9.3 THE EFFECTS OF VERTICAL RESOLUTION ON THE OPTIMIZATION OF TAMDAR DATA IN SHORT-RANGE MESOSCALE FORECASTS

Neil A. Jacobs[1*], Yubao Liu[2], and Cyrena-Marie Druse[1]

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

1. INTRODUCTION AND MOTIVATION

Upper-air observations are disproportionately sparse, both temporally and geographically, when compared to surface observations. The lack of data is likely one of the largest limiting factors in numerical weather prediction. Atmospheric measurements performed by the Tropospheric Airborne Meteorological Data Reporting (TAMDAR) sensors of humidity, pressure, temperature, winds, icing, and turbulence provide significant additional information in the lower troposphere. The meteorological data, along with the corresponding location, time, and altitude from built-in GPS, are relayed to a ground-based network operations center via satellite in real-time. The TAMDAR sensors were deployed on a fleet of 63 Saab 340s operated by Mesaba Airlines in the Great Lakes region as a part of the NASA-sponsored Great Lakes Fleet Experiment (GLFE) by 15 Jan. 2005. More than 800 soundings are generated from 400 flights to 75 regional airports daily.

A study of the impact of the TAMDAR data on mesoscale NWP is conducted using a mesoscale model which employs various assimilation techniques and the available TAMDAR data. Six parallel 12-h simulations, where the three experimental (control) runs include (withhold) TAMDAR data, are performed, with one of the three pair of experimental and control simulations having 36 sigma-levels, while the other two pair of experimental and control simulations have 48 sigma-levels. One of the 48-sigma-level pairs has a single domain of 10-km grid spacing, while the other two pair have a 36-km outer domain which feeds a 12-km inner domain.

The objectives of this study are to (i) optimize impacts that TAMDAR data may have on the forecast system by increasing the horizontal distribution of vertical atmospheric profiles during initialization, and (ii) to isolate the ability of the model to utilize a higher vertical resolution during summer convective forecasts.

2. METHODOLOGY AND MODEL CONFIGURATION

There are six different parallel model runs in this study (Table 1). The AirDat-standard run (AD) features an outer domain of 36-km grid spacing and a two-way nested 12-km inner domain. The AD run has 36 sigma-levels and assimilates the TAMDAR data. The AirNot-standard run (AN) has an identical model configuration except AN does not include TAMDAR data. The AirDat-2 (AD2) and AirNot-2 (AN2) runs have an identical nested-domain structure to both the AD and AN runs; however, 12 additional sigma-levels have been added in both AD2 and AN2 for a total of 48 levels. As in AD and AN, the only difference between AD2 and AN2 is that AD2 assimilates TAMDAR data, while AN2 does not. The majority of the additional sigma-levels for AD2 and AN2 are added in the lowest 3 km, or between the 1000 and 700-hPa pressure-levels. This spacing was chosen to best utilize the increased observation density provided by the TAMDAR data. The final two parallel simulations are the AirDat-10 (AD10) and the AirNot-10 (AN10). These two runs have only one domain of 10-km grid spacing. This domain has the same latitude and longitude dimensions as the outer domain of the previously discussed runs. Both the AD10 and AN10 runs have 48 sigma-levels, which are identically spaced to those in the AD2 and AN2 runs.

RUN	OUTER DOMAIN	INNER DOMAIN	LEVELS	TAMDAR
AD	36 KM	12 km	36	Yes
AN	36 KM	12 km	36	No
AD2	36 KM	12 km	48	Yes
AN2	36 KM	12 km	48	No
AD10	10 KM	None	48	Yes
AN10	10 KM	None	48	No

Table 1. The six different parallel model runs in this study

There are multiple angles to this study. First, comparisons are drawn between the like-runs (e.g., AD to AN, AD2 to AN2, etc.) to ascertain any TAMDAR-related impacts. Additionally, comparisons are drawn between AD, AD2, and AD10, with the AN runs as controls, to quantify the effects of increased vertical resolution on the utilization of TAMDAR data. Finally, comparisons are drawn between the AD and AD2 runs to the AD10 run to quantify any change in forecast skill as a function of finer horizontal radial influence during the observation assimilation stage.

The study covers the entire month of May 2006. Twice-daily MM5 simulations were initialized at 0600 and 1800 UTC for all model runs. All simulations were initialized with identical analysis fields provided by the NCAR/AirDat RT-FDDA-MM5. The NCAR/AirDat RT-FDDA system is built around the Fifth Generation of the Penn State/National Center for Atmospheric Research (PSU/NCAR) Mesoscale Model (MM5, Dudhia 1993; Grell et al. 1994). The outer domain of the RT-FDDA-

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

MM5 has a 100 x 97 grid spacing of 36 km, and is centered over the Great Lakes region (Fig. 1). A continuous data ingestion system using Newtonian relaxation is utilized during an analysis period to generate balanced 4-D analyses. This method greatly reduces the time and errors associated with typical model spin-up (Stauffer and Seaman 1994; Cram et al. 2001; and Liu et al. 2002). The NCAR/AirDat RT-FDDA is run on a 3-h cycle with cycle times occuring at 23Z, 2Z, 5Z, 8Z, 11Z, 14Z, 17Z, and 20Z. For this study, the 1-h output from the 0500 and 1700 UTC RT-FDDA cycles, valid 0600 and 1800 UTC, were used as first-guess fields. These files include 1 hour of additional 4DDA nudging; however, they do not include TAMDAR data (i.e., AIRNOT cycles). Boundary conditions were obtained from the ETA 212 via NCEP. After the first-guess field is generated, it is then passed through a 3DVAR-style technique to assimilate additional observations and construct the analysis from which each MM5 simulation is initialized. All of the additional observations are identical for all runs with the exception of the TAMDAR data, which is only assimilated by the AD runs.

