Assimilation of ARM WVIOP-96 Data with the MM5-4DVAR System

Y.-R. Guo, Y.-H. Kuo, J. Dudhia, and D. B. Parsons
National Center for Atmospheric Research
Boulder, Colorado

Introduction

In the development of an Integrated Data Assimilation and Sounding System (IDASS) in support of the Atmospheric Radiation Measurement (ARM) Program, assimilation of heterogeneous mesoscale observations collected during the Water Vapor Intensive Observation Period of September 1996 (WVIOP-96) is one of the main tasks. In this study, the MM5-4DVAR system was used to assimilate these mesoscale observations and create a complete and accurate four-dimensional (4-D) dataset for diagnostic study in support of ARM.

Mesoscale Observations

During the WVIOP-96, there was a variety of mesoscale observations collected by the ARM Experiment Center (AEC), such as the ground-based Global Positioning System (GPS) precipitable water (PW), wind profiler data, surface parameters from several mesonets, and hourly rainfall, etc. These datasets have different spatial and temporal resolutions as well as measurement accuracy. The distribution of the observations used in this study are shown in Figure 1 and listed as follows:

- GPS PW data: 15 sites, 30-min. temporal resolution
- Wind profiler data: 20 sites, 250-m vertical resolution and 60-min. temporal resolution
- Surface data: 15 sites from Kansas mesonet with 60-min. temporal resolution, 111 sites from Oklahoma mesonet, and 13 sites from ARM mesonet with 30-min. temporal resolution, 35 sites from the National Weather Service (NWS) with 180-min. temporal resolution
- We also have the hourly rainfall data from Arkansas Basin Red River Forecast Center (ABRFC). Figure 2 shows the hourly rainfall ending at 0300, 0600, 0900 and 1200 UTC 19 September 1996.

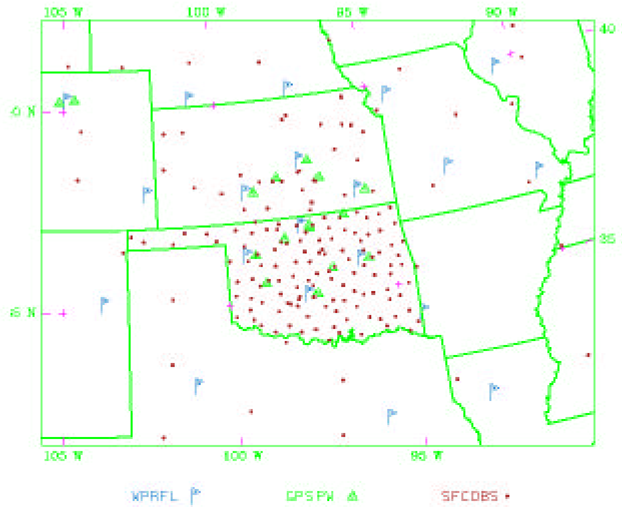

Figure 1. Model domain and mesoscale observations. (For a color version of this figure, please see *http://www.arm.gov/docs/documents/technical/conf_9803/guo-98.pdf*.)

MM5-4DVAR System and Experiment Design

Synoptic Case

The ARM WVIOP-96 lasted 3 weeks from 10 to 30 September 1996. The hourly precipitation over this period shows a well-defined squall line that developed and passed over the Kansas-Oklahoma area on 19 September (Figure 2). This convective case was chosen for this data assimilation study.

MM5-4DVAR System

The MM5-4DVAR system is a 4-D variational data assimilation system based on the Penn State/National Center for Atmospheric Research (NCAR) mesoscale model and its

