8.5 PRELIMINARY VERIFICATION OF THE NCAR-AIRDAT OPERATIONAL RTFDDA-WRF SYSTEM

Meredith Croke[1], Neil Jacobs[1], Peter Childs[1], Allan Huffman[1],
Yubao Liu[2], Yuewei Liu[2], and Rong-Shyang Sheu[2]

[1]AirDat LLC, Morrisville, NC 27560
[2]National Center for Atmospheric Research, Boulder, CO 80307

1. INTRODUCTION

AirDat began operationally running an explicit, CONUS-Scale version of the Advanced Research WRF (ARW) known as the NCAR-AirDat RTFDDA-WRF during the summer of 2009. The system is built upon the WRF model framework, but uses a Newtonian relaxation observational nudging data assimilation engine, which allows the model to more effectively assimilate asynoptic measurements such as those collected by the Tropospheric Airborne Meteorological Data Reporting (TAMDAR) sensor.

The TAMDAR sensor measures humidity, pressure, temperature, winds aloft, icing, and turbulence, along with the corresponding location, time, and altitude from built-in GPS. These observations are transmitted in real time to a ground-based network operations center via a global satellite network. The TAMDAR temperature, winds and humidity reports are continuously assimilated into the NCAR-AirDat RTFDDA-WRF modeling system. The operational modeling system cold-starts once a week, and produces 4 forecast cycles a day with each cycle producing a 6 h analysis and 72 h forecast from the dynamically consistent and cloud "spun-up" analysis produced by 4D continuous data assimilation.

The initial verification of the NCAR-AirDat RTFDDA-WRF operational forecasts was conducted by interpolating the model output to WMO/GTS standard radiosonde and METAR station points and computing statistical error metrics including BIAS, MAE and RMSE. The objectives of this study are to (i) Identify the impacts that TAMDAR data may have on the WRF-ARW forecast system, (ii) quantify any gains in forecast skills provided by more adequately assimilating asynpotic observations, and (iii) monitor the accuracy, contribution, and health of the TAMDAR QA system. Preliminary verification results suggest that the proper assimilation of the TAMDAR data improves the analyses and forecasts of the NCAR-AirDat operational 4 km CONUS RTFDDA-WRF system.

2. MODEL BACKGROUND

Since early 2009, AirDat and NCAR have worked together to implement a version of RTFDDA-WRF, which is an "observation-nudging" FDDA–based method built around the WRF-ARW core. This system is able to assimilate synoptic and asynoptic observational data sets, including various surface data (e.g., METAR, SYNOP, SPECI, ship, buoy, QuikScat seawinds, mesonets, etc.), and various upper-air observations (e.g., TEMP, PILOT, wind profilers, aircrafts (TAMDAR), satellite winds, dropsondes, radiometer profilers, RAOBS, Doppler radar VAD winds, etc.).

Several recent improvements have been made to the observation nudging scheme, including the ability to assimilate multi-level upper-air observations using vertical coherency principles. Additional improvements have been made to the terrain-dependent nudging weight corrections, including a ray-searching scheme, which eliminates the influence of an observation to a model grid-point if the two sites are physically separated by a significant mountain ridge or a deep valley.

RTFDDA "observation-nudging" is built for multi-scale mesoscale data assimilation. The multi-scale features are represented by differing influence radii for different grids and employs a revised "double-scan" approach.

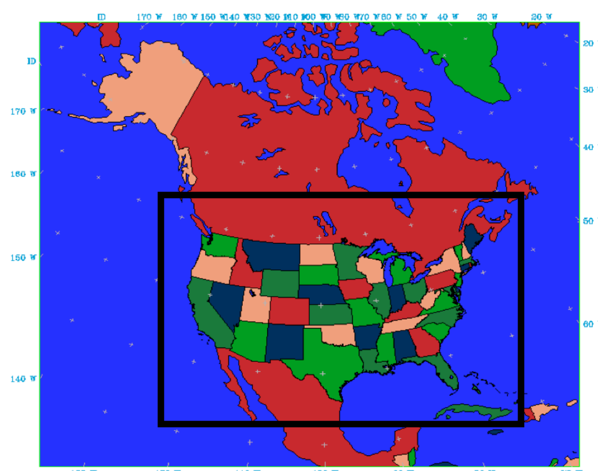

Fig. 1. The operational North America NCAR-AirDat RTFDDA-WRF grid configuration.

