

7.1 THE DEVELOPMENTAL TESTBED CENTER WINTER FORECASTING EXPERIMENT

Ligia R. Bernardet^{1+2&}, Louisa Nance²⁺³, Hui-Ya Chuang⁴, Andrew Loughe¹⁺²⁺⁵, Meral Demirtas²⁺³,
Steven Koch¹⁺² and Robert Gall²⁺³

¹NOAA Research – Forecast Systems Laboratory, Boulder, CO

²Affiliated with the Developmental Testbed Center, Boulder, CO

³National Center for Atmospheric Research, Boulder, CO

⁴National Centers for Environmental Prediction, Camp Springs, MD

⁵Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO

[&]Also cocontract with Systems Research Group, Inc., (SRG), Colorado Springs, CO

1. INTRODUCTION

The Developmental Testbed Center (DTC – Nance et al. 2005) was established so the Numerical Weather Prediction (NWP) research and operational communities can interact to accelerate testing and evaluation of new models and techniques for research applications and operational implementation. One method of testing and evaluation employed by the DTC is real-time forecast experiments, such as the DTC Winter Forecast Experiment (DWFE), which ran from January to March 2005.

DWFE was designed in close consultation with operational forecasters, researchers, and DTC staff, and was motivated by the needs of the National Weather Service (NWS) for improved model guidance to support their winter weather forecast and warning mission. The objectives of DWFE were to:

- provide experimental model guidance for winter weather forecasting over a large domain using two variants of the Weather Research Forecast (WRF) model with explicit convection only (no convective parameterization scheme),
- expose forecasters to the nature and behavior of the WRF model at very high resolution prior to the first scheduled operational implementation at NCEP,
- determine whether encouraging results seen earlier from 4-km WRF runs in the warm season provide forecast value during the winter for lead times out to 48 h, and
- identify the forecast value of small-scale flow features resolved by the high-resolution grids.

2. EXPERIMENT SETUP

DWFE employed an end-to-end forecast system with the following components: data preprocessor, forecast model, postprocessor, product dissemination, forecast verification, and archival.

2.1 Data ingest and forecast models

For this experiment, forecasts over the CONUS were generated using two dynamic cores within the

WRF framework: the Nonhydrostatic Mesoscale Model (NMM) developed by NCEP (Janjic 2003) and the Advanced Research WRF (ARW) developed by NCAR (Skamarock et al. 2005). More information on WRF can be obtained at <http://www.wrf-model.org/>. The setup of each dynamical core, which was run with its own set of physics, is shown in Table 1. The characteristics of the operational Eta model (Black 1994), run operationally at NCEP, are also summarized in Table 1 for comparison. The WRF-NMM core was run on NOAA-FSL's supercomputer, which is a Pentium IV Linux cluster. The model was integrated on 529 Xeon 2.2 GHz processors with 1 GB memory each and the model-only processing time was typically 5 h, 40 min. The WRF-ARW core was run on NCAR's IBM SP-cluster system. The model was integrated on 256 Power4 1.3 GHz processors with 2 GB memory each. The model-only processing time was typically 4 h, 45 min.

Both cores were run daily for the 0000 UTC cycle. The initial and boundary conditions for the 48-h forecasts were based on Eta 212 data (40-km grid spacing) processed using the WRF Standard Initialization (SI). In addition to the Eta 212 data, the initial land surface fields for the ARW core were obtained from the High-Resolution Land Data Assimilation System (HRLDAS; Chen et al. 2004). HRLDAS utilizes observations and land characteristics to drive the NOAA Land Surface Model (LSM) in uncoupled mode to capture fine-scale heterogeneity in land state. By executing on the WRF grid, HRLDAS-assimilated land state variables can be ingested directly into the coupled WRF/NOAH LSM forecast system without interpolation.

As detailed in Table 1, the dynamical cores use different grid projections, so even though the grids for both cores were set up to cover a virtually identical area, a postprocessing system was necessary to bring the forecasts onto a common grid.

2.2 Postprocessing

The WRF-POST (Chuang et al. 2004), developed at the National Centers for Environmental Prediction (NCEP), was used to postprocess DWFE forecasts. It serves three main purposes:

- vertically interpolate the forecasts from their native grid to isobaric levels and shelter level,

- horizontally interpolate the forecasts from their native staggered grid to an unstaggered grid defined by the user, and
- compute derived meteorological fields.

During DWFE, the WRF-NMM and WRF-ARW forecasts were both horizontally interpolated to a common Lambert Conformal Grid (G163), with 5 km grid spacing covering an area very similar to the native grids of each model. Forecasts on G163 were used for dissemination to users and for verification.

The pressure reduction to sea level was computed in the WRF-POST using the membrane method described in Chuang et al. (2004), while the reductions in temperature, RH, and wind to shelter level were computed by the WRF model itself and are compatible with the physics used in each setup. To meet the needs of the forecasters, new products, such as visibility and precipitation type, were incorporated into the WRF-POST. One of the most popular forecast products was simulated composite reflectivity, which is included here (Fig. 1) to illustrate the domain of G163.

The last step in post-processing was the conversion of the output of the WRF-POST from its native Grib-1 format to Grib-2. Since Grib-2 is significantly more compact, this format is superior for data transfers.

