2.8 ASSIMILATION OF DIVERSE METOROLOGICAL DATASETS WITH A FOUR-DIMENSIONAL MESOSCALE ANALYSIS AND FORECAST SYSTEM

Yubao Liu*, Wei Yu, Francois Vandenberghe, Andrea Hahmann, Tom Warner and Scott Swerdlin
National Center for Atmospheric Research/RAP, Boulder, Colorado

1. INTRODUCTION

In collaboration with the Army Test and Evaluation Command (ATEC), NCAR/RAL has been developing a multi-scale, rapid-cycling, Real-Time Four-Dimensional Data Assimilation (FDDA) and forecasting system (RTFDDA). In the last five years, the system has undergone continuous enhancement, and has been successfully deployed at six Army test ranges. It has also been used to support operations of various homeland-security missions, including the anti-terrorism support during SLC-2002 and Athens-2004 Olympics, and Joint-Urban 2003 Field Experiment in Oklahoma City, and a number of other government and international applications. Some recent improvements to this system are described in Liu et al. (2002) and Bowers et al. (2003).

Mesoscale (10 - 2000 km) meteorological data assimilation and prediction are challenging due to the sparseness of observations, especially in the upper-air atmosphere. In the past 15 years, a number of instruments, e.g. wind profilers, commercial aircraft reports, satellite measurements, Doppler radars, and most recently, TAMDAR (Tropospheric Airborne Meteorological Data Reporting), have been developed to enhance the upper-air observations. Various surface mesonets from public agencies and private companies make a very high density and frequency of surface weather observations available for mesoscale data analysis and forecast. The diverse observations from these conventional and unconventional, synoptic and asynoptic, observation platforms have been incorporated into the RTFDDA model system. In this paper, impacts of different observation platforms are evaluated and compared with numerical studies using the RTFDDA system. Sensitivity experiments are also carried out with sub-groups of observations to study the relative roles of different observations, such as surface data versus upper-air data, temperature versus winds observations, and in-situ observations versus remote-sensing volume measurements.

2. SCIENTIFIC REVIEW OF RTFDDA SYSTEM

The RTFDDA is an MM5/WRF-based, multi-scale (grid-sizes varying from 0.5 km to 45 km), rapid cycling, real-time, four-dimensional data assimilation and forecasting system. By employing continuous data assimilation through the Newtonian-relaxation method, the system produces four-dimensional, dynamically and physically consistent, analyses and short-term (0 - 48 hour) weather forecasts. The forecast system can be cycled (forecasts initiated) at time intervals of 1 - 12 hours, and the use of continuous data assimilation minimizes the spin-up problem that can be associated with intermittent data assimilation.

Data assimilation and forecasting on the meso-beta and gamma scales face many challenges. Observations are sparse relative to grid sizes, and they are sometimes available irregularly in time and space. Local circulations associated with orography and other surface heterogeneities may experience serious "spin-up" problems if improperly initialized. There are also spin-up problems associated with cloud and precipitation. The shortage of observations and small simulation domains also make it extremely important to properly simulate the synoptic weather processes as well, and their forcing through lateral boundary.

The RTFDDA system was designed to address these challenges. It makes use of two-way nested grids to model multi-scale interactions, from synoptic-scale to meso-gamma-scale circulations. Data assimilation is performed on all domains. Because the sparse observations are not sufficient to produce accurate three-dimensional analyses of the complex circulations, conventional intermittent analysis methods are not applicable. Instead, the so-called "station nudging" method was employed in the RTFDDA system. A basic station nudging procedure was firstly introduced to the MM4(5) system by Stauffer and Seaman (1994). It was adapted to real-time applications in RTFDDA by Cram et al. (2001), and has undergone significant enhancement since then (Liu et al 2002).

The station-nudging approach allows sequential insertion of each observation into a continuously running, full-physics, MM5/WRF model, with proper temporal and spatial weights. This algorithm is especially adequately for assimilating irregular observations such as ACARS and TAMDAR weather reports along flight

* Corresponding author address: Yubao Liu, National Center for Atmospheric Research, P. O. Box 3000, Boulder, CO 80307-3000. Phone: (303) 497-8211 Email: yliu@ucar.edu

tracks. As observations are nudged into the model solution, the model state is forced toward the observations. The non-observed model variables adjust under the constraints imposed by the full equations of the model. As time progresses, the model spreads the observation information in three-dimensional space according to the local weather circulations. This four-dimensional, multivariable analysis method allows the model physics to represent the local and synoptic conditions in regions where no observations are available, and corrects the model solution in the regions where observations exist. Experiments show that this four-dimensional data assimilation method is not only capable of properly assimilating information from sparse observation networks, but it is also well behaved when observations are abundant.