The outer domain (Fig. 1), which began cycling on 6-h intervals to 72 hours in late July 2009, features a grid spacing of 12 km with 74 vertical sigma levels, of which the highest concentration reside in the mixed layer, as well as near the jet stream level. The inner domain has a 4 km grid spacing that also has 74 vertical levels.

The initial configuration used the Lin microphysics scheme, the Kain-Fritch cumulus scheme (no CP for the 4 km), the YSU boundary layer parameterization, and the NOAH LSM. The radiation was handled by the

* *Corresponding author address*: Neil Jacobs, AirDat, LLC, 2400 Perimeter Park Dr. Suite 100, Morrisville, NC 27560. Email: njacobs@airdat.com

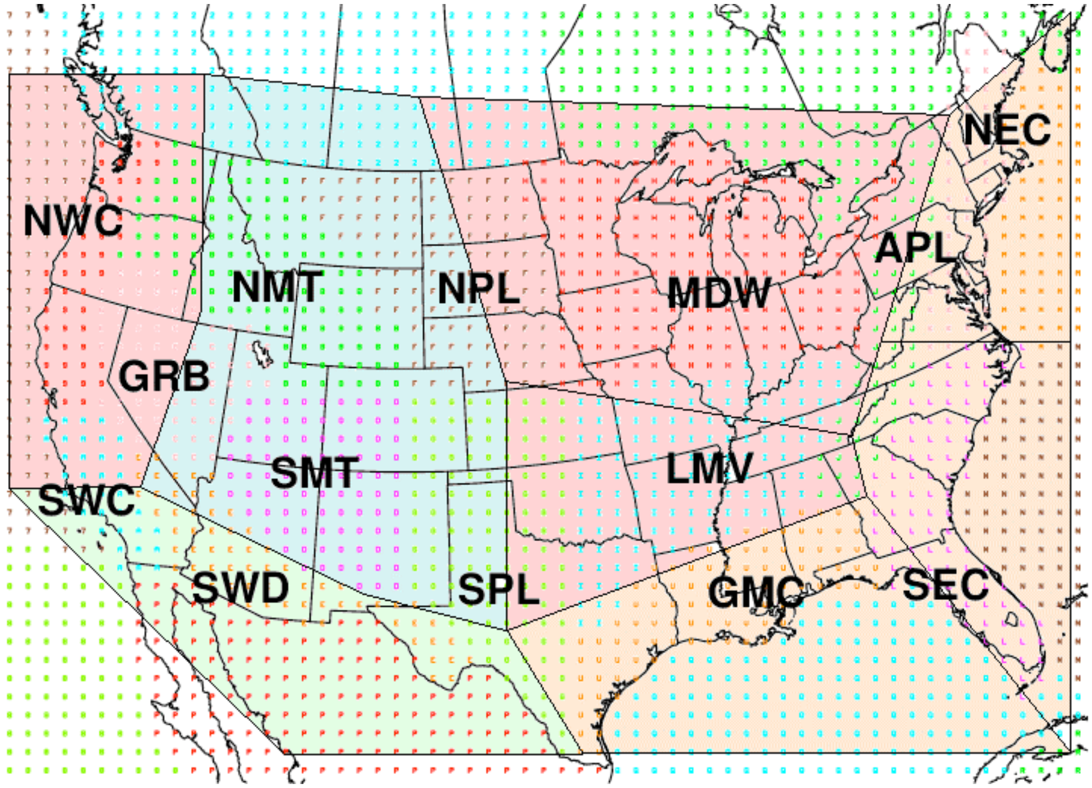

Fig. 2. The model verification grid containing the 14 locations (colored dots) defined by NCEP, as well as subdomains based on the volume of TAMDAR observations (red/orange having the most, and blue having the least as of 2009).