2.3 Dissemination

Realtime displays for the forecasts were available in four ways:

- Images for a select number of fields and levels were generated with NCAR Command Language (NCL) and made available through the DTC website (<http://www.DTCenter.org>). These images, as well as corresponding radar composites and Eta forecasts, are available via the DWFE catalogue hosted by the Joint Office for Science Support (JOSS; <http://www.joss.ucar.edu/dwfe/catalog/>).
- User-specified images were made available through FX-Net, an AWIPS-like interface that allows the user to interrogate the full three-dimensional model grids. FX-Net uses a client-server protocol to generate images on demand and to efficiently transfer the images to the users' display using a wavelet compression technique.
- The WRF-NMM forecasts for a select subset of two-dimensional fields (primarily surface and precipitation fields) were available through AWIPS in the Central, Eastern and Southern region National Weather Service (NWS) forecast offices. This limited AWIPS distribution was due to severe bandwidth constraints in the local forecast offices.
- Three-dimensional postprocessed grib files were available through scp to registered users, such as the NWS Hydrometeorological Prediction Center (HPC).

Additionally, the forecasts on their native grids and the post-processed grids were transferred to the NCAR Mass Store System (MSS) for archival and

future use.

2.4 Verification

Forecasts were evaluated through subjective and objective verification. Subjective evaluation was collected from the operational and research communities via online forms accessible through the DWFE catalogue. The responses to the forms are available via the DWFE catalogue and are summarized in Koch et. al (2005).

Objective verification of surface and upper air fields was performed for 74 of the 76 cycles of DWFE using the NCEP Verification System (Chuang et al. 2004). This package employs a grid-to-point verification approach in which forecast fields are bilinearly interpolated to station location.

The NCEP Verification System generates a Verification Statistics Data Base (VSDB) file, which contains the raw numbers from which the final statistics are computed. In the DTC, the VSDB files are ingested into a MySQL database and can be queried and visualized through a web interface (http://www-ad.fsl.noaa.gov/users/loughe/projects/wrf/DWFE_4/). This approach facilitates the sharing of information among the DTC members, who are distributed through a variety of locations.

At the surface, forecasts of mean sea level pressure (MSLP), 2-m temperature, 2-m relative humidity (RH), and 10-m winds were compared against METAR observations. The verification of 2-m temperature takes into consideration differences between model terrain height and observation height, and computes an adjustment to the temperature based on any discrepancy.

For upper levels, forecasts of temperature, RH, and winds were compared against rawinsondes.

Objective verification of QPF, as discussed by Demirtas et al. (2005a), uses two approaches:

- The NCEP Precipitation Verification System (Chuang et al. 2004) uses a grid-to-grid approach in which the forecasts and a gridded precipitation analysis are interpolated to a common grid and compared. Stage II data are used to verify 3-h accumulations and the River Forecast Center's one-eighth degree analysis is used to verify 24-h accumulations.
- The Realtime Verification System (RTVS) developed at NOAA FSL (Loughe et al. 2001) uses a grid-to-point approach to compare forecasts interpolated to rain gauge locations against the Hydrometeorological Automated Data System.

The results of the DWFE QPF Verification are presented in Demirtas et al. (2005b).

3. OBJECTIVE VERIFICATION RESULTS

3.1 Wind

All models presented positive wind speed biases

at 10 m (Fig. 2). The Eta and the WRF-NMM had very similar values of approximately 0.3 m s^{-1} . The WRF-ARW had much larger biases that followed a diurnal cycle, with maxima of 1.4 m s^{-1} at night and minima of 0.7 m s^{-1} during the day

The RMSE increased from 2.6 m s^{-1} for the Eta and the WRF-NMM and 3.2 m s^{-1} for the WRF-ARW in the beginning of the forecast up to 3.4 m s^{-1} for the Eta and the WRF-NMM and 4.0 m s^{-1} for the WRF-ARW by the end of the forecast period.

The excessive wind speeds at the surface contrasted with the results at upper levels. All models had near zero bias at 850 hPa with the bias becoming more negative up to 250 hPa (Fig. 3 – top), where the values reached -1.3 m s^{-1} (WRF-ARW), -1.1 m s^{-1} (WRF-NMM) and -0.8 m s^{-1} (Eta). Further investigation should be made as to the cause of the differences between the bias characteristics at the surface and those for the rest of the tropospheric column. Excessive vertical mixing is a possible explanation for the presence of high winds at the surface and low winds at 850 hPa. However, the negative bias throughout the entire tropospheric column suggests the deficiency at 850 hPa can be attributed to more general problems rather than the planetary boundary layer. The consistency between METAR and rawinsonde wind observations should also be investigated, since an inconsistency between these observation types could be the source of the discrepancy.

The tropospheric profile of wind vector RMSE was remarkably similar for all models (Fig. 3 – bottom). The smallest errors occurred at 850 hPa (approximately 4.7 m s^{-1}), while the highest RMSE occurred at jet level (approximately 8.6 m s^{-1}). This behavior is typically found in model verification studies, which show forecast wind errors vary in proportion to the wind speed.