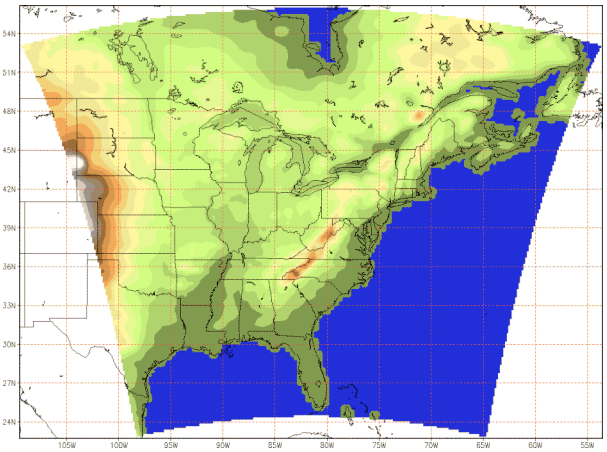

Fig. 1. The Lambert conformal grid for the outer domain of all the simulations.

The MM5 was employed as a means to ascertain optimal combinations and settings for various ingestion, weighting, resolution, and parameterization options. Extensive testing of various parameterizations has been performed to optimize the impact of TAMDAR data (Jacobs and Liu 2006; Jacobs et al. 2006). Those studies suggest that the Kain-Fritsch (KF) cumulus parameterization (CP) is better suited for 20 to 30-km grid spacing, while the Grell CP is better suited for 10-km spacing, and are consistent with findings from Kain and Fritsch (1993). Kain-Fritsch, which generates more convection, may be acceptable if only using the 36-km domain, but when that domain is used as boundary conditions for a finer nested domain such as 12-km, too much convective feedback may occur. Additionally, some CP schemes are better for summer and tropical convection, and some are better for winter mid-latitude convection. Grell tends to handle thunderstorms much better, while KF handles winter frontal systems better (Mahoney and Lackmann 2005). Based on these findings, as well as the season of the study, the Grell cumulus parameterization was chosen for its handling of convective precipitation at smaller grid scales. The MRF planetary boundary layer scheme, as well as the Mixed-phase (Reisner-1) microphysics were also chosen to be consistent with the analysis field generation methods. All simulations were integrated for 12 hours; however, only the forecast-hour 6 (i.e., 0000 and 1200 UTC for the 1800 and 0600 UTC runs, respectively) was used for verification purposes.

To verify the forecast output, model-generated soundings were produced at three locations: Minneapolis/St. Paul, MN (KMPX; 44.8489 N, 93.5656 W; Elev: 946'), Detroit/Pontiac, MI (KDTX; 42.6997 N, 83.4717 W; Elev: 1072'), and Green Bay, WI (KGRB; 44.4983 N, 88.1114 W; Elev: 682'). The motivation for choosing the locations of KMPX and KDTX was based on data density. Both airport locations are regional hubs for Mesaba Airlines. It is assumed that the largest possible data-density-related impacts will be realized in the vicinity of these hubs. KGRB was also included as a 3rd non-hub location because it is located between the other two, and quite often is in line with direct flight paths between KMPX and KDTX. The model-generated soundings were then compared to soundings taken from the North American Regional Reanalysis (NARR), which was obtained from NCDC's NOMADS[1] archive (Kalnay et al. 1990). The NARR/Regional Climate Data Assimilation System (R-CDAS) is an extension of the NCEP Global Reanalysis (GR), which is run over the CONUS. The NARR uses the high resolution NCEP Eta Model (32km/45 layer) together with the Regional Data Assimilation System (RDAS). The NARR output is a native E-grid, and is interpolated to a Northern Lambert Conic Conformal (NCEP Grid 221) projection before the data is archived (Mesinger et al. 2006). The variation in grid orientations resulted in minor rounding (<0.2%) errors as the non-grid-relative U and V wind values were backed out. Verification with the NARR is considered a crude, yet objective and consistent way to quantify performance differences between the simulations. A more precise method would employ raw RAOB data; however, that would require varying x and y since the balloon drifts as it ascends. The NARR method allows for direct comparison between true vertical profiles. The potential error and inaccuracy from the coarse grid spacing of NARR is still lower than the potential difference seen as RAOBs can sometimes drift large horizontal distances. The reliability of the upper-level NARR data is considered very good since the verification times are 0000 and 1200 UTC, and coincide with the largest RAOB-based analysis adjustments.

