11.9 EVALUATION OF BOUNDARY LAYER STRUCTURE IN SEVERAL OPERATIONAL AND RESEARCH WEATHER PREDICTION MODELS

James Wilczak[1], Irina Djalalova[1], Robert Zamora[1], Jian-Wen Bao[1],
Jeff McQueen[2], Stan Benjamin[3], and Georg Grell[3]

[1]Environmental Technology Laboratory, NOAA
[2]National Centers for Environmental Prediction, NOAA
[3]Forecast Systems Laboratory, NOAA

1. INTRODUCTION

During the summers of 2002 and 2003 intensive meteorological observation and modeling programs took place in the New England region. The goals of these studies [the 2002 New England Air Quality Study (NEAQS-2002) and 2002 and 2003 New England High Resolution Temperature Programs (NEHRTP-2002, NEHRTP-2003)], were to develop a better understanding and forecast skill for the meteorological transport processes associated with high ozone events, and to improve model forecasts of surface temperature. Since the atmospheric boundary layer plays a key role in both of these forecast problems, an important component of each of these programs was an assessment of the ability of present operational and research models to predict boundary layer growth and structure.

Our model evaluation includes an assessment of the atmospheric boundary layer (ABL) depth, the diurnal variation of winds and temperatures within the ABL, and model solar irradiance parameterizations. The model evaluation includes assessments of each model's bias statistics over the summer field season. The statistical evaluations are used to identify specific parameterizations within the models that are sources of significant model error that can affect ozone predictions.

In this paper we focus on the summer 2002 field studies. Similar instrumentation and models were used during the 2003 studies, and the analysis of that data is in progress. During the summer of 2004 both surface temperature and air quality studies are continuing in New England. As part of these field programs an expanded set of meteorological instrumentation is being deployed, which centers on a boundary layer super-site at Plymouth, MA. The instrumentation for this site will be briefly described during the conference.

* *Corresponding author address*: James Wilczak, NOAA/ETL; 325 Broadway, Boulder 80305, CO; e-mail: james.m.wilczak@noaa.gov.

2. OBSERVATIONAL DATA

The core set of instrumentation used for the NEAQS-2002 and NEHRTP-2002 model assessment studies was a network of seven 915 MHz wind profilers (Fig. 1). The wind profilers provided hourly averages of wind speed and direction, typically to heights of 3000m AGL. In addition to winds, the profilers measured the vertical profile of virtual temperature using the Radio Acoustic Sounding System (RASS) technique, which typically reached heights of 1000 m AGL. Hourly wind averages were computed after first visually inspecting each days moment data for bird contamination (Wilczak et al., 1995). The moment data were then processed using the Weber and Wuertz (1992) algorithm, which uses a pattern recognition scheme first to determine hourly values, and then applies a second pass of a similar algorithm to eliminate outliers from the hourly winds and temperatures. The data were then visually examined and a few remaining outliers in the hourly values were eliminated.

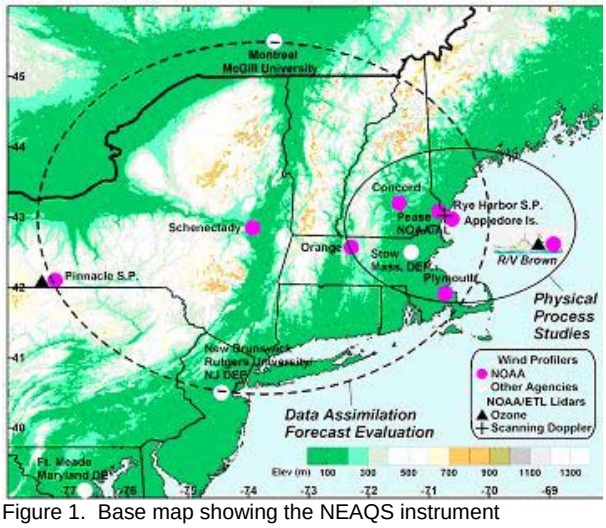

Figure 1. Base map showing the NEAQS instrument deployment. The wind profilers used in this assessment study are the land-based NOAA wind profilers (magenta circles).