3.2 Temperature and Relative Humidity

The 2-m temperature bias for all models presented a diurnal cycle with the Eta model exhibiting the smallest bias (Fig. 4 – top). The Eta and the WRF-NMM had the same diurnal trend of bias: the NMM particularly tended to be cold at night and warm during the day, revealing an exaggerated diurnal cycle. The WRF-NMM positive temperature biases during the day reached 1.4°C on the first day and 1.7°C on the second, while the negative biases were -1.4°C on the first day and near zero on the second. The resemblance of the Eta and WRF-NMM model verification statistics is likely due to similarities in their physics packages.

A bug in the radiation parameterization for the WRF-NMM was discovered during DWFE that caused both the shortwave and the longwave radiation not to interact with ice clouds. As a consequence, excessive solar radiation arrived at the surface during the day, and there was a deficit in longwave radiation loss to space at night. Certainly this bug affected the temperature biases seen during DWFE. Following DWFE, NOAA/FSL started the NMM5-CONUS real-time experiment using a new version of the WRF-

NMM code. The radiation bug was one of several updates and bug fixes that were made for NMM5-CONUS but otherwise the configuration was identical to that of the WRF-NMM model during DWFE. The results of NMM5-CONUS forecast verification for April through July (Fig. 5 – top) show an improvement of the WRF-NMM temperature errors, especially during the daytime. Comparing DWFE with NMM5-CONUS, it is noticeable that the WRF-NMM daytime and nighttime biases were reduced, and that the WRF-NMM values were brought closer to the Eta values.

The diurnal cycle of 2-m temperature bias for the WRF-ARW during DWFE (Fig. 4 – top) was of the opposite sign to that of the WRF-NMM's: the WRF-ARW was too warm during the night and slightly too cold during the day, that is, it underestimated the diurnal cycle. The positive bias at night was approximately 2.3°C , and during the day, -0.7°C . It is well known that the near surface temperature fields can be very sensitive to the temperature and moisture content of the soil. Since the WRF-ARW and the WRF-NMM used different sources for LSM initialization, the 2-m temperature differences between the two setups could be due to three factors:

- Differences in land surface initialization;
- Factors intrinsic to each setup's dynamics and physics;
- Differences in reduction to shelter level.

The contributions of each of the factors listed above are being researched in the DTC.

All models presented a diurnal cycle of 2-m temperature RMSE (Fig. 4 – bottom) superimposed on a slight overall increasing trend with forecast time with maxima at 1200 UTC (sunrise). WRF-ARW had the highest RMSE, with 3.8°C on the first day and 4.1°C on the second. Eta had the lowest RMSE, with maxima of approximately 3.2°C on the first day and 3.4°C on the second. The WRF-NMM also had secondary maxima at 21 and 45 UTC, corresponding to its daytime positive bias. Quite different diurnal temperature forecast errors were found in the warm season NMM5-CONUS runs.

At upper levels, the Eta and the WRF-ARW had similar biases at most levels, whereas the WRF-NMM was generally warmer than the Eta and the WRF-ARW by approximately 0.4°C . While the sign of the temperature bias for the Eta and WRF-ARW varied with height, the WRF-NMM temperature bias was positive at all levels except 150 hPa, where it reached -0.8°C . The maximum WRF-NMM bias was 0.7°C at 250 hPa. It is interesting to note that the pattern of temperature bias changes little from the initial time (Fig. 6 - top) to later times in the forecast (Fig. 7 - top), indicating that errors present at the initial time persist during the forecast. The reason for the differences (0.4°C) between WRF-ARW and WRF-NMM temperature biases at initial time can be attributed to differences in the initialization procedure for each core. Both models start from the Eta 212 analysis; however, the WRF-ARW obtains temperature directly from the Eta 212 temperature, while the WRF-NMM retrieves temperature from the Eta 212 geopotential using the hydrostatic approximation.

The upper air temperature forecast RMSE was

similar in all models, with maxima at 850 (1.9°C) and 200 hPa (2.2°C) and a minimum at midlevels of 1.3°C (Fig. 7 - bottom). Although the RMSE increased with forecast time, the shape of the curve remained unaltered.

The bias of forecasted upper air RH shows a strong intermodel variability. This is possibly due to the difficulty in verifying this quantity, which has a high spatial variability. The WRF-ARW had positive bias at all levels, with values up to 2.1 % at 500 hPa (Fig. 8 – top). The WRF-NMM has a negative bias at 850 and 700 hPa (minimum of –1.1 %) and positive above, with a maximum of 0.6 at 300 hPa. The Eta model had positive bias at all levels except 850 hPa, with the magnitude of the bias not exceeding 2.5%.

All models exhibit RH RMSE minima of approximately 19.0% at 850 and 300 hPa and a maximum at 500 hPa of approximately 21.0% (Fig. 8 – bottom). The Eta has the lowest RMSE at all levels, while the WRF-NMM had the highest RMSE at low and upper levels and the WRF-ARW had the highest midlevel RMSE.

Unlike the behavior of temperature, the bias and RMSE of RH at 24 h was very different than its initial time counterpart (not shown). While at the initial time the largest RMSEs were found at the lower and upper levels, at 24 h they are found in midlevels.

4. DISCUSSION AND CONCLUSIONS

The central mission of the DTC is to bridge the gap between research and operations. This objective was certainly accomplished in DWFE, where forecasters, scientists, and DTC staff came together to design an experiment that would familiarize NWS forecasters with a new Numerical Weather Prediction model and supply forecasters with new high-resolution products permitting state-of-the-art diagnostic and prediction of wintertime mesoscale weather systems, such as narrow reflectivity bands, topographically forced weather, and lake effect snow bands. Koch et al. (2005) present a summary of the NWS participation in DWFE.