The forecast bias is simply defined as the 6-hour forecast value minus the NARR value. In the case of the 1800 UTC simulations, differences are calculated at 0000 UTC. The forecast mean absolute error is defined as the averaged magnitude of this difference. Thus, it

[1] http://nomads.ncdc.noaa.gov/

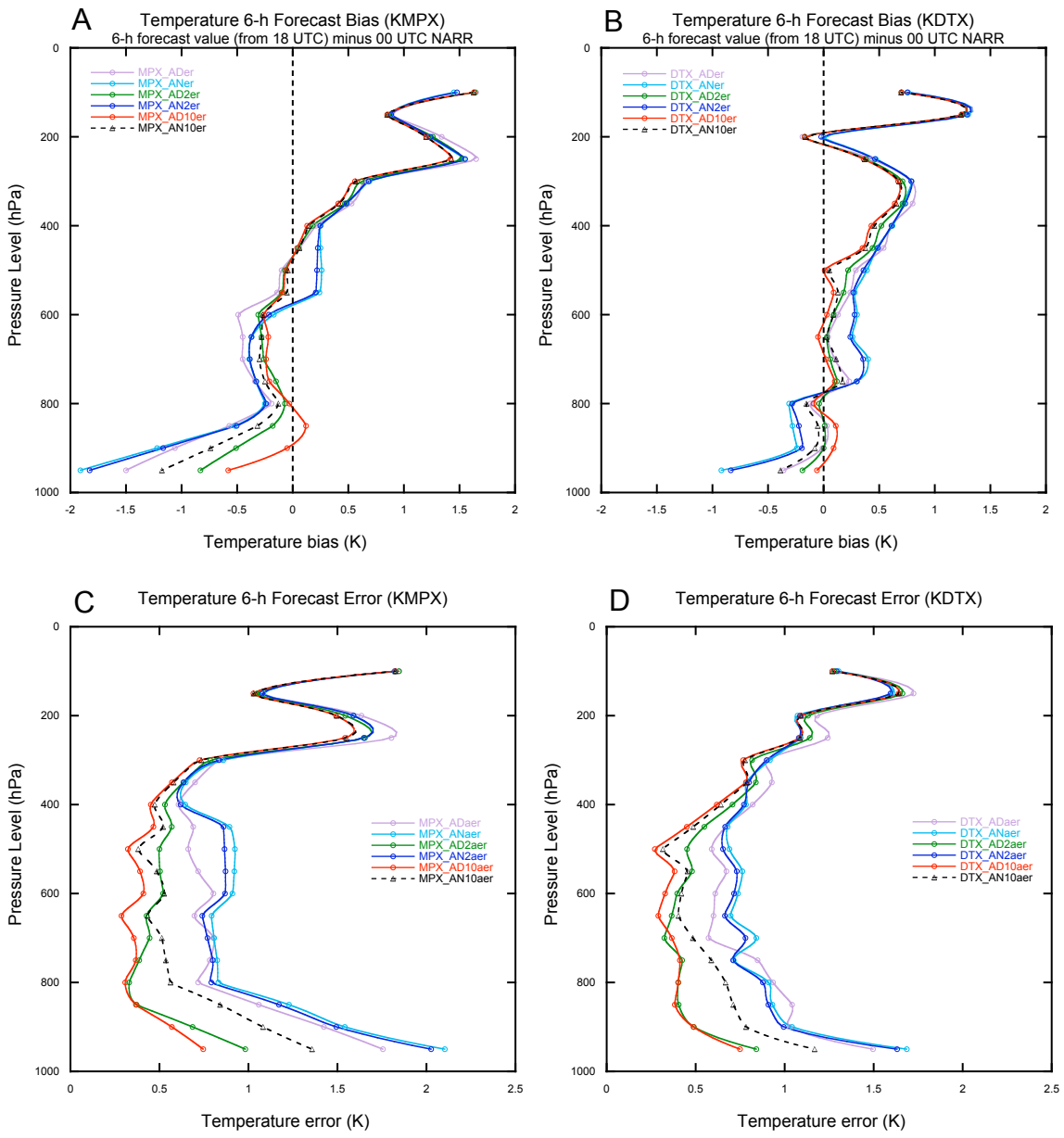

Fig. 2. Vertical profile of the 6-h temperature forecast bias (K) for May 2006 for the KMPX location (A) and the KDTX location (B) valid 0000 UTC, and the 6-h temperature forecast MAE (K) for the KMPX location (C) and the KDTX location (D) valid 0000 UTC.

should be taken as the divergence from the NARR. A percent reduction in error is seen as a percentage increase in forecast skill. This percent improvement is defined as

$$\%IMP = -\frac{\alpha - \beta}{\beta} \times 100, \qquad (1)$$

where simulation α is compared to simulation β, and appears contextually below as α versus β (Brooks and Doswell 1996).