3. DATA SOURCES

The conventional data types assimilated by the system include the twice-daily radiosondes and pibal winds, hourly surface observations at synoptic stations, data from ships and buoys, and data from other special reports disseminated by WMO (World Meteorological Organization). High-frequency measurements from special observation platforms at the Army test ranges, such as 15-minute averages from surface mesonet stations and boundary-layer profilers, are utilized, as are range radiosonde soundings. The four-dimensional data assimilation approach used here makes full use of the above conventional and special observations that are available at both synoptic and asynoptic times. Nevertheless, these data sources still only represent a very limited portion of the atmospheric at the time and space scales of interest. For example, between the 12-hourly synoptic radiosonde observation times, we hardly obtain any upper-air information from these data sources. It is obvious that incorporating more data sources should be one of the major objectives in enhancing this system.

The capability of the RTFDDA system to utilize observations at irregular times and locations gives us the flexibility to explore the usefulness of a few advanced, non-conventional observation platforms that have emerged in recent years. At surface, the high-density and high-frequency observation from a large number of public and private agencies, are collected in the MADIS (Meteorological Assimilation Data Ingest System) data set of the NOAA/FSL. A total of about 18,000 surface weather stations over the CONUS are collected in real-time at present. For upper-air, in collaboration with NOAA/NESDIS (National Environmental Satellite, Data, and Information Service; Chris Velden and Jaime Daniels), the experimental hourly rapid-scan GOES 8 and 9 (now 10 and 12) winds that are derived from visible, IR and water vapor images, were ingested into the system. Beginning at about the same time, data from the NPN (NOAA Profiler Network) and the CAP (Cooperative Agency Profilers) network, disseminated by NOAA/FSL, were incorporated. Aircraft reports (ACARS/AMDAR - Aircraft Communications Addressing and Reporting System/Aircraft Meteorological Data Relay) processed and disseminated by NOAA/FSL were also successfully tested with the RTFDDA system. Recently, NEXRAD VAD, Radiometrics Radiometers, NASA/JPL QuikScat sea surface winds, and AirDat LLC TAMDAR data are also incorporated into the RTFDDA system. These new data sources enhance greatly the time and space coverage, and with their high quality and real-time availability the data sources fill the temporal gap between the synoptic radiosonde observations. Fig. 1a summarizes the observations used in the Army White Sand Missile Range (WSMR) RTFDDA system at 00Z 28 July 2003. Fig. 1b

Observation totals used at 2003072800 a
Eta run used for lateral bc: 2003072712
Cycle stage: final

Observation Totals	Grid 1	Grid 2	Grid 3
Total obs	13606	761	197
Metar obs	1610	93	22
Special obs	66	8	0
Ship obs	41	0	0
Mesowest obs	1661	49	9
Raob T,Td,Wind obs	36	2	0
Sat-wind & ACARS obs	10011	490	49
RANGE SAMS obs	115	115	115
Profiler obs	64	4	2

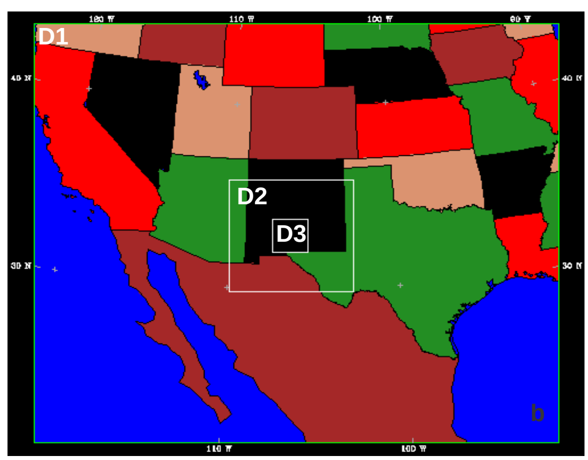

Fig.1 (a) Observations used in the Army's White Sand Missile Range (WSMR) RTFDDA system at 00Z 28 July 2003. (b). WSMR RTFDDA system domain configuration.

shows the model domain configuration. The sum of the observations from the new upp-air platforms is more than five times the total number of observations from radiosonde reports in Domain 1 (radiosondes have 70 levels, on average). There are significant additions of data to the areas of the fine-mesh grids also. It is worth pointing out that there are more ACARS and TAMDAR reports and satellite winds reports in the daytime than at night. The reported data numbers are only those that passed rigorous quality control checks, and were used in the system during an 80-minute time window.