The objective forecast verification of DWFE presented here reveals that the setups of the WRF models performed comparably. The WRF-ARW model had more near surface errors, with a flat temperature diurnal cycle and excessive winds. The WRF-NMM and Eta temperature forecast errors were similar near the surface, especially after a radiation bug was fixed in the WRF-NMM. At upper levels, all models performed similarly in wind forecasts. For the temperature forecasts, the WRF-NMM was consistently half a degree warmer than the other models. While this small difference is important, it was traced to differences in the initialization method of the WRF-NMM and the WRF-ARW and does not reflect significant differences in performances between the dynamic core/physics package configurations. Relative humidity is the field with the largest differentiation among the models, with the WRF-NMM presenting more errors at lower levels and the WRF-

ARW more errors at midlevels.

The verification statistics computed during DWFE provide feedback to model developers and scientists about areas that need improvement. An area of special interest is to advance the understanding of the differences in surface temperature bias between the WRF-NMM and the WRF-ARW and to determine whether they are intrinsic to the model physics and dynamics or are caused by model initialization and/or postprocessing.

The results of DWFE also call for further advancement in the area of forecast verification. The use of RH alone does not provide a good understanding of the moisture errors, because RH is influenced by temperature. The use of specific humidity or dewpoint (that is actually reported by the observations) is desirable. Moreover, non-traditional verification based on spectral analysis, object-oriented approaches and others are needed to shed further light on the characteristics and value of the WRF model forecasts in its different setups. The DTC is currently fostering an active Visitor Program (Nance et al. 2005) to investigate these issues (Gallus 2005, Bernardet et al. 2005).

5. REFERENCES

- Bernardet, L. P. Bogenschutz, J. Snook and A. Loughe, 2005. WRF forecasts over the Southeast United States: Does a larger domain lead to better results? *6th WRF / 15th MM5 Users' Workshop*, Boulder, CO, Natl. Center for Atmos. Res., CD-ROM, P2.10.
- Black, T., 1994: The New NMC Mesoscale Eta model: description and forecast examples. *Wea. Forec.*, **9**, 265–278.
- Chen, F., K. W. Manning, D. N. Yates, M. A. Lemone, S. B. Trier, R. Cuena, D. Niyogi, 2004. Development of High-Resolution Land Data Assimilation System and its application to WRF. Preprint, *16th Conference on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc..
- Chuang, H., G. DiMego, M. Baldwin, and WRF DTC team, 2004. NCEP's WRF Post-Processor and Verification Systems. *5th WRF / 14th MM5 Users' Workshop*, Boulder, CO.
- Demirtas, M., L. Nance, L. Bernardet, Y. Lin, H. Chuang, A. Loughe, R. Gall, and S. Koch, 2005a. Verification Systems used for the Developmental Tested Center. *6th WRF / 15th MM5 Users' Workshop*, Boulder, CO, Natl. Center for Atmos. Res., CD-ROM, P3.23.
- Demirtas, M., L. Nance, L. Bernardet, Y. Lin, H. Chuang, A. Loughe, J. Mahoney, R. Gall, and S. Koch, 2005b. Quantitative precipitation forecast (QPF) verification of DWFE. *21st Conference on Weather Analysis and Forecasting*, Washington, D.C., Amer. Meteor.Soc..

Gallus, W., 2005. Comparison of WRF dynamic core, physics package, and initial conditions on warm season rainfall forecasts. *21st Conference on Weather Analysis and Forecasting*, Washington, D.C., Amer. Meteor. Soc..

Janjic, Z. I., 2003: A nonhydrostatic model based on a new approach. *Meteorol. Atmos. Phys.* **82**, 271–285.

Koch, S., R. Gall, G. DiMego, E. Szoke, J. Waldstreicher, P. Manousos, B. Meisner, N. Seaman, M. Jackson, R. Graham, A. Edman and D. Nietfeld, 2005. Lessons learned from the DTC Winter Forecast Experiment. *21st Conference on Weather Analysis and Forecasting*, Washington, D.C., Amer. Meteor. Soc..

Loughe, A., J. Henderson, J. Mahoney, E. Tollerud, 2001. A verification approach suitable for assessing

the quality of model-based precipitation forecasts during extreme precipitation events. Preprint, *Symposium on Precipitation Extremes: Prediction, Impacts and Responses*, Albuquerque, NM. Amer. Meteor. Soc.

Nance, L., L. Bernardet, H. Chuang, G. DiMego, M. Demirtas, R. Gall, S. Koch, Y. Lin, A. Loughe, J. Mahoney and M. Pyle, 2005. The WRF Developmental Testbed Center: A Status Report. *6th WRF / 15th MM5 Users' Workshop*, Boulder, CO, Natl. Center for Atmos. Res., CD-ROM, P2.8.

Skamarock, W.C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, J. G. Powers, 2005: A Description of the Advanced Research WRF Version 2. NCAR Tech Note, NCAR/TN-468+STR, 88 pp. [Available from UCAR Communications, P. O. Box 3000, Boulder, CO 80307].