In addition to the wind and virtual temperature profiles, the radar wind profilers also can provide measurements of the depth of the daytime, convective ABL (White, 1993; Angevine et al., 1994; Bianco and Wilczak, 2002). Values of range-corrected signal to noise ratio, vertical velocity (which is large within the convective ABL), and radar spectral width (which is a measure of turbulence intensity) were visually inspected to determine the ABL depth.

At each of the wind profiler sites, surface meteorological observations were also taken that are used in the model evaluation study. These included the standard 2m temperature and humidity, 10m winds, pressure and precipitation, and solar and net radiation.

3. MODEL DESCRIPTION

Four meteorological models were evaluated during NEAQS 2002. The first of these is the operational Eta model, which was obtained in real-time from NCEP and which will be used in NOAA's initial 2004 operational ozone forecast system. The second model is a version of the developmental Weather Research and Forecasting (WRF) model. The third is MM5, which has frequently been used by state and local agencies for EPA mandated ozone assessment studies, and therefore provides a good benchmark of historical methodologies. Finally, we also have included the RUC model, which is run operationally. The Eta and MM5 model data used in this study are from real-time forecasts made during the summer of 2002, while both the WRF and RUC models were run in a retrospective mode during the autumn of 2003. Between 2002 and 2003 updates have been made to the operational Eta model that alleviate some of the biases that are documented in this study.

Details of each of the models and their parameterization schemes can be found Table 1. The WRF model used for the NEAQS 2002 simulations was configured with the intent to make its parameterizations similar to those in the Eta. Both use the same surface layer scheme (Monin-Obhukov as implemented by Janic in the Eta model), and the same boundary-layer TKE parameterization (Mellor-Yamada-Janjic). For a land-surface model WRF uses the OSU parameterization, while the Eta uses a later derivative (NCEP's Noah parameterization) which is very similar. For shortwave radiation both the Eta and WRF use implementations based on the original Lacis and Hansen code, while for longwave the Eta uses a GFDL parameterization and the WRF the RRTM code.

	Eta	WRF	MM5	RUC
LSM	Noah	OSU	Smirnova	Smirnova
Surface layer	MO	MO (Eta)	MO (Hi-res PBL)	MO (Pan 94)
PBL	MY 2.5	MY 2.5 (Eta)	MY 2.5	MY 2.5 (BT)
Short-wave	Lacis& Hansen	Dudhia	Dudhia	Dudhia
Longwave	GFDL	RRTM	RRTM	Dudhia
Microphysics	Ferrier	NCEP 3-class	Reisner 5 class	Reisner 5 class
Cumulus precip	BMJ	BMJ	Grell-Dev. Ens	Grell-Dev. Ens
Initial Cond	EDAS	RUC 3Dvar	RUC (OI)	RUC (OI)
Boundary Cond	GFS	Eta	Eta	Eta
Horiz grid	12km	27km	27km	20km
Vert levels	60	25	30	50
Lvls <1km/2km	20/26	12/16	11/14	10/17
Horiz adv		5th	2nd	Smol. pos-def
Vert adv		3rd	2nd	2nd
Time integr		Runge-Kutta (150s)	Leapfrog (80s)	Adams-Bashforth (30s)

Table I. Summary of model parameterization schemes.

The MM5 and RUC models also had physical parameterizations that tended to be similar to each other, but different from those of the Eta and WRF. Both used the Smirnova LSM, and Monin-Obhukov similarity theory and Mellor-Yamada level 2.5 turbulence schemes for the surface layer and PBL. For radiation, MM5 used the same long and shortwave parameterizations as the WRF, while the RUC used the Dudhia longwave scheme.

For initial conditions the Eta model uses its own EDAS assimilation system, and it uses the GFS model for boundary conditions. In turn, the WRF, MM5 and RUC all use boundary conditions from the Eta model. The RUC and MM5 models were initialized using the RUC's Optimal Interpolation analysis, while the WRF model was initialized using a latter RUC 3Dvar technique. Finally, we note that the Eta model had higher resolution, including almost twice the vertical resolution in the lowest km compared to the other three models.