RRTM (longwave) and Dudhia (shortwave). The general performance of the configuration was very good, but there were several adjustments and upgrades that have been (and will be) implemented over the past (next) 6 months. Several modifications have been made to the system over the duration of this initial analysis (Childs et al. 2010).

Height-based data assimilation code was added to replace the conversion of the height levels to pressure levels, which allows the model to assimilate the observations directly on the native observation levels. This upgrade will have the greatest impact on wind profiler data assimilation, and should produce a slight improvement in the low-level wind forecasts.

The sea surface temperature (SST) initialization data was upgraded from the NCEP standard 40 km grid to the latest high resolution 9 km RTG SST composite product from the Polar Center at NCEP. In late August 2009, the WRF-ARW executables were upgraded to latest release version from NCAR (Version 3.1.1).

The final adjustment for (December) 2009 was to change the microphysics scheme from Lin to the Morrison two-moment scheme, which predicts the mixing ratios of rain, ice, snow and graupel, as well as the number concentration of these hydrometeor species. Future upgrades and modifications to the existing system are discussed in Childs et al. (2010).

3. VERIFICATION TOOL DEVELOPMENT

AirDat has recently set up an operational system for forecast model verification that employs the Model Evaluation Tools version 2.0 (METv2) to calculate MAE, bias, and RMSE for every forecast cycle from 0 - 72 hours every 6 hours. The input file to the statistical verification tool is a gridded GRIB1 standard destaggered pressure-level gridded output from the WRF post-processor. The point observation file is in the PrepBufr format. This file is reformatted using the PB2NC tool, which stratifies the observations according to the configuration file and writes them out in NetCDF format. This enables the Point-Stat program to compare the NetCDF observation file to the GRIB model forecast. The verification statistics are calculated via the Point-Stat tool by matching the gridded forecast output to the location of the observation. There are several interpolation options available, and they are currently being tested with the RTFDDA-WRF output.

Verification of surface observations is less complex since no vertical interpolation is required. However, upper-air verification is highly dependent on the method of vertical interpolation. AirDat will begin testing various methods of verification in 2010 to add an additional layer to the TAMDAR quality control, as well as model skill assessment. Since TAMDAR observations are both asynpotic, and typically recorded on non-standard pressure levels, verification will rely heavily on the ability to properly interpolate the output to the most sensible space-time position.

The MAE, bias, and RMSE are calculated for all 14 regions of the NCEP verification grid, as well as CONUS. Within the gridded verification domain (Fig. 2), there are groupings of subsections binned based on the abundance of TAMDAR observations. The red zones have the most flights, while the blue zones have the least. It is important to note that the zone is relevant when considering the skill for a particular forecast duration, since the influence of the observation will propagate downstream. It is also worth mentioning that many of these areas are inherently tough for the models to handle regardless of data (e.g., GRB, NMT, etc.). The opposite can also be true (e.g., SWC, GMC), and the level of difficulty is frequently linked to complex terrain, or lack thereof.

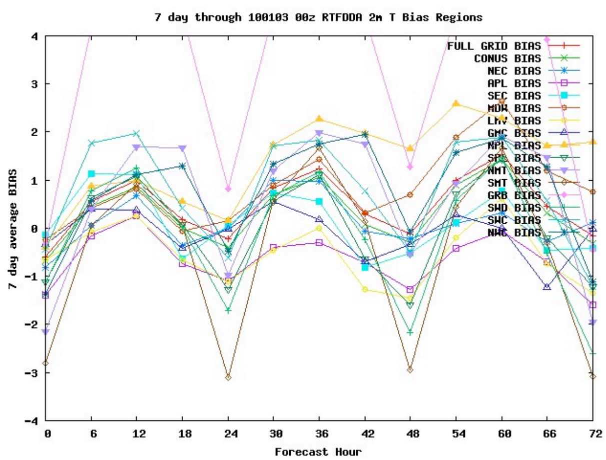

Fig. 3. An example of the 7 day running mean of 2-m temperature bias for the 14 regions and CONUS based on the 00Z cycle out to forecast hour 72.