Table 1. Setups of the WRF-ARW, WRF-NMM and Eta during DWFE.

	WRF-ARW	WRF-NMM	Eta
Land-Surface Model	NOAH 5 layer*	NOAH 5 layer*	NOAH 5 layer*
Boundary Layer	YSU	MYJ 2.5	MYJ 2.5
Microphysics	WSM5	Ferrier	Ferrier
Cumulus Parameterization	None	None	BMJ
Shortwave	Dudhia	Lacis-Hansen	Lacis-Hansen
Longwave	RRTM	Fels-Schwartzkopf	Fels-Schwartzkopf
Projection	Lambert Conformal	Rotated Lat-Lon	Rotated Lat-Lon
Grid Staggering	C	E	E
Vertical Coordinate	Terrain following sigma	Hybrid: terrain following sigma at low levels and isobaric above	Step mountain
Horizontal Grid Spacing	5 km	5 km	12 km
Number of verticals levels	37	37	37
Initial Conditions	Eta 212 + HRLDAS	Eta 212	EDAS
Boundary Conditions	Eta 212 – 3 hourly	Eta 212 – 3 hourly	GFS

*The version of the NOAH LSM is not identical in the WRF-ARW, WRF-NMM and Eta models.

DWFE ARW 03/28/2005 (00:00) 3 hr fcst

Valid 03/28/2005 03:00 UTC

Composite Reflectivity (dBZ)

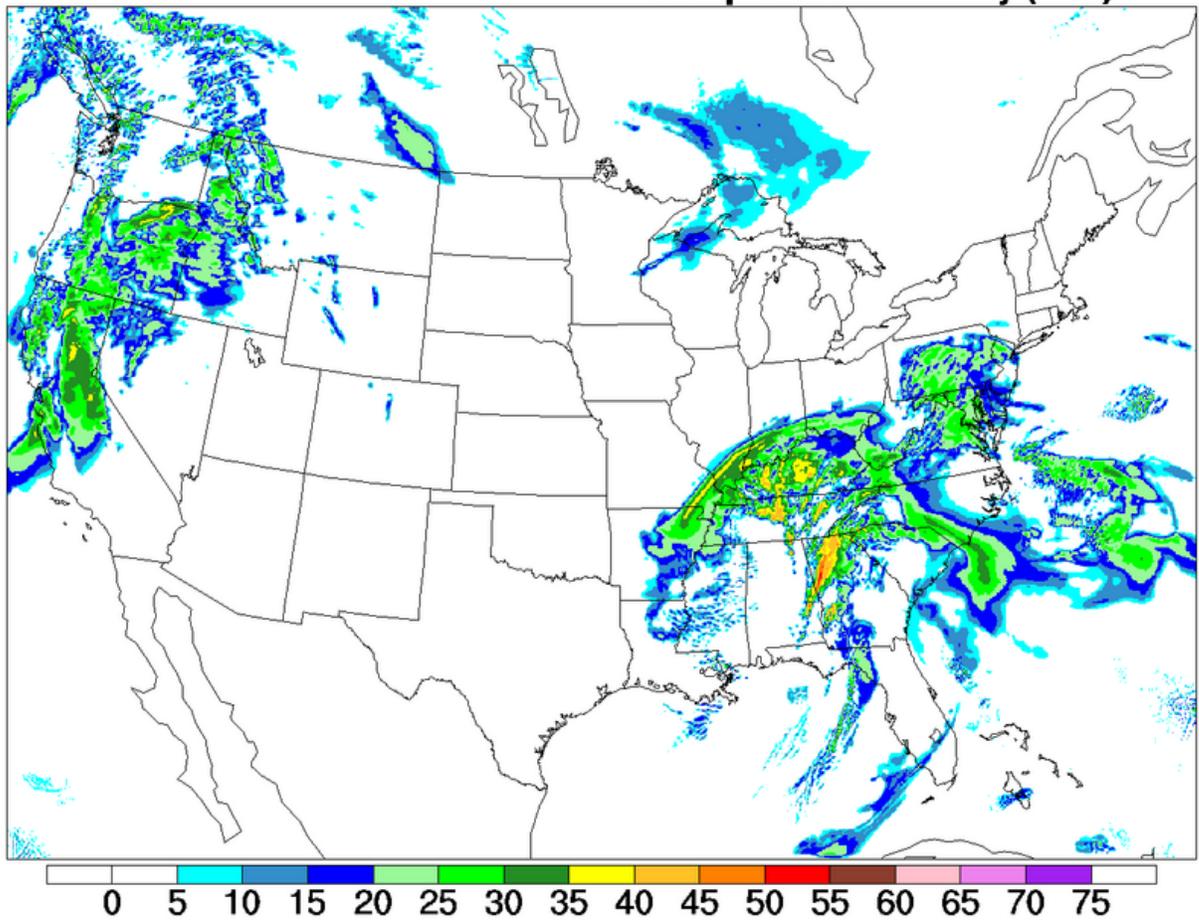


Fig. 1. Three-hour simulated composite reflectivity (dBZ) for the WRF-ARW model valid at 0300 UTC March 28, 2005.