4. STATISTICAL EVALUATION

For each of the models, surface and vertical profile data were extracted at the model gridpoints closest to the coordinates of each

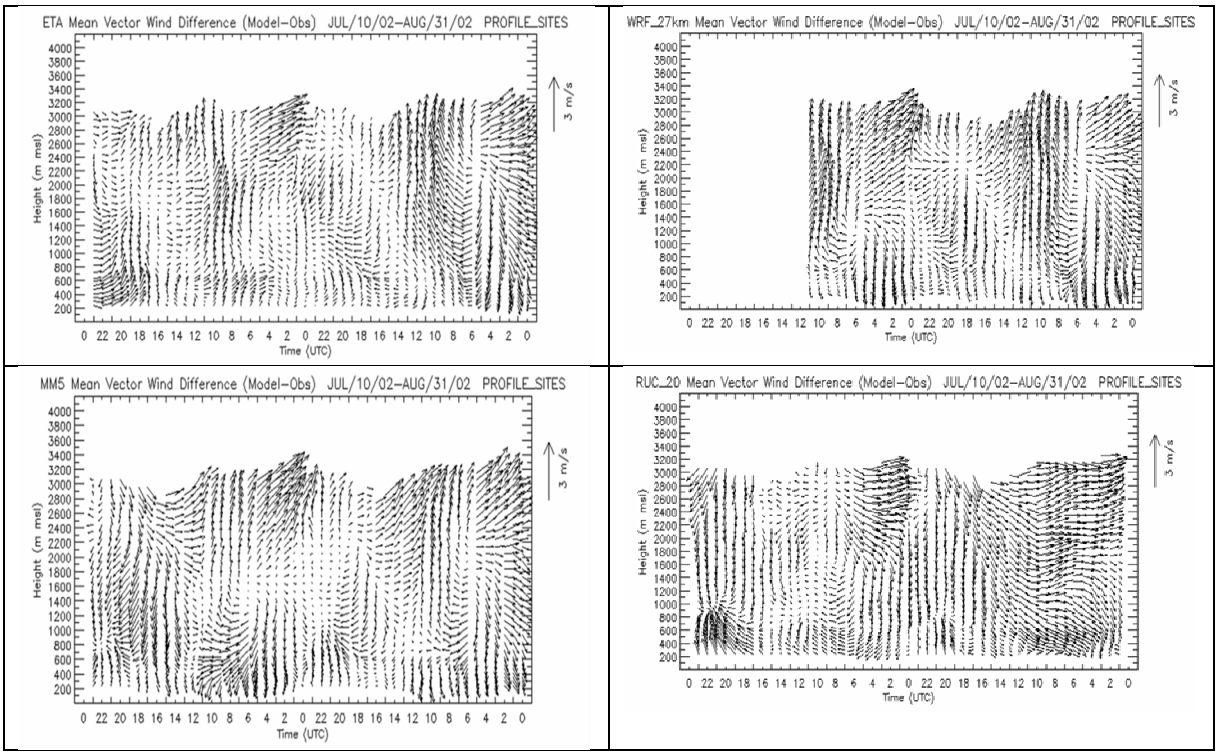

Figure. 2 Vector wind biases for the for the 00 UTC model initialization cycle of the Eta, WRF, MM5, and RUC models, averaged for all 7 profiler sites for 53 days.

observation site. Because each of the models have their own unique vertical coordinates, for comparison purposes we have interpolated the model values to the exact heights of the wind profiler and RASS observations. Bias errors were then calculated using each 48 hour (36 for WRF) forecast made on 53 consecutive days between 10 July and 31 August, 2002. The bias is defined as the model minus observed value, and is calculated at hourly intervals (the output frequency of each of the models). Separate statistics are calculated for the 00 UTC and 12 UTC cycle runs of each model. For the profiler statistics, a minimum of 60% of the possible total observations is required to plot the data.

