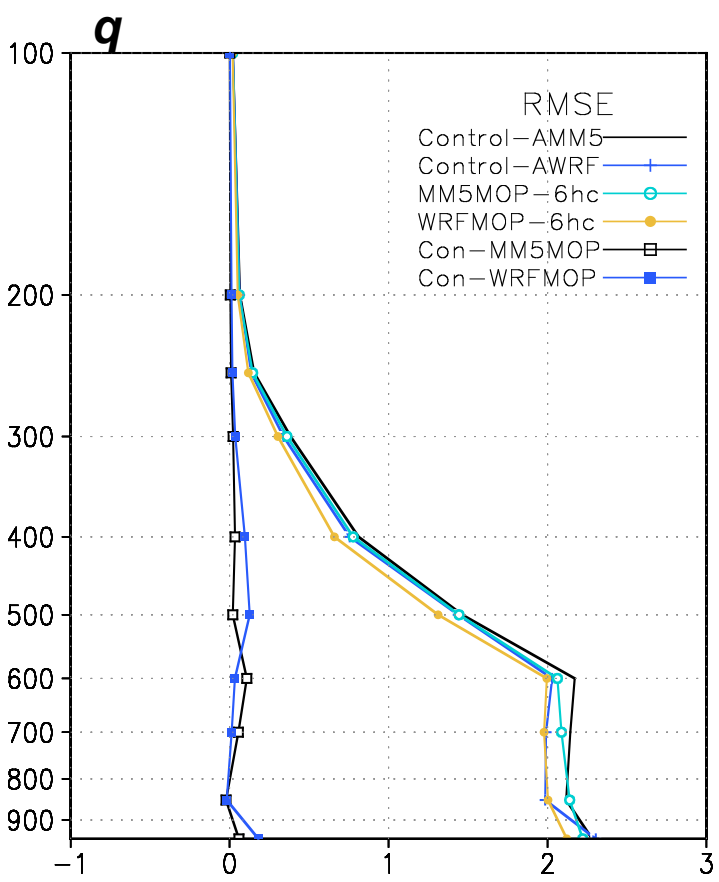
Averaged RMS error profiles at analysis time



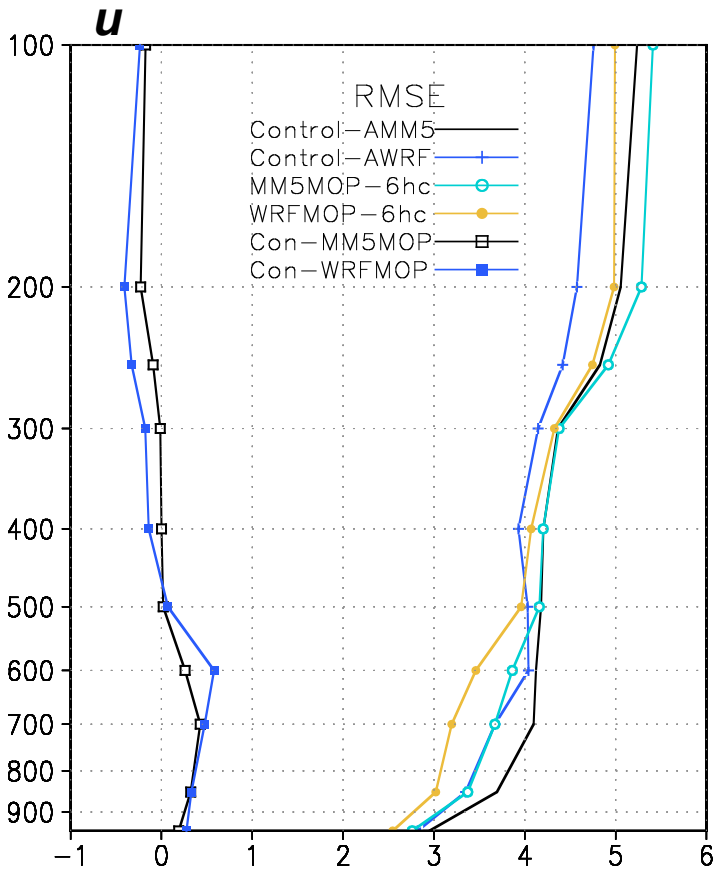


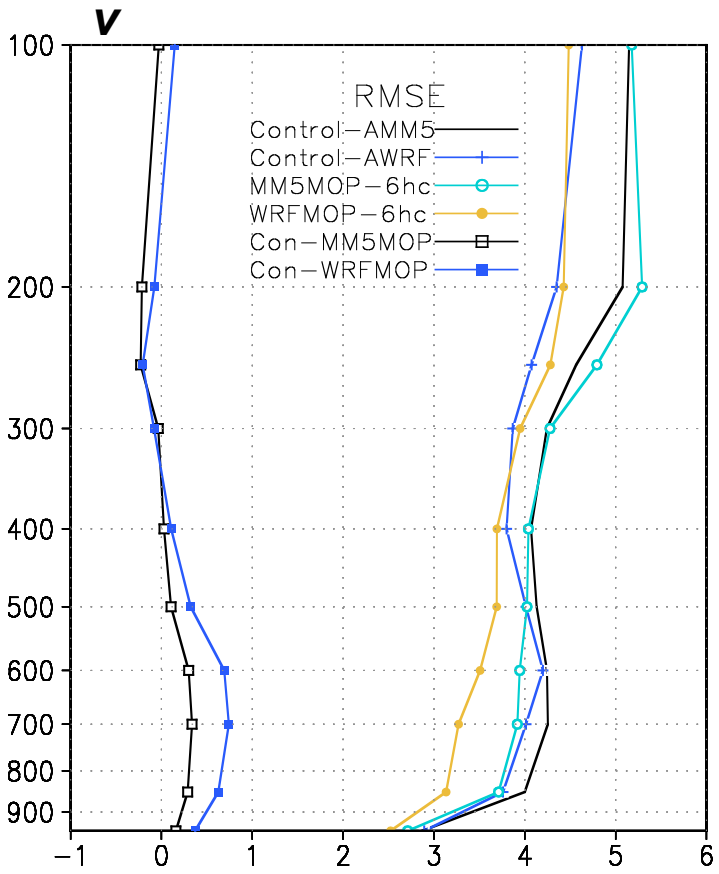
Averaged RMS error profiles at 12 h FCST





Averaged RMS error profiles at 12 h FCST





Compared with “non-identical” twin experiments

- The control experiment has smaller errors.
- The observation impact is larger.
- Identical-twin experiments are not too bad … may still be useful?!