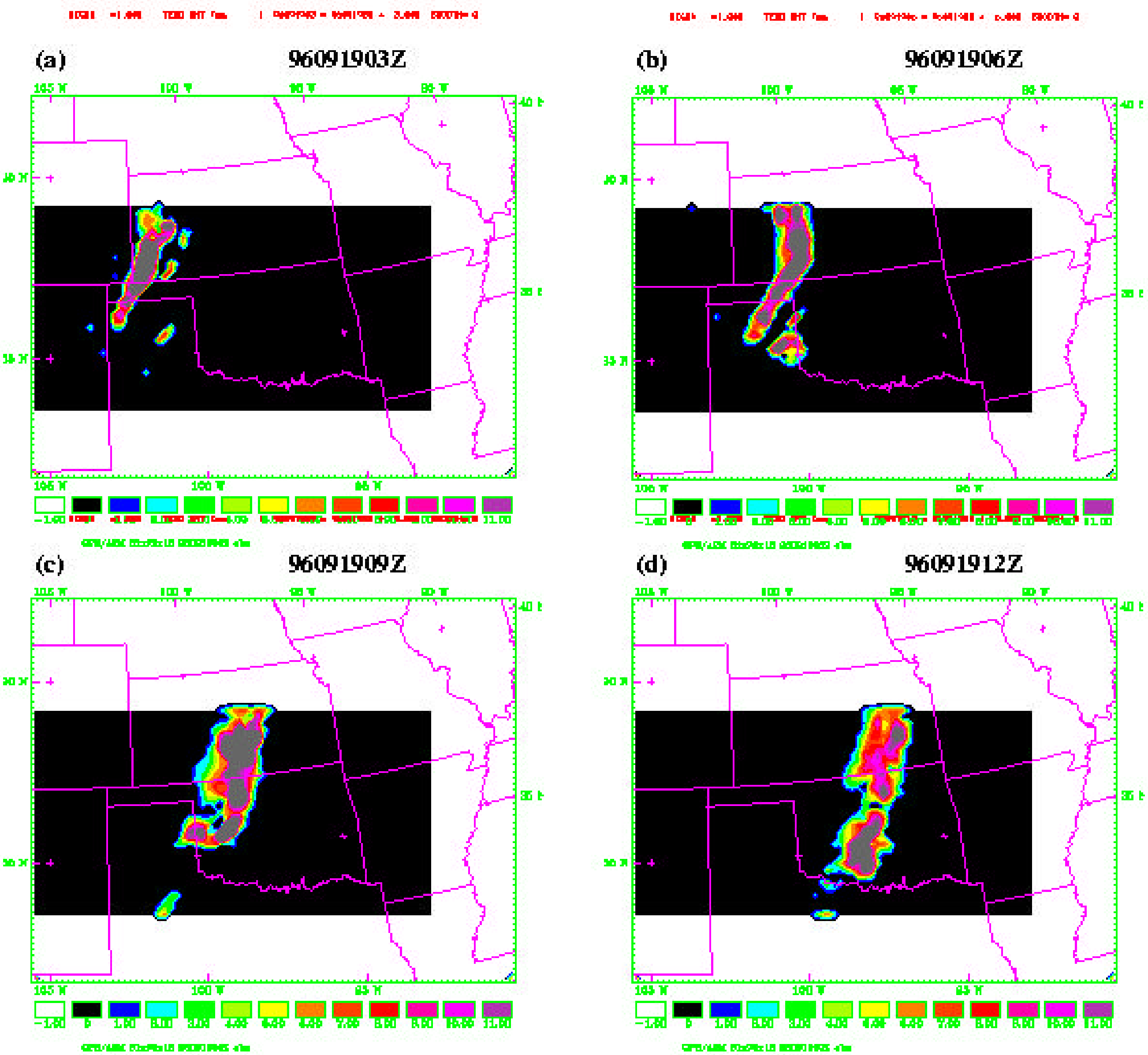

Figure 2. Hourly rainfall at (a) 0300, (b) 0600, (c) 0900, and (d) 1200 UTC 19 September 1996. The black area shows the data coverage. (For a color version of this figure, please see *http://www.arm.gov/docs/documents/technical/conf_9803/guo-98.pdf*.)

full-physics adjoint. The system includes many physics options. In this study, the physics packages of Bulk planetary boundary layers (PBL) with surface fluxes, Grell cumulus parameterization scheme, and Dudhia's explicit moisture scheme with ice effects, etc., were used. When one wants to assimilate a variety of observations with the MM5-4DVAR system, the observation operators, which transform the model variables to the observational quantities, and their adjoint must be developed. And also the weightings to each of the observations must be pre-specified. Here we defined five terms of the cost function: J_0, J_1, J_2, J_3, and J_4:

- J_0 is the background term. The gridded data from the model initial condition (obtained from an objective analysis of upper-air and surface observations) are used to represent this term.

- J_1 is the term for the GPS PW observations over the 6-hr assimilation window.

- J_2 is the term for the wind profiler observations over the 6-hr assimilation window.

- J_3 is the term for the surface dew point observations over the 6-hr assimilation window. We did not use other surface observational parameters because 1) we focus our attention on water vapor and 2) the direct assimilation of surface pressure and temperature can degrade the results based on previous experience.

- J_4 is the term for the hourly rainfall observations over the 6-hr assimilation window.