4. DATA IMPACT EXPERIMENTS

Measurements from different platforms and sensors possess different qualities, and temporal and spatial densities. The sample volumes can be very different. And a platform may observe only one or two weather variables. Therefore, it is important to investigate how these observations with their particular attributes affect the RTFDDA forecast skill. Numerical experiments were carried out to study the impact of 1) subsets of the observations from different platforms, 2) surface versus upper-air data, and 3) different variables. In this section, results from the experiments with the Army's WSMR RTFDDA system, running with a single coarse domain (grid size of 30 km) for a 6-day period, will be presented. In all experiments shown below, the RTFDDA system ran with 3-hourly cycling (forecast initiation frequency), with 12 hour forecasts in each cycle. The 6-day model runs are evaluated by verifying the model 10-12 hour forecasts against all hourly surface observations and 12 hour upper-air forecasts against all 00Z and 12Z radiosonde observations. Note that the shorter-duration forecasts, and the analyses, are more accurate (Cram et al. 2001). Finally, we point out that these experiments were conducted a few months ago when only 3-hourly satellite winds (rather than hourly) and a subset of ACARS reports (those from United Airlines only) were available.

Two sets of experiments, with two different weather scenarios respectively, were carried out to study the data impact. The first set is made up of a total of 8 experiments. The second set of the experiments contains 20 experiments with a summer convection case in which severe storms wiping through the northeastern States. We are in process of analyzing model output of the second set. This paper focuses on the first set. The experiments were run with the same version of the system, with the same model physics. Each experiment was run with a subset of the observations. The experiment abbreviations are listed below,

A). CTRL -- control experiment, uses all observations,

B). NOobs -- no observations are assimilated,

C). SFConly -- uses (all types of) surface observations only,

D). PROFSFC -- the same as SFConly, but with the addition of wind-profiler data,

E). SATWSFC -- the same as SFConly, but with the addition of satellite winds,

F). UPRonly -- uses (all types of) upper-air observations only,

G). TQonly -- assimilates (both surface and upper-air) temperature and moisture only, and

H). WINDonly -- assimilates (both surface and upper-air) wind only.

As will be seen from the model results, the control experiment (CTRL) that made use of the full data set, and the experiment NOobs that used no observations at all, represent the upper and lower bounds of the system performance, respectively.

4.1 Impact of Surface and Upper-air observations

One interest of the data assimilation and NWP (Numerical Weather Prediction) communities is the impact of surface and upper-air observations on forecast skill. Specifically, it is important to investigate how effectively surface data can correct the model error near the surface, and how these data influence the three-dimensional upper-air forecast. Similarly, we are also interested in how upper-air observations can correct the upper-air model error, and how these data affect the forecast near the surface. The continuous FDDA approach allows the surface layer and the free atmosphere to interact through the PBL dynamics. Thus, surface-observation information can propagate into the upper-atmosphere, and vice versa.

Four experiments, CTRL, SFConly, UPRonly and NOobs are compared here. The verification results (Figs. 2 and 3) are very interesting. First of all, as one could anticipate, by fully assimilating all observations, the CTRL runs produced the best results in terms of both upper-air and surface forecast verification statistics. Conversely, the runs without assimilation of any observations, NOobs, generated the largest forecast errors. Assimilating surface observations only, SFConly, dramatically improves the quality of the surface analyses and forecasts. As compared with NOobs, SFConly reduces the moisture RMSE error of the 10-12 h surface forecasts by 20-35% and temperature by ~0.5-1C (about 20-30%). The surface winds are minimally affected. Similarly, UPRonly was able to improve the upper-air model wind, temperature and moisture forecasts. The RMSE (Root Mean-Squared