3. SENSITIVITY TO σ-LEVEL DENSITY AND HORIZONTAL GRID SPACING

It should be noted that all of the plots below correspond to the 0000 UTC comparison. The 0000 UTC time historically shows a larger difference because (i) there are more flights between 0700 and 1900 EDT, and (ii) the lower-troposphere is less stable.

A vertical profile of the 6-h temperature forecast bias (K) averaged for the entire month of May 2006 for KMPX is shown in Fig. 2A. A warm bias trend exists above 450 hPa and a slight negative bias exists below 500 hPa, which increases (negatively) below 850 hPa. In the majority of profiles, there appears to be minimal

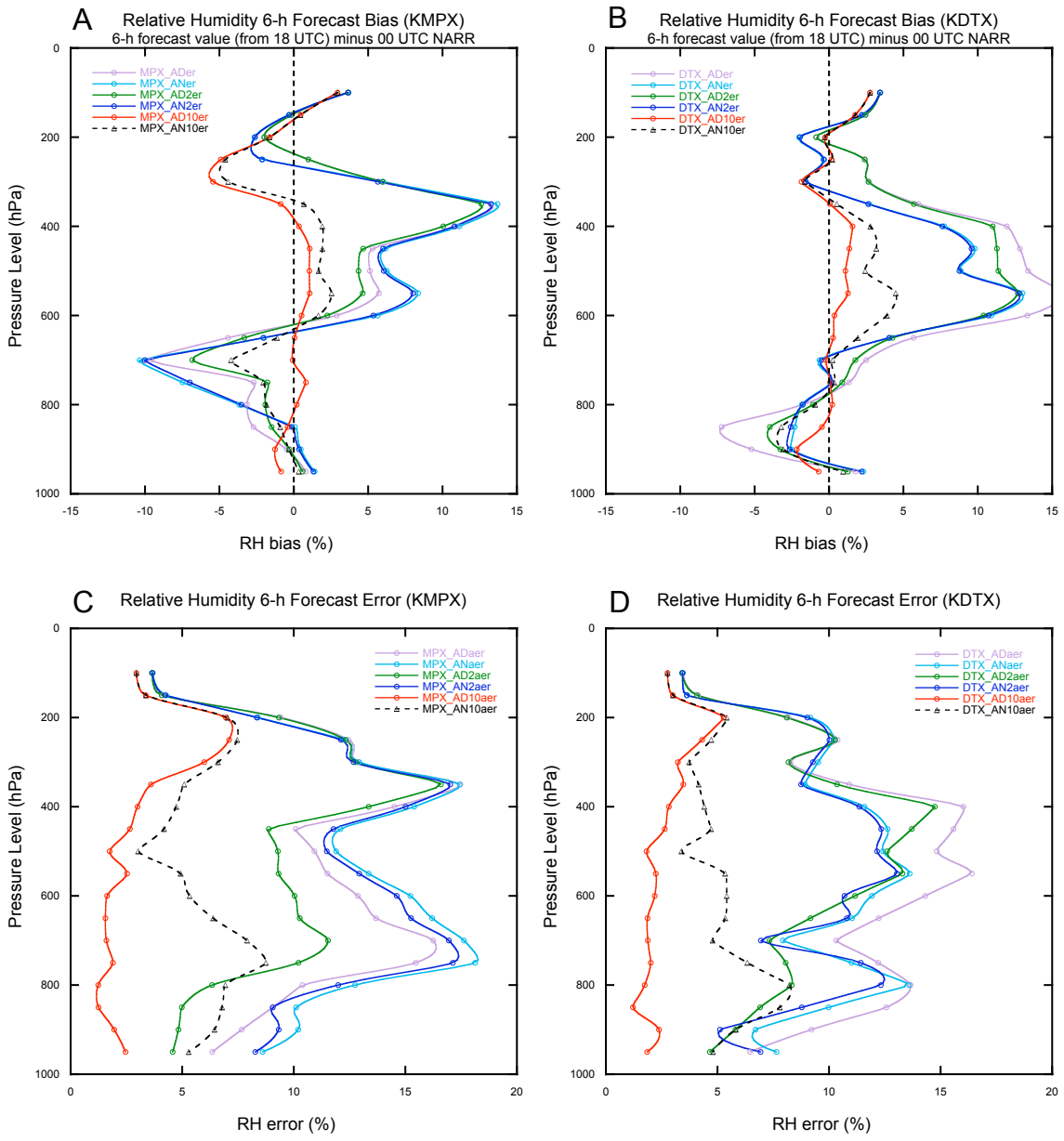

Fig. 3. Vertical profile of the 6-h relative humidity forecast error (%) during May 2006 for the KMPX location (A) and the KDTX location (B) valid 0000 UTC, and the 6-h relative humidity forecast MAE (%) for the KMPX location (C) and the KDTX location (D) valid 0000 UTC.