Experiment Design

We performed a series of data assimilation experiments over a 6-hr time window from 0000 UTC to 0600 UTC 19 September 1996. Five 12-hr forecasts are conducted starting from the original initial condition and the optimal initial conditions from each of the 4DVAR experiments (Figure 3 and Table 1). All experiments are conducted over a grid mesh of 61 x 79 with a 20-km grid distance and are driven by the hourly lateral boundary conditions from a 60-km coarse mesh model integration.

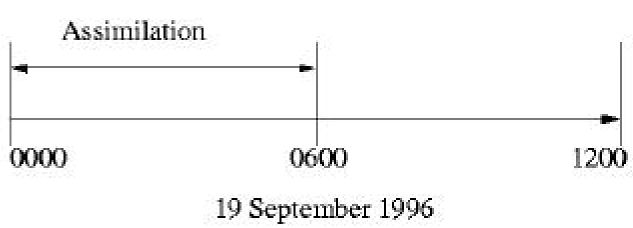

Figure 3. Diagram for showing the experiment design.

Table 1. Experiments.

Exp.	Name	Cost Function
1	NO4DVAR	forecast from original init. cond.
2	GPSPW	$J = J_0 + J_1$
3	PW+WPRF	$J = J_0 + J_1 + J_2$
4	PW+WPRF+TD	$J = J_0 + J_1 + J_2 + J_3$
5	PW+WPRF+TD+RN	$J = J_0 + J_1 + J_2 + J_3 + J_4$

The minimization procedure in 4DVAR experiments was terminated at 30 iterations.

4DVAR Performance

Figure 4 shows the performance of the 4DVAR experiments. The total cost function and the gradient norm decreased during the minimization process in all of the 4DVAR experiments (Figure 4a and Figure 4b). The MM5-4DVAR system worked well with these heterogeneous observations. But the absolute values of the cost function and gradient for Exp. 5, including all observations, are much larger than those for Exp. 2, including only the background and the GPS PW terms. Figures 4c and 4d show changes of each term in the cost function during the minimization process for Exp. 2 (J_0 and J_1) and Exp. 5 (J_0, J_1, J_2, J_3, and J_4). The largest term is the wind profiler term, and the smallest term is the GPS PW term. At the end of minimization, J_0 for Exp. 5 is much larger than that for Exp. 2. This means that the final analysis ("optimal" initial condition) is more deviated from the original analysis in Exp. 5 after all the observations are assimilated.

Results

Precipitable Water Errors

Figure 5 shows the time series of the PW root mean square (rms) errors verified against the GPS PW observations. During the assimilation window from 0000 UTC to 0600 UTC 19 September 1996, all the 4DVAR Exps reduced the rms errors significantly as compared with Exp. 1 (NO4DVAR). This shows that the MM5-4DVAR system incorporated these observations into the model well. The lowest errors are from Exp. 2 and the highest from Exp. 5. This is not surprising based on the cost function reductions (Figure 4). Beyond the assimilation window from 0600 UTC to 1200 UTC 19 September 1996; however, the errors for all the experiments grow rapidly. There are many factors that may cause this fast error growth, some of them are related to the 4DVAR component of the system, and some related to the MM5 model itself. The Exp. 3 (PW+WPRF), however, gives a steady improvement beyond the assimilation window even though the improvements are rather small. The wind profiler observations are the only 4-D dataset that is broadly spread in the model domain. Other data used here are very localized or available only at a single level (Figure 1).