Although 7 profiler sites collected data during NEAQS, the Appledore Island site, which is located approximately 10 km off-shore of New Hampshire (Fig. 1), has been excluded from the surface analysis because although the island is small (6 km by 1 km), it nevertheless affected the surface meteorological observations sufficiently such that the site is not representative of either an ocean gridpoint nor a land gridpoint for comparison to the models. In addition, with the dominant off-shore flow, the boundary layer at the site is typically one of transition from a deep continental convective ABL to a shallow marine ABL. This made it difficult to determine what was the true depth of the turbulent boundary layer in the observations, and so the site has also been excluded from the ABL depth analysis.

Figure 2 shows the vector wind difference plots for the 00 UTC cycle of all four models, averaged over all 7 wind profiler locations. Time in all of the figures increases to the left, following the standard meteorological convention. The most striking aspect of this figure is the remarkable similarity of the wind biases among the Eta, WRF, and MM5 models for heights greater than about 1 km. In contrast, the RUC has vector wind biases about the same magnitude as the other three models, but with a significantly different diurnal time-height pattern for times beyond the first 4-5 hours of the forecast period.

The greatest dissimilarities among the four models shown in Fig. 2 tend to be in the lowest 1 km. This is especially true during the nighttime hours (0-10 UTC), and it suggests that a significant difference between the models is how they handle the nocturnal boundary layer.

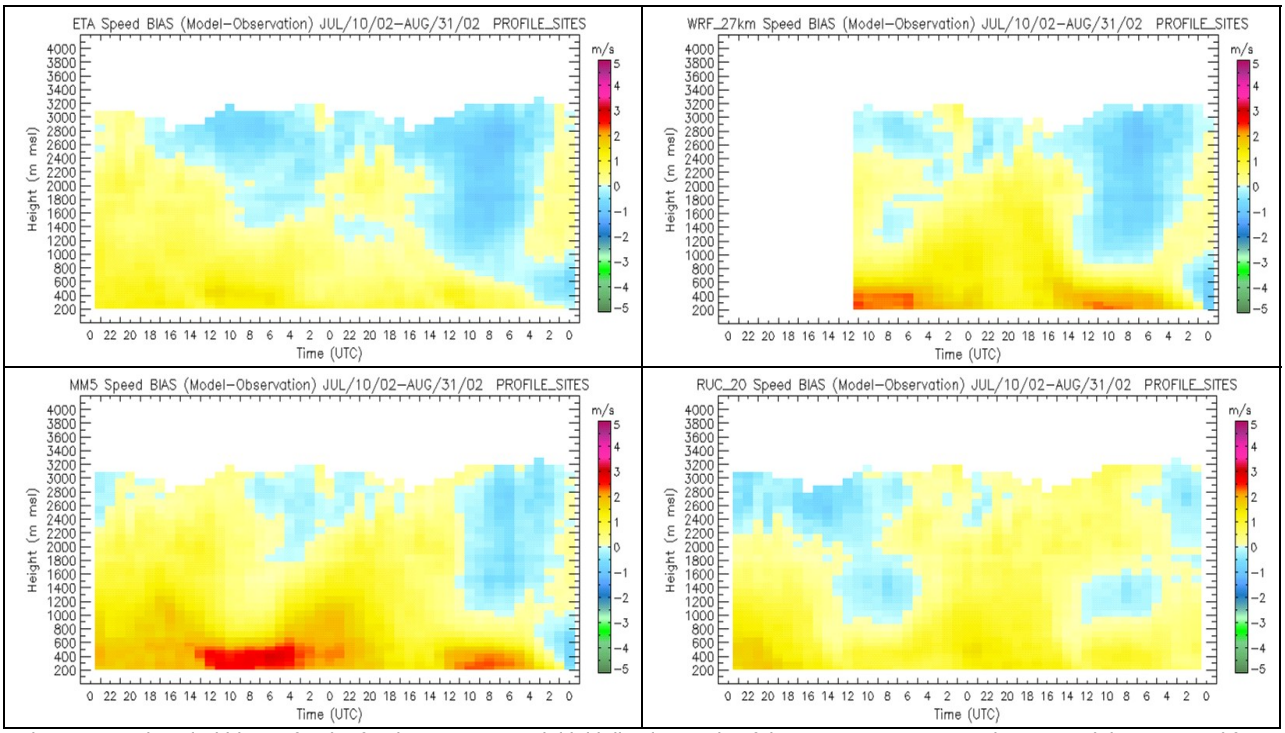

Figure. 3 Scalar wind biases for the for the 00 UTC model initialization cycle of the Eta, WRF, MM5, and RUC models, averaged for all 7 profiler sites for 53 days.