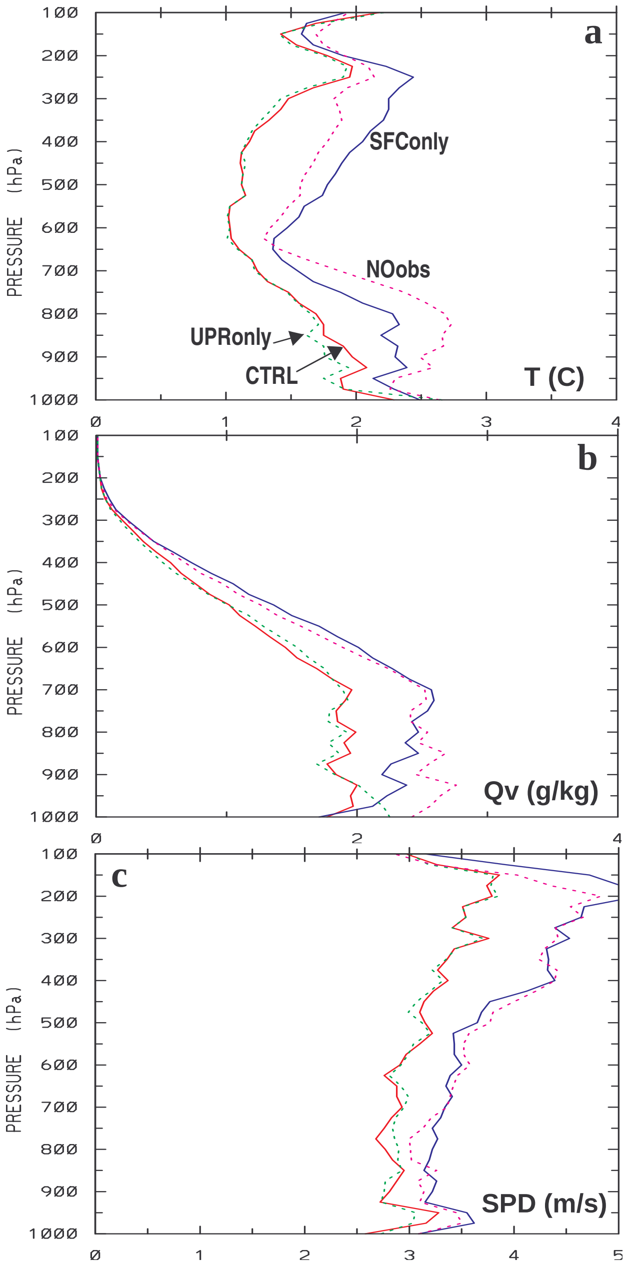

Fig.2 RMSE of the 12th hour upper-air forecasts of CRTL, SFConly, UPRonly and NOobs, verified against 00Z and 12Z soundings.

Error) errors are reduced by 20-40%, with the largest corrections for temperature and moisture in the lower troposphere, and for winds in the upper-troposphere.

The positive impact of upper-air data on upper-air forecast skill, and the impact of surface data on the skill of surface forecasts is easy understood. It is perhaps more interesting to explore how the surface data affect upper-air forecasts and what benefit the upper-air observation can bring to the surface-forecast skill. Comparing the RMSE errors of the upper-air, hour-12 forecasts for SFConly with NOobs, it is clear that the