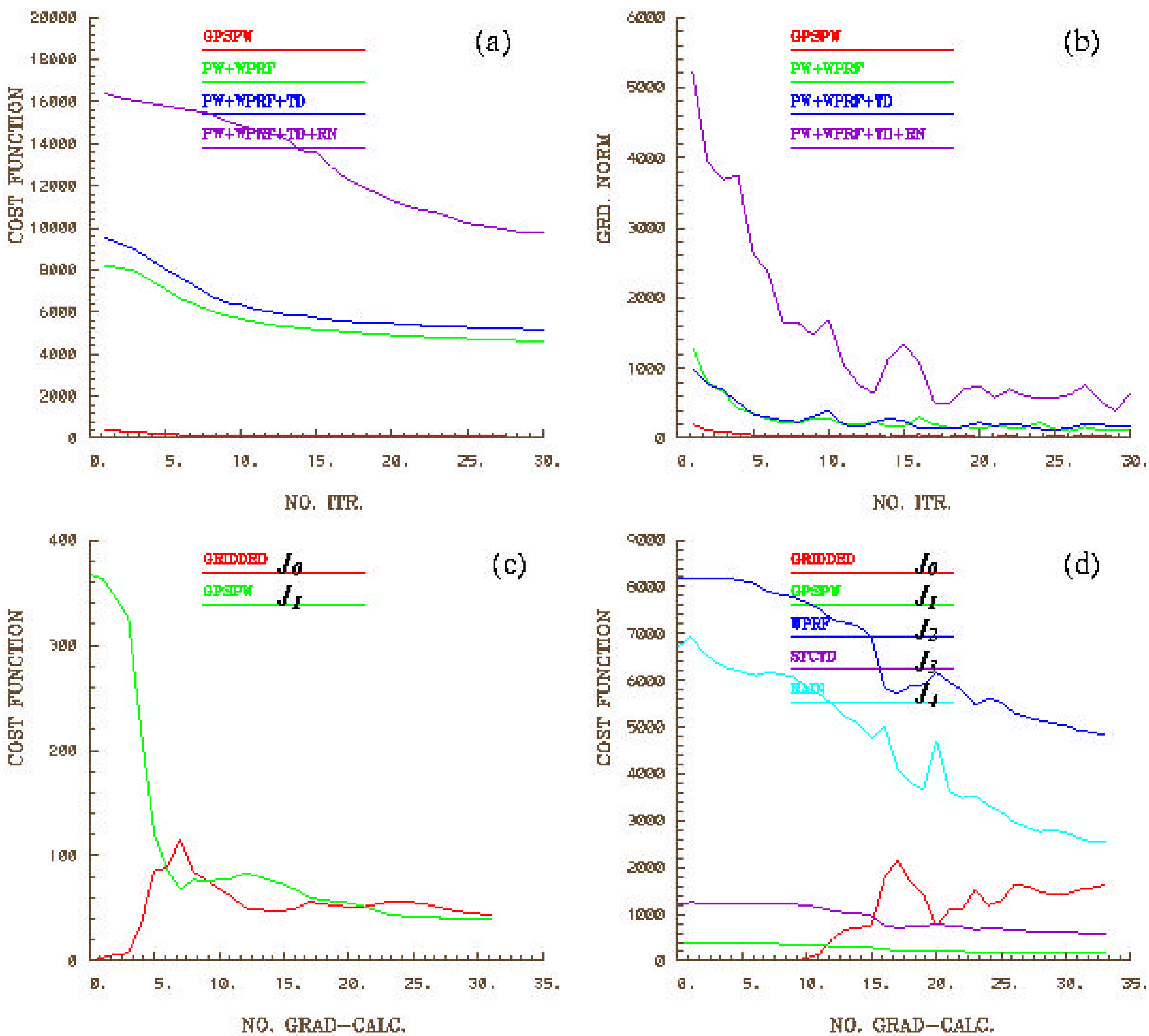

Figure 4. 4DVAR performance. (a) The cost function J and (b) gradient norm for Exps. 2, 3, 4, and 5. (c) The term J_0 and J_1 for Exp. 2, and term J_0, J_1, J_2, J_3, and J_4 for Exp. 5. (For a color version of this figure, please see *http://www.arm.gov/docs/documents/technical/conf_9803/guo-98.pdf*.)

Wind rms Errors

We also calculated the wind errors verifying against the wind profiler observations. Figure 6 shows that the 4DVAR experiments with the cost function including the wind profiler data give significant error reductions during the assimilation period. Without the wind profiler data included in Exp. 2, there is no error reduction even in the assimilation period, and error is even a little bit greater than that in Exp. 1 (NO4DVAR). The very localized GPS PW observation did not help to improve the wind structure through the 4DVAR procedure. With the rainfall data included, an error increase is found when compared with Exps. 3 and 4. This might be caused by an improper weighting specification for rainfall data or inadequacy in the model precipitation physics (Figure 4). Again, Exp. 3 (PW+WRPF) gave the best result in terms of the wind error reduction during the entire 12-hr period. Addition of surface dew point data does not lead to further reduction of wind and PW errors (Figures 5 and 6).