difference between the AN and AN2 runs. Around the 500-hPa level, all of the AD runs, as well as the AN10 run, perform quite well. Below 850 hPa, AD10 does the best at avoiding the cold trend followed by the others. There is noticeably less divergence between the model runs at the KDTX location (Fig. 2B). This could possibly be related to fewer flights, as well as less thermal variability from Great Lakes-enhanced boundary layer moisture content. The 6-h temperature forecast error for the KMPX location appears to be greater in magnitude above 300 hPa; however, it is assumed to be unrelated to the TAMDAR data and grid variations since similar trends are seen for all simulations. Divergence between the simulations begins to occur below 400 hPa. Between 400 and 800 hPa, the error magnitude does not change. The AN and AN2 runs have an average error around 0.8 K, while the AD and AD2 runs have errors less than 0.5 K. Below 850 hPa, the error increases, but appears to increase at more than twice the rate in the TAMDAR-withheld simulations. Similar trends are also seen above 400 hPa for the KDTX location (Fig. 2D). Below 400 hPa, both AD2 and AD10 show a significant decrease in error over AN and AN2, as well as even the AD run.

A vertical profile of the 6-h relative humidity forecast bias (%), averaged for the entire month of May 2006 for

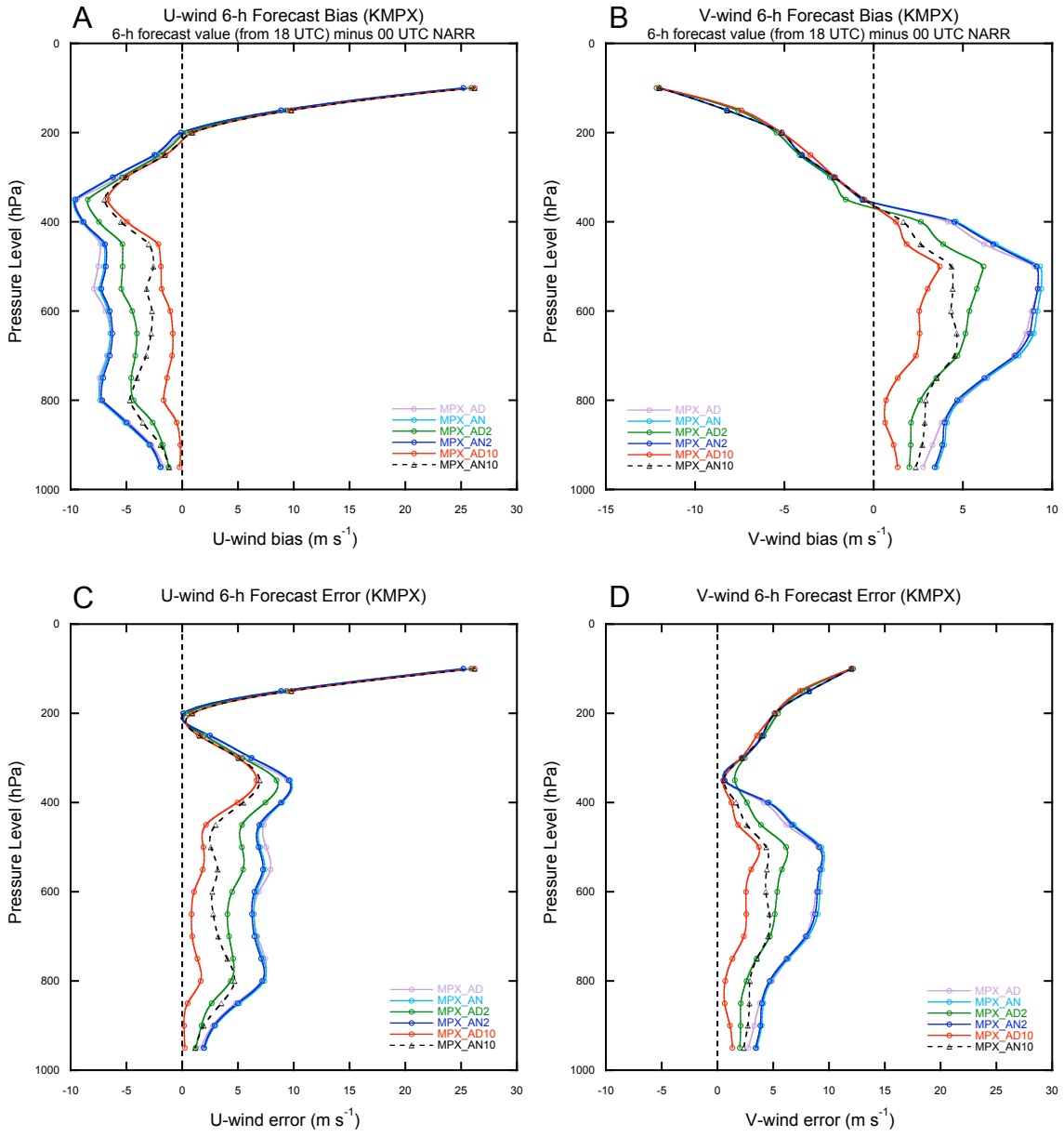

Fig. 4. Vertical profile of the 6-h wind forecast bias (m s^{-1}) during May 2006 for U-wind (A) and V-wind (B) valid 0000 UTC for the KMPX location, and the 6-h wind forecast MAE (m s^{-1}) for U-wind (C) and V-wind (D) valid 0000 UTC for the KMPX location.