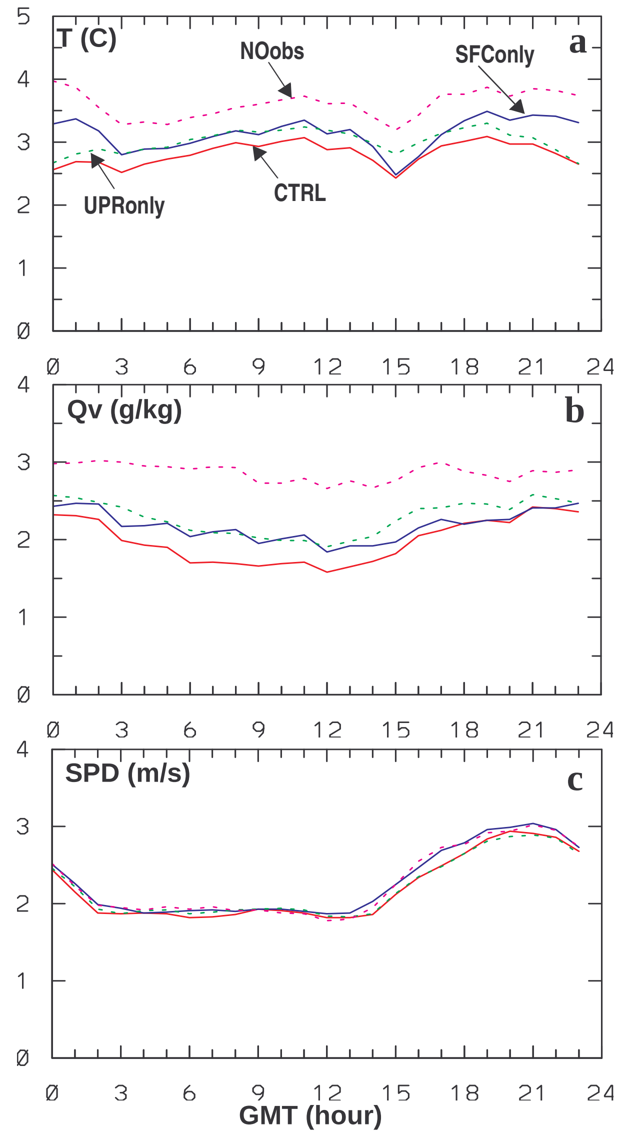

Fig.3 RMSE of the 10th-12th hour surface forecasts of CRTL, SFConly, UPRonly and NOobs, verified against all surface observations

effect of the surface data extends throughout the model layers. Significant improvements can be observed in the lower-tropospheric (below 700 hPa) temperature and humidity. The largest correction of moisture occurs at the surface, gradually decreasing upward. In general, surface observations do not improve the upper-air winds. Surface data degrade the forecasts of temperature above 600 hPa, and winds around the tropopause. This might be explained by the fact that surface observations are not capable of correcting the phase errors of synoptic systems, and the correction of the model error in the lower troposphere by the surface observations may cause a dynamic inconsistency in the vertical structure.

By comparing the diurnal evolution of the RMSE for the model 10th-12th hour surface forecasts of UPRonly and NOobs, it can be seen that, unlike the surface data impacts on the upper-air model state, the upper-air data consistently benefit the surface-weather forecasts. UPRonly significantly reduced the errors in the model surface-layer temperature and moisture. The near-surface wind speeds and directions (not shown) are also improved. The positive impact of the upper-air observations on both surface and upper-air forecasts indicates that the use of the upper-air data can correct the phase errors of the synoptic weather systems, and that the model PBL transfers this benefit to the surface.

Finally, we compare the SFConly and UPRonly with the CTRL runs. Using the full set of observations produces model results that are better than both SFConly and UPRonly. In particular, CTRL further reduced the RMSE from SFConly by 0.3 C for surface temperature, 0.3 g/km for specific humidity and 0.1 m/s for wind speed. Obviously, the good performance of the CTRL run results from the correction of the error in the synoptic systems by the upper-air data, and from the acceptance of the surface data by the PBL physics.

4.2 Impact of Wind Profilers and Satellite Winds

Both wind profiler and satellite-based wind observations provide winds only. Wind profilers measure winds at fixed horizontal locations and height levels. Most of the profilers are over the Central Plains region. In contrast, the locations and heights of the satellite observed winds depend on the weather systems. Therefore, it is of great interest to see how these two observation platforms affect the RTFDDA results.

The Experiments PROFSFC and SATWSFC are compared with the Experiments SFConly and CTRL. The RMSE of upper-air temperature, specific humidity and wind speed of the 12th hour forecasts from the experiments are presented in Fig.4. Assimilating upper-air winds from both satellite and wind profiler platforms significantly improves the upper-air flow, and the temperature and moisture fields (through model internal adjustment) benefit as well. It is interesting to see that, although there are large differences in the temporal and spatial distributions and other attributes of the wind observations from the two platforms, their overall impacts on the model winds, temperature and moisture are similar. On the other hand, differences do exist. For example, because satellites report fewer in the lower troposphere than do the profilers, the improvement from