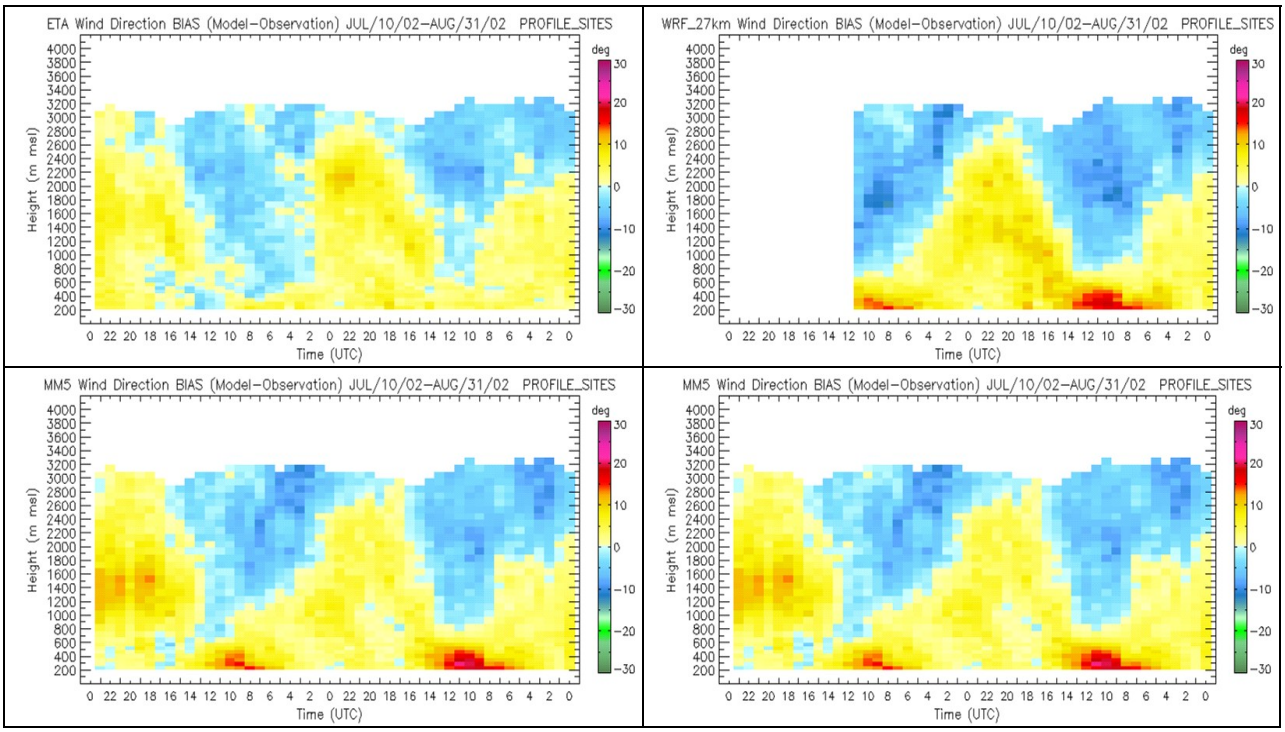

Figure 4. Wind direction biases for the for the 00 UTC model initialization cycle of the Eta, WRF, MM5, and RUC models, averaged for all 7 profiler sites for 53 days.