KMPX, is shown in Fig. 3A. A large moist bias (>10%) is seen between 250 and 600 hPa for both AD and AD2, as well as AN and AN2. These same runs also show a dry bias (~5%) at 700 hPa. The AD2 and AN10 are the two best simulations at reducing this bias, as the AD10 run appears to be unaffected by any noticeable bias trends below 400 hPa and remains less than ±2%. Similar trends are seen above 500 hPa for the RH bias profile at KDTX where all the simulations, with the exception of AD10 and AN10, are moist (Fig. 3B). Interestingly, the TAMDAR-included runs AD and AD2 have a larger moist bias than AN and AN2. Since all of the simulations began with identical first-guess fields, this suggests that possible sources for this difference are either related to the 3DVAR data assimilation, or a bias in the NARR. Since the better-performing AD10 and AN10 runs appear to correct this issue, it is assumed that the problem is not linked to the NARR, but is likely a function of upstream TAMDAR observations interfering with the proper assimilation of downstream non-TAMDAR observations by interacting with the background error covariances. Similar, but dry, bias trends are seen for this location below 850 hPa. A vertical profile of the 6-h relative humidity forecast error (%), averaged for the entire month of May 2006 for KMPX, is shown in Fig. 3C. All simulations follow a similar trend above 200 hPa. Between 200 and 500 hPa, significant divergence between the two 10-km

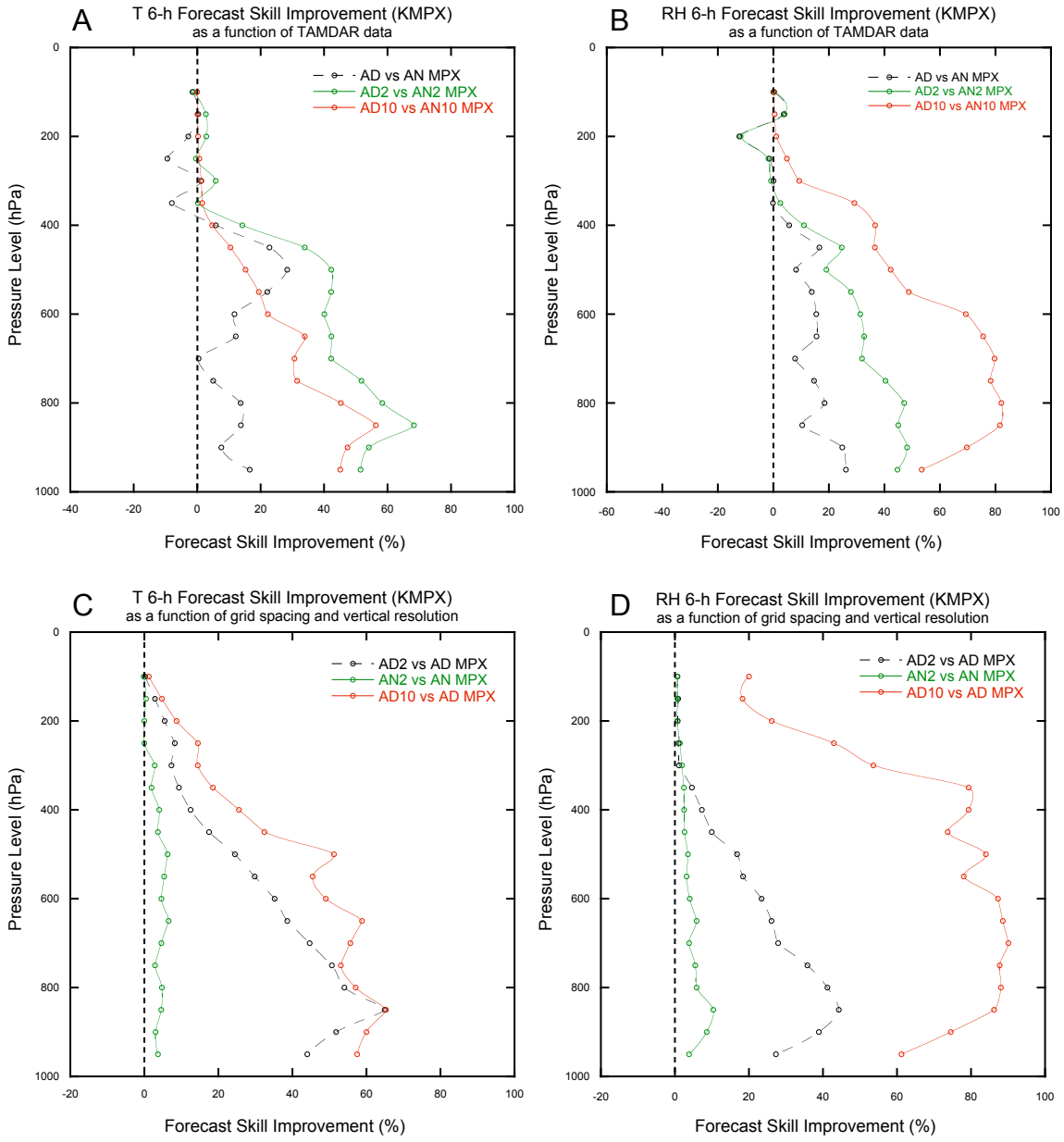

Fig. 5. Vertical profile from the KMPX location during May 2006 of the 6-h forecast skill improvement, as a function of TAMDAR data, w.r.t. NARR for temperature (A) and relative humidity (B), and the 6-h forecast skill improvement, as a function of grid spacing and vertical resolution, w.r.t. NARR for temperature (C) and relative humidity (D).