SFC Vector Wind Lead Time (fcsthr):ALL 15 JAN - 31 MAR 2005 Natl Domain

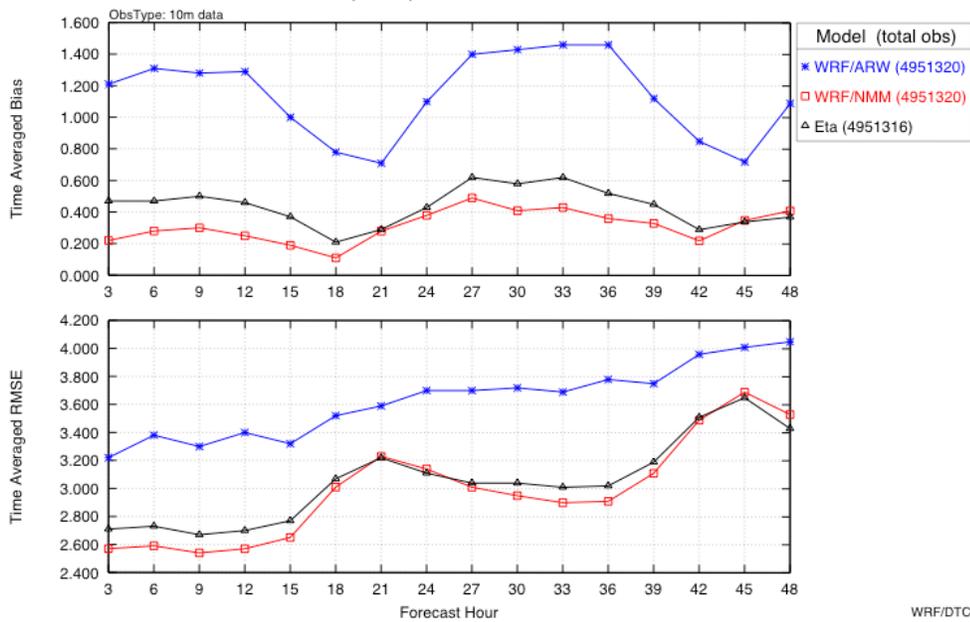


Fig. 2 Ten-meter wind bias (top) and RMSE (bottom) averaged for 74 days of DWFE. WRF-ARW in blue, WRF-NMM red, and Eta in black.

Vector Wind Lead Time (fcsthr):ALL 15 JAN - 31 MAR 2005 Natl Domain

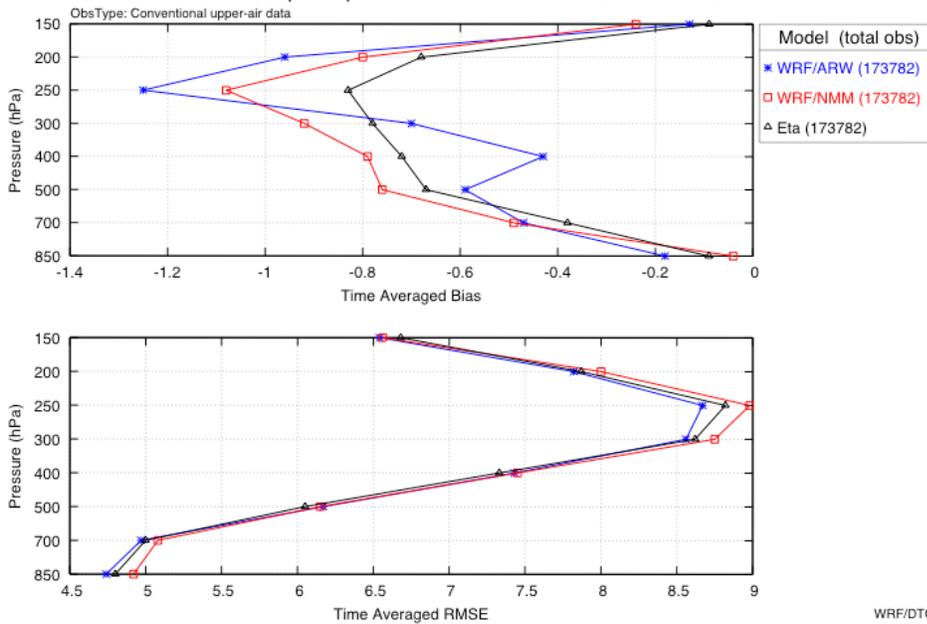


Fig. 3. Vertical profile of wind bias (top) and RMSE (bottom) averaged for all forecast hours over 74 days of DWFE. WRF-ARW in blue, WRF-NMM red, and Eta in black.