The next set of four panels in Fig. 3 show the scalar wind biases for each of the models, again averaged for all of the profiler sites over the summer field campaign. The Eta's speed bias is small, mostly varying between +- 1 m/s, with a positive bias at lower levels and a negative bias at upper levels, with the positive bias slowly increasing with time. A small diurnal variation in the bias is present. In comparison, the WRF wind speed bias shows the same diurnal tendency, but with a considerably larger bias of up to 3 m/s in the nocturnal boundary layer. MM5's wind speed bias is very similar to that of WRF, again with a large positive speed bias in the nocturnal boundary layer. Interestingly, the RUC model's scalar wind speed bias is small overall. Given that the RUC has the same coarser low-level vertical resolution as WRF and MM5, it does not seem likely that the larger errors in WRF and MM5 can be attributed to vertical resolution. The mix of LSM's, surface layer and PBL turbulence schemes shown in Table 1 for the various models also suggests that none of these is alone responsible for the larger low-level speed errors in WRF and MM5. Also, we note that there is a small diurnal variation in the speed bias in the RUC, but that it is out of phase with those in WRF and MM5, with a minimum during the nighttime hours and a maximum in the afternoon around 21 UTC.

In Fig. 4 we show the scalar wind direction bias for the four models. The scalar direction bias is determined by taking the difference in the wind direction between the model and the observation, with positive values if the model direction is clockwise from the observed, and then averaging these values regardless of the wind speed. Here a diurnal variation is readily apparent at upper levels, more so than was seen in the wind speed. The direction bias for all models is most negative near sunrise and most positive near sunset. This diurnal variation is smallest in the Eta and then the RUC models, and largest in the WRF, where the amplitude of the variation is as large as 10 degrees. At low levels the Eta and RUC model biases have little diurnal variation and are small overall, typically less than 5 degrees. In contrast, both MM5 and WRF have a larger diurnal variation, with biases approaching 20 degrees during the early morning hours. The positive speed bias and clockwise direction bias of MM5 and WRF suggest that there is insufficient frictional coupling with the surface, deceleration, and cross-isobaric flow, but only during periods of stable stratification.

Biases of virtual temperature calculated from the RASS observations and model predictions (derived from the model's temperature and humidity fields) are shown in Fig. 5. All 4 models tend to have the same diurnal variation in the bias, with a relatively warm bias at lower altitudes beginning in the mid to late afternoon continuing through the night, and a relatively cold bias starting near sunrise and continuing until the afternoon. In the Eta model the amplitude of this diurnal variation is about 1 C, and it is slightly larger in the other 3 models. In addition, the WRF and MM5 models have a cooling trend of approximately 1 C/day superimposed on the diurnal variation.

In Fig. 6 we show the ABL depth biases for the 4 models for the 00 UTC cycle, averaged for all profiler sites except Appledore Island. The method that NCEP uses to diagnose the ABL depth is based on the TKE, and we use the same method here. This method simply searches for the model level where the TKE changes from the model background TKE value to a higher value associated with boundary layer turbulence. Using this method, the Eta model is found to have small depth biases, on the order of 100m or less

For MM5, WRF, and the RUC models we have applied the exact same technique to determine the ABL depth as was used for the Eta model. WRF and MM5 show fairly good agreement with the observations initially, but large negative biases in the afternoon. At 19 UTC, which is the time of maximum ABL depth the depth bias is approximately -400m for MM5 and the WRF, which is approximately 35% of the observed mean depth at this time. Although this seems quite large, we note that because of the coarse resolution of MM5, 400m is only slightly greater than the vertical grid spacing of the model at the inversion height, while for the WRF model it is closer to a two grid-level error. In contrast, the RUC has a moderate positive bias of 100-200 m on the first day of the forecast. On day two the RUC starts out with a larger positive bias of 350 m, but the bias disappears completely by mid-day.