simulations and the other runs appears. The error in both AD10 and AN10 remains less than 5%, while the error in AD, AD2, AN, and AN2 follow similar trends, which average more than 10% and exceed 15% near 350 hPa. Since there are no TAMDAR observations near this altitude, and both the TAMDAR-included and TAMDAR-withheld runs align, it suggests that the source of the error is linked to horizontal grid spacing in conjunction with the assimilation of non-TAMDAR observations. AD10, and to a lesser degree, AN10 outperform the other runs throughout the profile with the exception of the AD2 run surpassing the skill of AN10 below 800 hPa. Similar trends are seen for the KDTX location with the exception of the AD and AD2 errors discussed above (Fig. 3D).

Vertical profiles of the 6-h wind forecast bias (m s^{-1}) at KMPX during May 2006 for U-wind and V-wind, valid 0000 UTC, are shown in Figs. 4A and 4B, respectively. A significant positive bias in the U-wind and negative bias in the V-wind are seen near the top of the respective profiles. Since all of the simulations align while showing identical trends, it is assumed that this bias is either MM5 code-related or a function of the Eta-32 boundary conditions. The important focus of the study deals with the differences between the runs, moreso than the error magnitude, within the lower and

middle troposphere. Within that region, the U-wind has a negative bias, while the V-wind has a positive bias, for all simulations. The vertical profile of the 6-h wind forecast error (m s^{-1}) for U-wind (Fig. 4C) and V-wind (Fig. 4D) show similar error magnitudes despite the mirror image of bias value. The 10-km simulations clearly outperform the other runs with the AD10 simulation being the better of the two. The AD2 run reduces the error by half as compared to the remaining three simulations (AN, AN2, and AD).

A crude method for quantifying forecast skill improvement is shown in Fig. 5, which is valid 0000 UTC for the KMPX location. As mentioned in section 2, this method was chosen for consistency and simplicity. The improvement as a function of TAMDAR data for temperature is shown in Fig. 5A. Minor positive impact is seen for the AD run over the AN run, while the largest improvement noted is between AD2 and AN2. The positive impact of TAMDAR data in the AD10 run over the AN10 was not as significant. This is because the AN10 had markedly less error to begin with, as well as the ability of AD2 to nearly match the error reduction seen in AD10. The most notable improvement as a function of TAMDAR data appears to be linked to the utilization of RH observations in the AD10 simulation (Fig. 5B). A substantial reduction in error is seen between AD10 and AN10, which reaches a maximum of more than 70% between the 650 and 900-hPa levels. The other two comparisons show similar trends, but to a lesser magnitude.

The improvement in temperature forecast skill as a function of grid spacing and vertical resolution is shown in Fig. 5C. It is evident that the increase in vertical resolution between AN and AN2 made only a very slight positive contribution. Whereas, the same addition of sigma-levels made a significant difference between the other simulations. This suggests that without the inclusion of a higher-density data set, the high vertical resolution only contributes a small amount to increasing forecast skill. The difference between the AD2-vs-AD line and the AD10-vs-AD line can be attributed to the increase in horizontal resolution (Fig. 5C). The linear increase in skill, which peaks for both near 850 hPa, is likely a function of the additional sigma-levels being evenly spaced between the previous lower and middle-tropospheric levels (Figs. 5C and 5D). A similar trend of negligible-to-slight improvement between AN and AN2 is seen for RH (Fig. 5D). As seen with temperature, the same addition of sigma-levels made significant differences between the other simulations. Once again, the difference between the AD2-vs-AD line and the AD10-vs-AD line for RH can be attributed to the increase in horizontal resolution. However, unlike temperature, the difference in horizontal resolution produced a much greater improvement in RH forecast skill (Fig. 5D). This increase in skill is seen throughout the profile suggesting that proper observation assimilation of RH data, whether it is from TAMDAR or not, is not only dependent on vertical resolution, but is also highly dependent or horizontal grid spacing.

4. CONCLUSIONS

The degree of forecast skill improvement presented here is seen as a best-case result with the current level of model optimization and data quality because the verification location (KMPX) is a Mesaba Airlines hub, which naturally hosts the greatest density of observations. It is assumed that additional regions surrounding (and downstream) of future fleet hubs will likely realize similar improvements.

Results suggest that the addition of TAMDAR data improves all three experimental simulations for key model variables, albeit to various degrees. In general, increasing the number of sigma-levels from 36 to 48 results in better utilization of the higher resolution TAMDAR data.