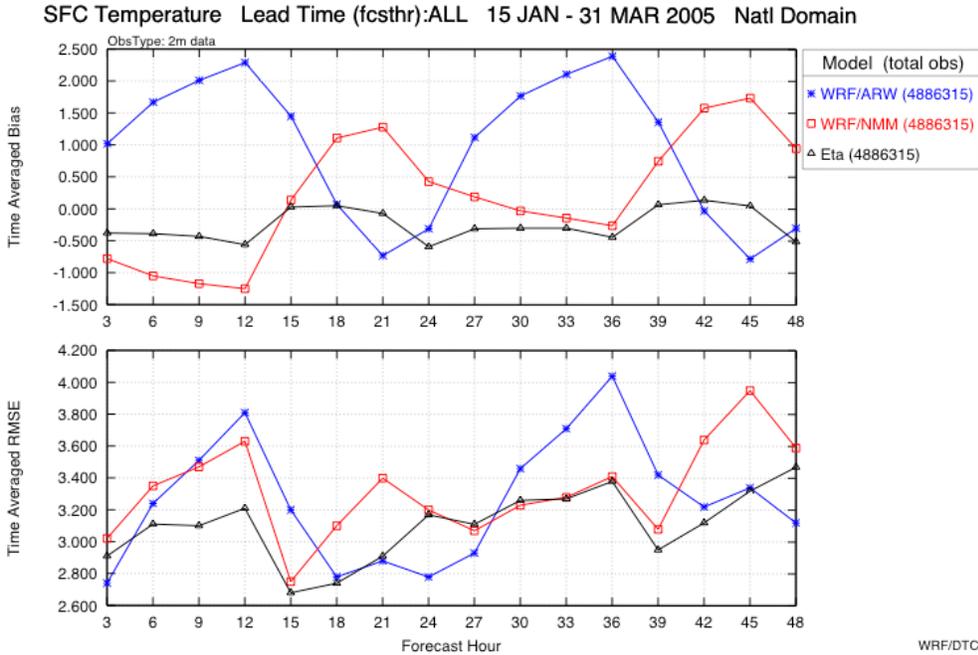


Fig. 4. Two-meter temperature bias (top) and RMSE (bottom) averaged for 74 days of DWFE. WRF-ARW in blue, WRF-NMM red, and Eta in black.

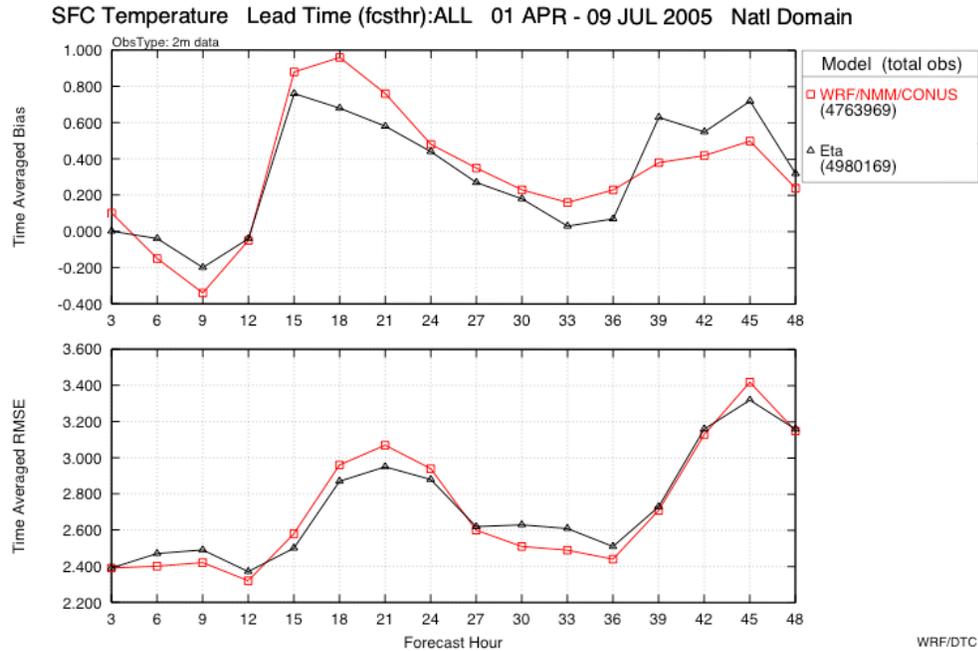


Fig. 5. Two-meter temperature bias (top) and RMSE (bottom) averaged for 69 days of the NMM5-CONUS Project. WRF-NMM red, and Eta in black.

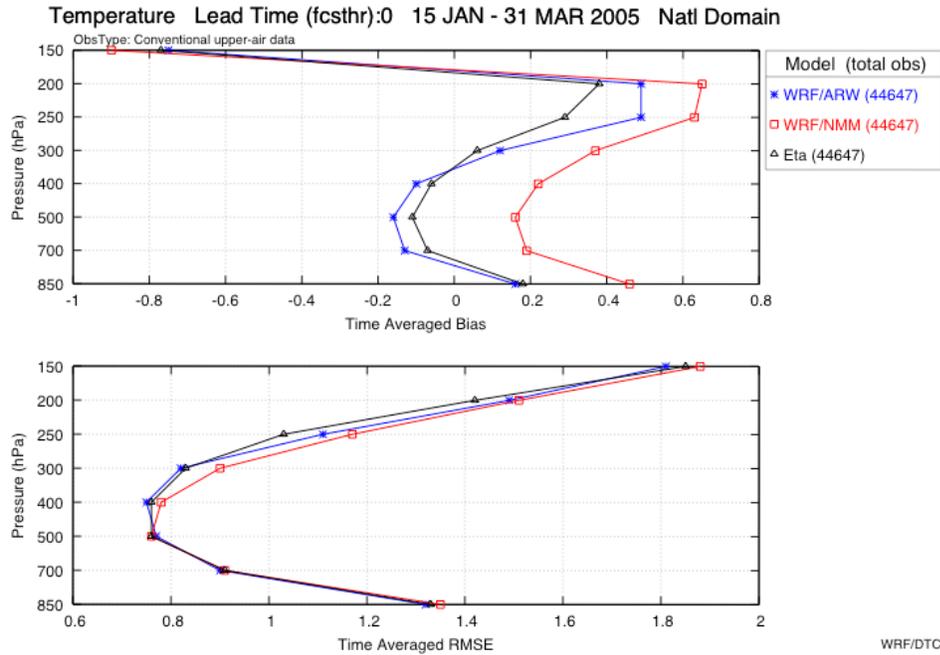


Fig. 6 Vertical profile of initial time temperature bias (top) and RMSE (bottom) averaged for 74 days of DWFE. WRF-ARW in blue, WRF-NMM red, and Eta in black.