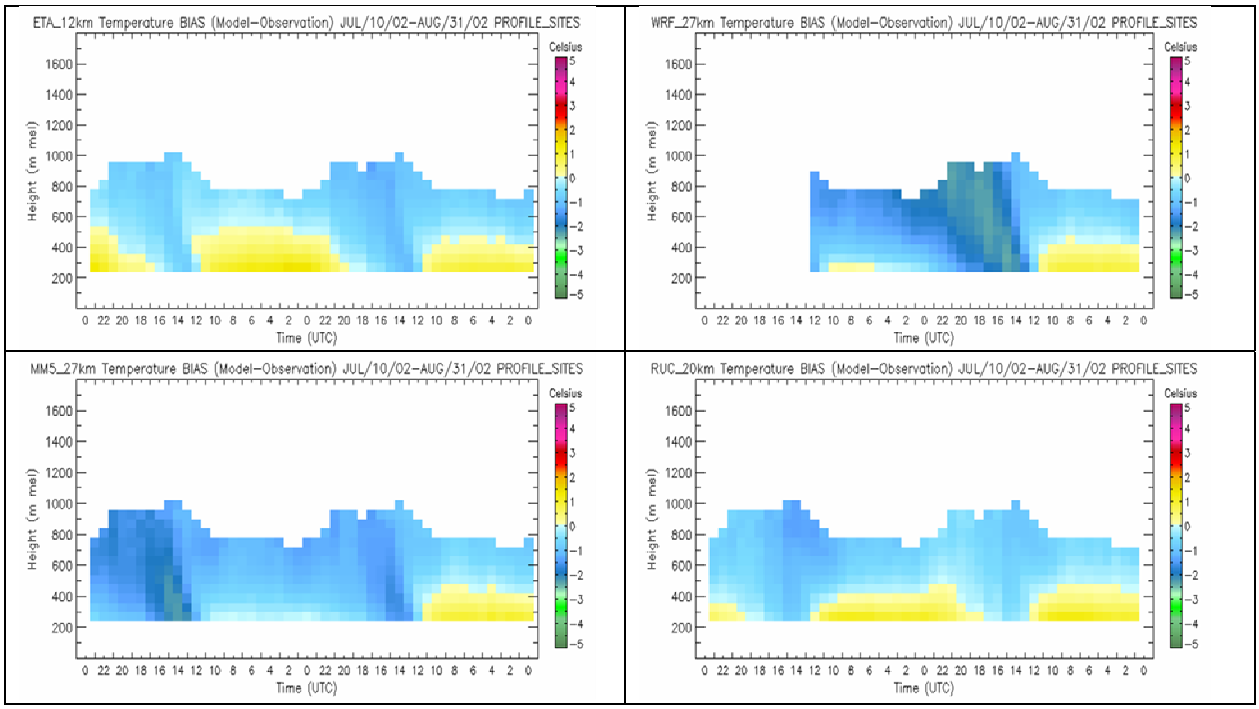

Figure 5. Virtual temperature biases for the for the 00 UTC model initialization cycle of the Eta, WRF, MM5, and RUC models, averaged for all 7 profiler sites for 53 days.

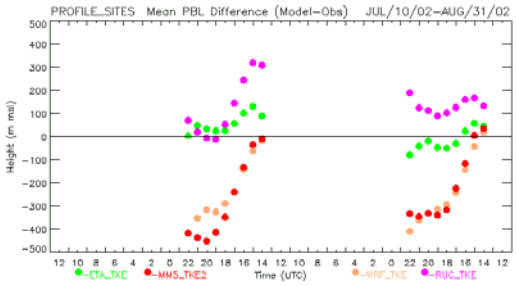

Figure 6. Boundary-layer depth biases for each model for the 00 UTC cycle.