The most notable improvements, when increasing the number of model sigma-levels, were found to occur in the vicinity of 850 hPa. This peak is likely a result of the balance between the increase in TAMDAR observations and the increase in expected model error (i.e., seen in the control runs) when approaching the surface layer from the middle-troposphere.

Changing the outer domain grid spacing from 36 km to 10 km reduces the bias and error for both the experimental (AD10) and the control (AN10); however, the magnitude of reduction is greater in AD10. This is likely because all of the non-TAMDAR observations assimilated into the first-guess fields of both runs improved the respective analysis fields. The additional reduction in error magnitude seen in AD10 is attributed to the TAMDAR data (the only difference between AD10 and AN10).

There are a few possibilities that may have influenced this outcome since the observations are only assimilated into the outermost grid. With a larger grid spacing, the error statistics of observations which are interpolated to identical (i, j, k) grid space can negatively influence the impact of redundant observations. The probability of this occurrence is reduced when assimilating data into a finer horizontal grid because the frequency of interpolating to identical (i, j, k)-space is reduced. The former appears to be a less problematic choice when using a background error-covariance matrix based on a larger grid in conjunction with TAMDAR data, which posses greater variability in wind-field errors when compared to the lower mass-field errors. Further cost function analysis is required to avoid pitfalls outlined by Xie et al. (2002). Additionally, some of the forecast skill success of the 10-km simulations is likely a function of the Grell CP scheme, which is better suited for this grid size versus the larger 36-km spacing (Grell 1993).

With respect to temperature from TAMDAR, the largest improvements were seen between AD2 and AN2. The additional sigma-levels made a notable difference; however, reducing the grid size from 36 to 10 km, along with the additional sigma-levels, had less of an impact. This is partially because the control (AN10) also had less error. This finding is consistent with the understanding that temperature, particularly in regions like KMPX during the summer months, can vary more in (z) than in (x, y).

With respect to relative humidity from TAMDAR, the largest improvements were seen between AD10 and AN10. This suggests that increasing the horizontal resolution is just as important, if not more important, than increasing the vertical resolution, and is consistent with the understanding that RH is extremely variable in both (x, y) as well as (z).

5. ACKNOWLEDGMENTS

The authors would like to thank Stan Benjamin and William Moninger (NOAA/ERL/FSL) for their comments and suggestions on this work. The authors are very grateful for the 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.

6. REFERENCES

Brooks, H. E., and C. A. Doswell, 1996: A Comparison of Measures-Oriented and Distributions-Oriented Approaches to Forecast Verification. Weather and Forecasting, **11**, 288-303.

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.

Dudhia, J., 1993: A nonhydrostatic version of the Penn State / NCAR mesoscale model: Validation tests and simulation of an Atlantic cyclone and cold front. *Mon. Wea. Rev.*, **121**,1493-1513.

Grell, G. A., 1993: Prognostic evaluation of assumptions used by cumulus parameterizations. *Mon. Wea. Rev.*, **121**, 764-787.

Grell, G. A., J. Dudhia, and D. R. Stauffer, 1994: A description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note NCAR/TN-398+STR,122 pp.

Jacobs, N. A., and Y. Liu, 2006: A Comprehensive Quantitative Precipitation Forecast Statistical Verification Study, Documentation and Tech. Note AirDat, LLC, 25 pp.

Jacobs, N. A, Y. Liu, and C.-M. Druse, 2006: Evaluation of temporal and spatial distribution of TAMDAR data in short-range mesoscale forecasts, AMS Annual Meeting, 10th Symp. IOAS-AOLS.

Kain, J. S., and J. M. Fritsch, 1993: Convective parameterization for mesoscale models: The Kain-Fritsch scheme. The representation of cumulus convection in numerical models, K. A. Emanuel and D. J. Raymond, Eds., Amer. Meteor. Soc., 246 pp.

Kalnay, M. Kanamitsu, and W.E. Baker, 1990: Global numerical weather prediction at the National Meteorological Center. Bull. Amer. Meteor. Soc., **71**, 1410-1428.

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, 12-16 August, 2002, San Antonio, Taxes, 399-402.

Mahoney, K.M., and G.M. Lackmann, 2006: The sensitivity of numerical forecasts to convective parameterization: A case study of the 17 February 2004 east coast cyclone. *Wea. Forecasting*, **21**, 465–488.

Mesinger F., G. DiMego, E. Kalnay, K. Mitchell, P.C. Shafran, W. Ebisuzajki, D. Jovic, J. Woollen, E. Rogers, E. H. Berbery, M.B. Ek, Y. Fan, R. Grumbine, W. Hyggins, H. Li, Y. Lin, G. Manikin, D. Parrish, and W. Shi, 2006: North American Regional Reanalysis, *Bulletin of the American Meteorological Society*, **87**: 343-360.

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

Xie, Y. F., C. Lu, and G. L. Browning. 2002. Impact of formulation of cost function and constraints on three-dimensional variational data assimilation. *Mon. Wea. Rev.* **130**, no. 10: 2433-47.