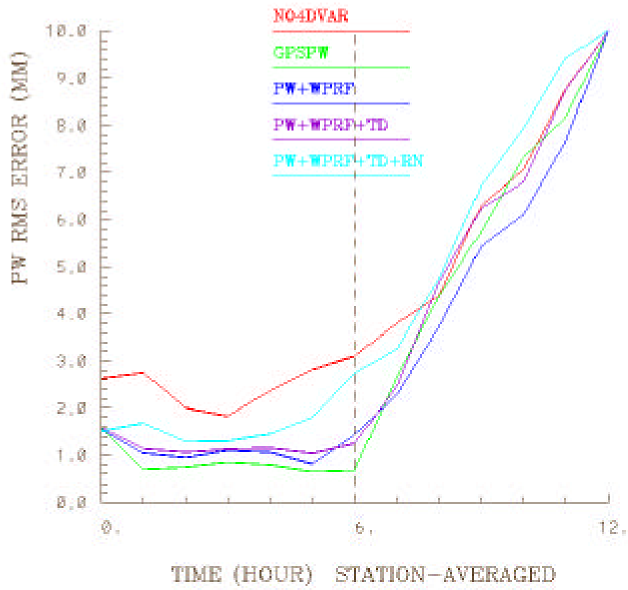

Figure 5. The time series of PW rms errors verified against the GPS PW observations. (For a color version of this figure, please see *http://www.arm.gov/docs/documents/technical/conf_9803/guo-98.pdf*.)

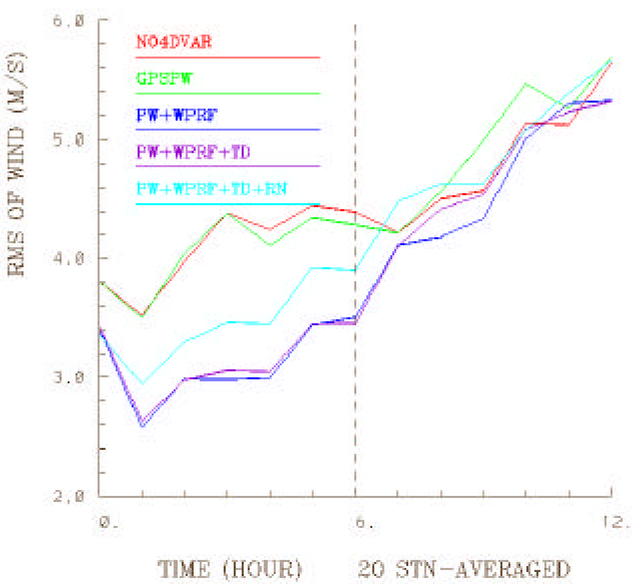

Figure 6. The rms errors of wind for all the experiments verifying against the wind profiler observations. (For a color version of this figure, please see *http://www.arm.gov/docs/documents/technical/conf_9803/guo-98.pdf*.)

Hourly Rainfall Forecasts

Figure 7 shows the hourly rainfall forecasts ending at 0600 UTC and 1200 UTC 19 September 1996, respectively. At the end of the assimilation window, 0600 UTC 19 September, Exp. 5, with the rainfall data included, gives the best rainfall forecast over the Oklahoma-Kansas area (Figures 2b and 7e). Figures 7a, 7b, and 7c show that the assimilation of the wind profiler data (Exp. 3) improves the rainfall forecast at the end of the assimilation period. By the end of the 12-hr forecast period, the strong convection over southern Oklahoma (Figure 2d) is not predicted in all the experiments, although the forecasts of Exps. 3, 4, and 5 are a little bit better than that of Exp. 1 (NO4DVAR).

Conclusions

1. The MM5-4DVAR system successfully assimilated the heterogeneous observations collected in ARM WVIOP-96. The minimization procedure worked well with these datasets, and the 4DVAR procedure led a significant error reduction during the assimilation period.

2. The assimilation of GPS PW plus wind profiler observations had the smallest errors of PW and winds during the entire 12-hr experimental period, although the error grew fast beyond the assimilation period. For the limited area model, the lateral boundary condition can seriously influence the forecast in the interior of the domain. In this study, we did not adjust the lateral boundary condition for the 4DVAR procedure. Optimal control of lateral boundary condition will be carried out in future work.

3. In this study, the wind profiler observations are shown to be more effective than other observations. Here the wind profiler data are the only 4-D dataset within the model domain. They can improve the upstream conditions for the convective region. Also the cost function term related to wind profiler data is the biggest one (Figure 4d), and the minimization may rely more heavily on this term. We may need to tune the coefficients or weightings in the 4DVAR system for other observations. Certainly, additional observations should be included in the future.