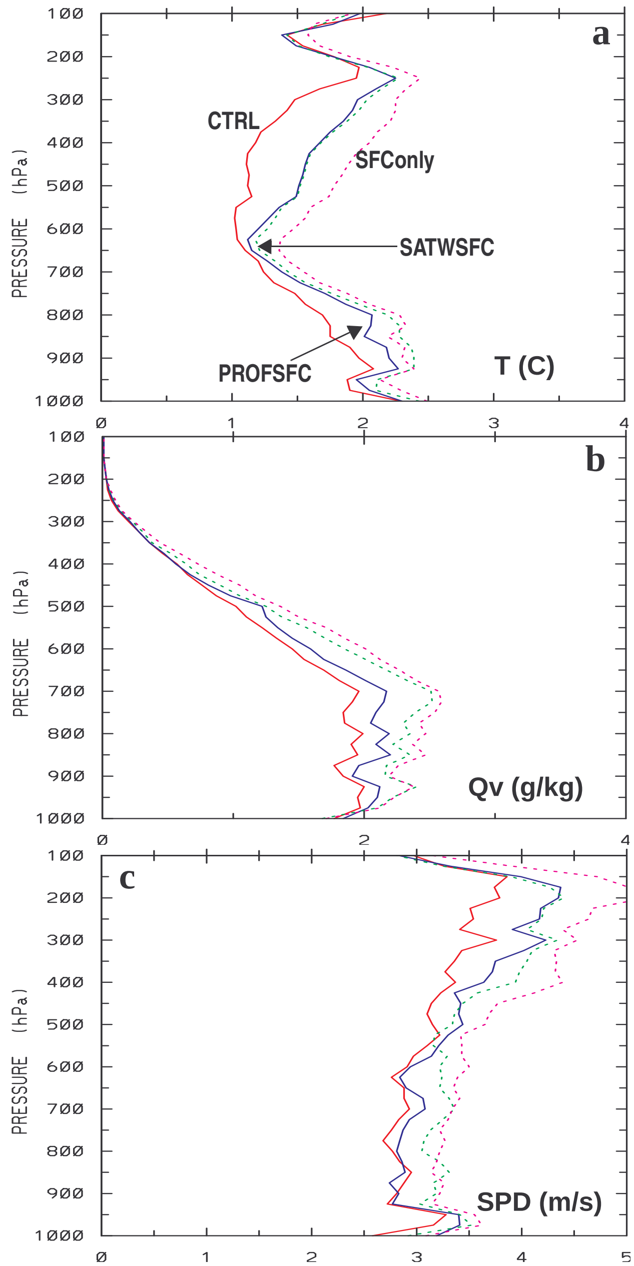

Fig.4 Same as Fig.2, but for CRTL, PROFSFC, SATWSFC and SFConly.

the profiler data appears to be much larger than from the satellites in the layer.

4.3 Impact of Different Observation Variables

Under the constraint of the full-physics model, the RTFDDA system is able to propagate observation information from the observed variables to non-observed variables. However, because of the inherent nature of the atmospheric dynamics and the model approximations, the roles and adjustment processes associated with assimilating different model variables can be different. To address the issue, the experiments