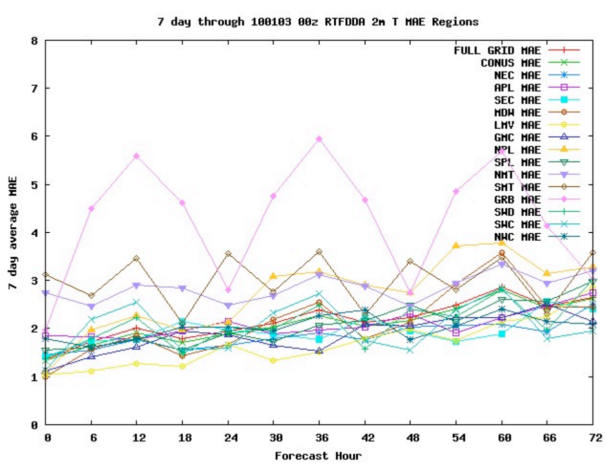

Fig. 4. An example of the 7 day running mean of 2-m temperature MAE for the 14 regions and CONUS based on the 00Z cycle out to forecast hour 72.

4. SUMMARY OF PRELIMINARY TESTING

An example of the 7-day running mean 2-m temperature bias tend can be seen in Fig. 3. There is a very slight cool bias with the analysis, but overall, the trend is a small upward warming near 0 to 0.5 C. Most evident is the diurnal error curve. There is typically a cool bias in the afternoon (around 18Z) and a warm bias in the morning (around 12Z).

The most extreme swings in the diurnal cycle 2-m temperature bias, and also (typically) the largest MAE (Fig. 4) and RMSE, are the regions GRB, SMT, NMT. The regions which consistently have the lowest bias (and MAE, RMSE) are LMV, SWC, SEC, GMC, MDW. Roux et al. (2009) have performed a similar analysis, and reported that errors, which appear to be highly dependent on the diurnal cycle, occur in higher altitude station locations. The long-term trends for the CONUS 2-m temperature MAE begin around 1.0 C and grow to approximately 2.3 C by forecast hour 72. Research is underway to determine whether these biases are model-related, or possibly linked to urbanization in the vicinity of the sites.

The 10-m winds within the CONUS verification region have consistently produced a positive bias of about 0.5 m s-1. The MAE over CONUS begins around 1.2 m s-1, but does not grow much with time, and by forecast hour 72, it is just below 2 m s-1.

The SEC region has the highest bias, as well as the strongest diurnal signal. Since the bias is present during the night, and nearly disappears completely during the day, it is possible that there are issues linked to differential heating consistent with the 2-m temperature bias of SEC, and the subsequent damping of the sea breeze over night since the SEC is bound on the east (and Gulf side) by warm coastal waters in the fall. NPL is a clear outlier, as it is the only region to produce a negative 10-m wind bias. As mentioned earlier, GRB, NMT, and SMT show the largest errors, while MDW, LMV, GMC have the lowest errors.

In 2006, Yubao et al. (2007) conducted an OSSE study using simulated TAMDAR observations to determine where the most improvement in forecast skill could be achieved based on existing, but yet-to-be-equipped flight routes. Figure 5a shows the simulation with no observation assimilation (control). The areas of green/black and pink/orange are regions with the largest errors. Figure 5b shows the same simulation with the inclusion of the simulated TAMDAR observations from the fleets that were to be equipped between 2006-09. As can be seen by comparing Fig. 5a to 5b, the majority of the existing error in the control simulation was significantly reduced with the exception of the region centered over UT, CO, and WY (white circle). Three years later, at the start of 2010, those proposed airline fleets have been equipped with TAMDAR, and are reporting data. Although this is just one event case that the OSSE was based on, it is consistent with the potential magnitude of error now being observed in GRB, NMT, and SMT from the same area as highlighted by the white circle (cf., Fig. 5b and Fig. 2).