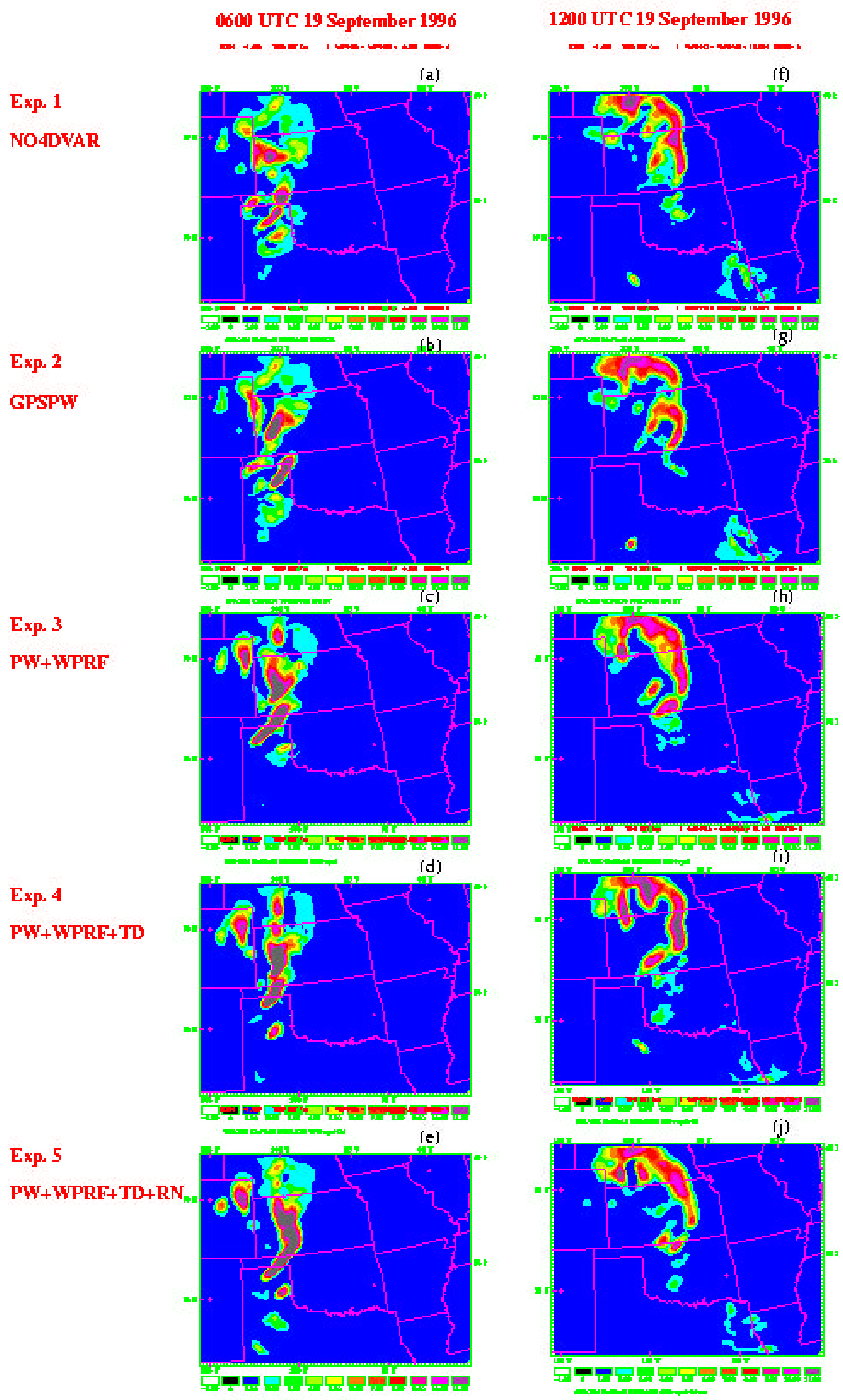

Figure 7. The hourly rainfall forecasts for all the experiments ending at 0600 UTC and 1200 UTC 19 September 1996, respectively. (For a color version of this figure, please see *http://www.arm.gov/docs/documents/technical/ conf_9803/guo-98.pdf*.)

4. The case of 19 September 1996 is a well-defined convection case. During the period of 0000 UTC to 0600 UTC 19 September 1996, the squall line is located upstream of the GPS PW network, ARM, and Oklahoma and Kansas mesonets. Assimilation of the observations from this observed network during 0000 UTC to 0600 UTC may not help to improve forecast of convection development in the next 6-hr period when the squall line moves downstream to the middle of the domain. It would be desirable to attempt the assimilation experiments over the period of 0600 UTC to 1200 UTC for this case.