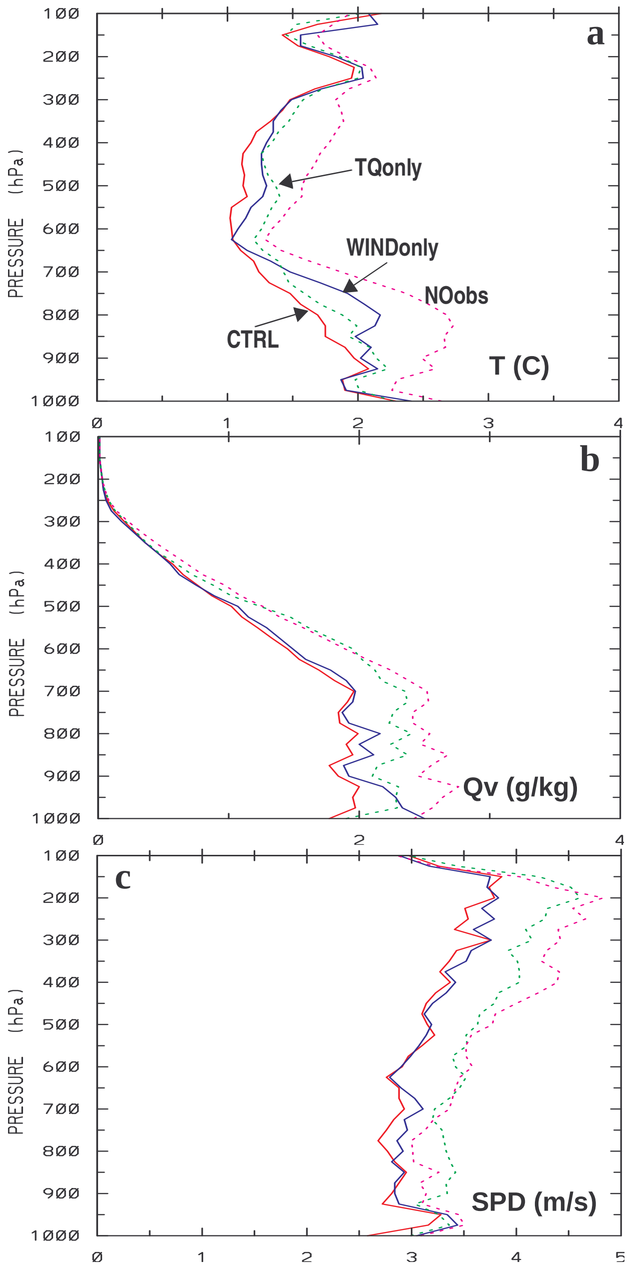

Fig.5 Same as Fig.2, but for CRTL, TQonly, WINDonly and NOobs.

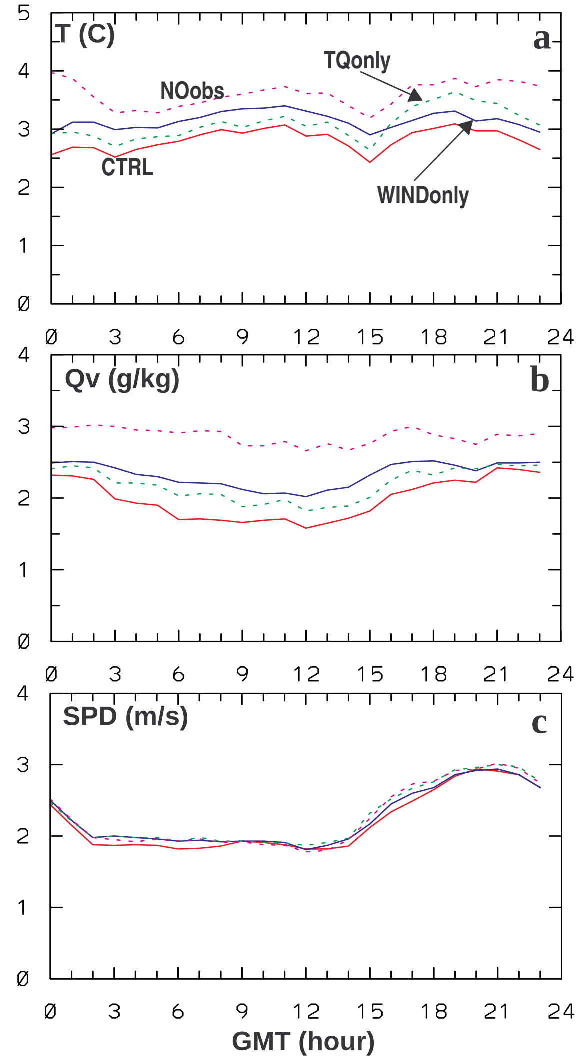

Fig.6 Same as Fig.3, but for CRTL, TQonly, WINDonly and NOobs.

TQonly and WINDonly are compared with each other and with the CTRL and NOobs runs. The TQonly and WINDonly experiments made use of both surface and upper-air data sources.

The RMSEs of the 12th hour upper-air forecasts from the four experiments are shown in Fig. 5. It is remarkable that assimilating wind only (WINDonly) not only corrects the upper-air wind error to the magnitude of the CTRL run, but it also reduces the upper-air temperature error in a similar extent to that in the TQonly experiment where temperature is assimilated directly. Even more interesting, WINDonly produces a moisture correction in the deep troposphere above 900 hPa, with values very close to the CTRL runs where all data are used. The moisture correction in the layer is even much larger than in the TQonly runs where moisture observations are directly assimilated. Below 900 hPa, the effect of WINDonly assimilation on moisture quickly diminishes toward the surface. Unlike the WINDonly runs, assimilating temperature and moisture only (TAonly) corrects the model thermal and moisture structure to a certain extent, but their impact on the model wind fields is negligible.