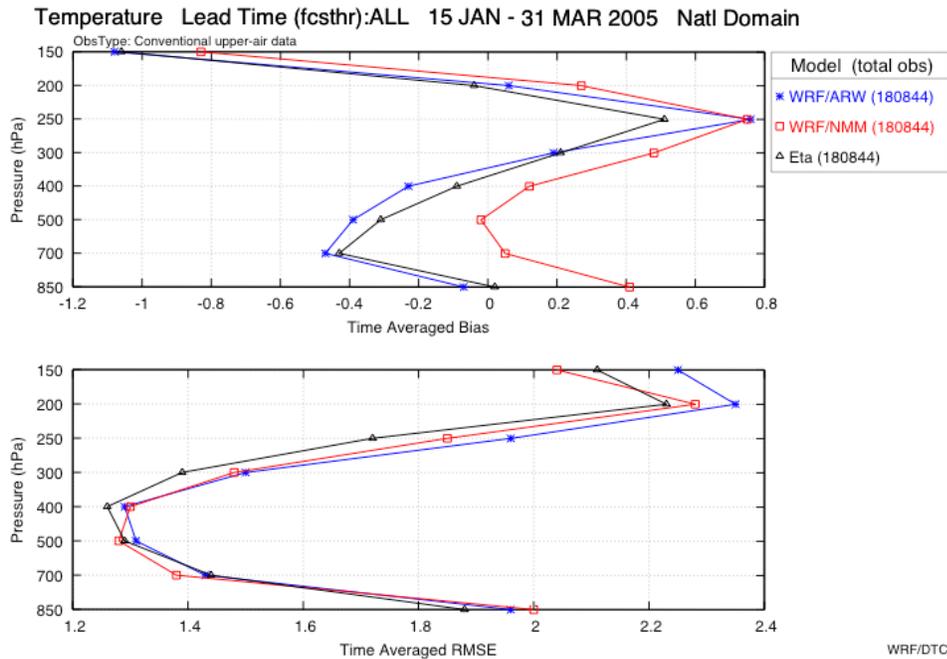


Fig. 7. Vertical profile of temperature bias (top) and RMSE (bottom) averaged for all forecast hours over 74 days of DWFE. WRF-ARW in blue, WRF-NMM red, and Eta in black.

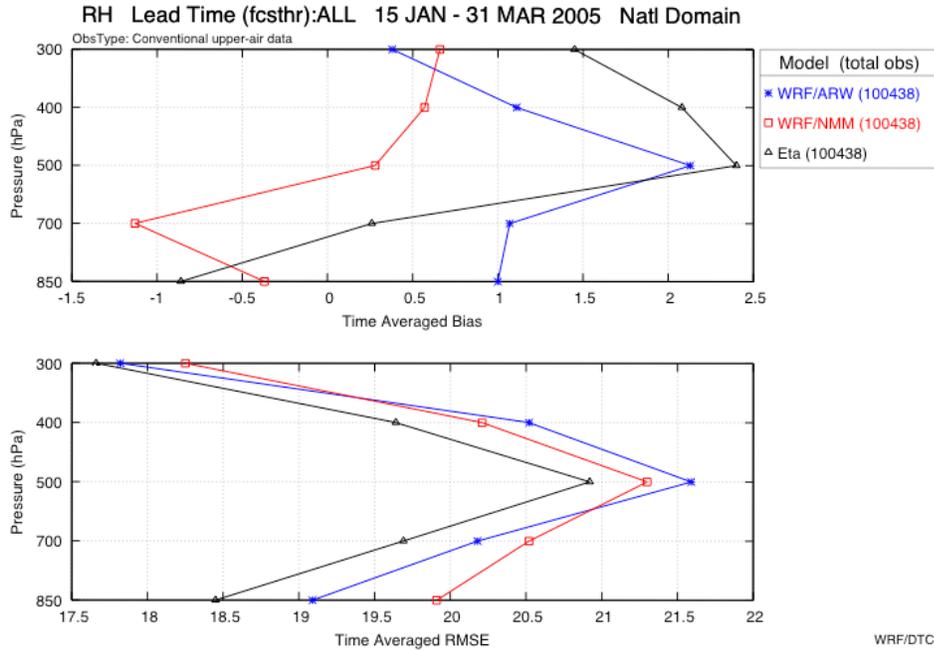


Fig 8. Vertical profile of RH bias (top) and RMSE (bottom) averaged for all forecast hours over 74 days of DWFE. WRF-ARW in blue, WRF-NMM red, and Eta in black.