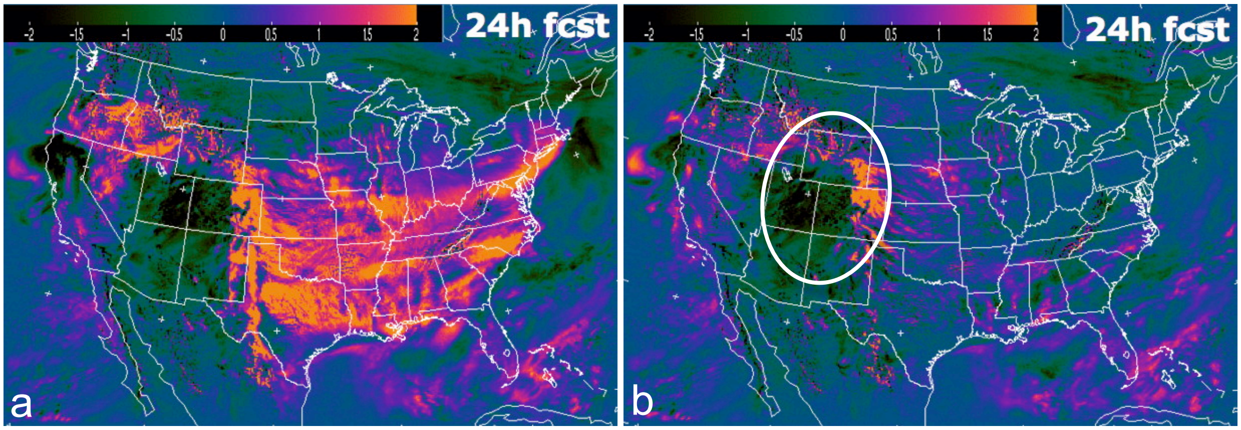

Fig. 5. Temperature errors in a 24h simulated OSSE study for the control (a), which had no observation assimilation and (b), which included simulated flight track data from all of the proposed airline fleets expected to be equipped by the end of 2009. Blue/green is no error. The error increases from purple to orange and from green to black.

In general, the errors and biases between the different cycles are lowest for the 00Z cycle, and highest for the 06Z. The 18Z and 12Z typically fall in the middle with the 12Z cycles showing slightly more skill. It is assumed that the 00Z and the 12Z have a clear advantage because of the inclusion of RAOB data. While the 18Z cycle has a significant portion of TAMDAR observations, the 00Z cycle contains the greatest impacts since some of the previous gains from the 18Z assimilation are retained.

5. ACKNOWLEDGMENTS

The authors are very grateful for the technical and computer support provided by NCAR. We would like to acknowledge the NASA Aeronautics Research Office's Aviation Safety Program, the FAA Aviation Weather Research Program, and AirDat, LLC for their support in this effort. Model Evaluation Tools (MET) was developed at the National Center for Atmospheric Research (NCAR) through a grant from the United States Air Force Weather Agency (AFWA). NCAR is sponsored by the United States National Science Foundation.

6. REFERENCES

Childs, P., N. Jacobs, M. Croke, A. Huffman, Y. Liu, W. Wu, G. Roux, and M. Ge, 2010: An introduction to the NCAR-AirDat operational TAMDAR-enhanced RTFDDA-WRF-ARW. 14th Symposium on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface (IOAS-AOLS), AMS, Atlanta, GA.

Liu, Y., N. A. Jacobs, W. Yu, T. T. Warner, S. P. Swerdlin, and M. Anderson, 2007: An OSSE study of TAMDAR data impact on mesoscale data assimilation and prediction. 11th Symposium on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface (IOAS-AOLS), AMS, San Antonio, TX.

Roux, G., Y. Liu, L. D. Monache, R.-S. Sheu, and T. T. Warner, 2009: Verification of high resolution WRF-RTFDDA surface forecasts over mountains and plains. 10th WRF users' workshop, Jun 20-23, 2009. Boulder, CO.