Surface errors (Fig. 6) of the 10th - 12th hour forecasts exhibit similar effects of assimilating temperature and moisture versus assimilating winds. Again, the WINDonly runs slightly improve the surface winds as well as the surface temperature and moisture properties, whereas the TQonly runs dramatically improve the surface temperature and moisture, but there is little effect on the surface winds. Unlike in the upper-air, the TQonly runs lead to much larger corrections of the surface temperature and moisture than those seen in the WINDonly runs.

It is worth pointing out that, aside from the 12-hourly radiosonde observations, the upper-air observations in the experiment are mostly dominated by wind observations (from wind profilers and satellite winds). Thus, the different effects of assimilating temperature, moisture and winds individually, seen in the current study, may result from the different data densities (data-set size). Nevertheless, it is interesting to point out that the larger effect of assimilating winds, relative to temperature, with our station-nudging approach agrees with many earlier studies that used the grid-nudging method.

5. SUMMARY AND CONCLUSIONS

The station-nudging approach of the NCAR/ATEC Real-time FDDA and forecast system (RTFDDA) provides a flexible method for incorporating all observations taken at irregular times and spatial locations. New unconventional data sources were introduced to the system, and many model features were refined to improve the RTFDDA analyses and forecasts. A set of eight comparison experiments were conducted to investigate the RTFDDA model response to the assimilation of the observations from satellite wind measurements and wind-profiler networks, from upper-air and surface platforms, and from temperature and moisture observations and wind observations. It can be concluded that, 1) Adding new data sources will always reduce the error, though to varying extents, no matter which variable is provided; 2) observations that are well distributed temporally and spatially appear to perform better than those clustered in time and space; 3) upper-air wind observations are most important and effective in driving the model toward the correct evolution (This indicates that data sources such as NESDIS satellite winds, FSL wind profilers and the FSL ACARS aircraft observations are especially valuable); and 4) surface observations (of temperature, moisture and winds) are critical to maintaining the proper evolution and accuracy of the model surface states and the physical structure of the lower troposphere; and lastly and interestingly, the simulations that employ the complete set of observations tend to produce the best analyses and forecasts.

Finally, we noted that, although the above conclusions appear to be robust for this case, it is important to investigate how the value of the each platform changes with synoptic weather and geographic locations by conducting more numerical experiments. The second group experiment mentioned earlier is one effort aiming at this. Also more complete data sources were tested in the second set. The result from this set of experiments will be reported at the conference and future publications.

8. ACKNOWLEDGMENTS

The authors would like to thank Wayne Bresky and Jaime Daniels for providing with the GOES satellite winds, Bill Moninger with ACARS data, Patty Miller and Michael Barth with MADIS data, Vicki Peck and Gene Legg with NASA QuikScat data, and Mark Anderson and his team with TAMDAR data.

REFERENCES

Bowers, J., S. Swerdlin, T. Warner, Y. Liu, L. Carson, A. Bourgeois, H. Fisher, D. Hahn, S. Low-Nam, D. Rife, T. Sandblom, B. Sharman, R. Sheu, F. Vandenberghe and M. Xu, 2003: Enhancement to and new application of, the ATEC data assimilation and forecast system. BACIMO-2003 conf., Monterey, CA.

Cram, J.M., Y. Liu, S. Low-Nam, R-S. Sheu, L. Carson, C.A. Davis, T. Warner, J.F. Bowers, 2001: An operational mesoscale RT-FDDA analysis and forecasting system. *Preprints 18th WAF and 14th NWP Confs*., Ft. Lauderdale, AMS, Boston, MA.

Liu, Y., S. Low-Nam, R. Sheu, L. Carson, C. Davis, T. Warner, S. Swerdlin, J. Bowers, M. Xu, H.-M. Hsu and D. Rife, 2002: Performance and enhancement of the NCAR/ATEC mesoscale FDDA and forecast system. *15th Conference on Numerical Weather Prediction* , San Antonio, Taxes, 399-402. 12-16 August, 2002.

Stauffer, D.R., and N. L. Seaman, 1994: Multiscale four-dimensional data assimilation. *J. Appl. Meteor., 33, 416-434.*