Time-height cross-sections of virtual potential temperature (solid black lines), model TKE (color background), observed ABL depth (yellow dots), and model diagnosed ABL depth (solid black dots) are shown in Fig. 7 for one set of model simulations for the Eta, WRF, and RUC models. Although these are only examples from a single simulation, they are nevertheless informative. The Eta model's ABL predictions follow the observed values quite well, as was also found in the statistics. The WRF's predicted ABL depths are considerably too small, also as was found in the statistics. We note that the TKE values within the ABL are only about half as large in the WRF as in the Eta. The RUC's ABL depth predictions follow the observed values well for this case, which is also true for the RUC's ABL statistics. TKE values in the RUC are similar to those in WRF. We also note that the RUC tends to have higher levels of TKE in the surface layer than do the other models. This can be seen in Fig. 7 not only during the daytime hours, but especially during the nighttime hours on day 2 of the RUC simulation, when a low-level well-mixed layer exists. This is a phenomenon that has been found to occur frequently in the RUC, but not in the other models. Finally in the last panel of Fig. 7 the TKE and ABL depths for MM5 are shown. The model underestimates the ABL depth on these two days, as found in the mean statistics, despite the fact its TKE values come closer to the larger values found in the Eta.

In Fig. 8 we show the solar radiation biases for the 4 models. The Eta has the largest bias, approaching 200 Wm^{-2}. This bias was previously known at NCEP and was due to the parameterization of low-level clouds in the Eta model. In 2003 a change to the code was made that has significantly reduced the magnitude of this bias. MM5 also has a positive but smaller bias, while the WRF and RUC models had the smallest overall bias. The solar radiation biases among the four models do not show any strong correlation with their ABL depth biases.

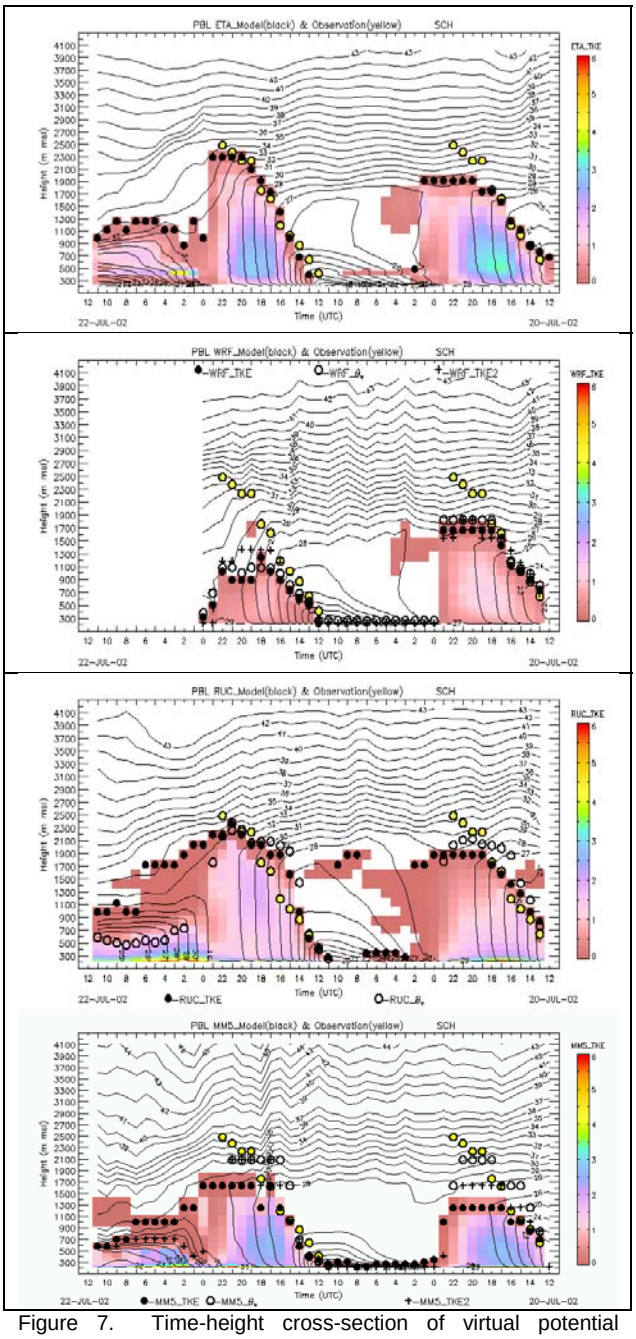

Figure 7. Time-height cross-section of virtual potential temperature (black lines), model TKE (color shading), observed ABL depth (yellow dots), and TKE-based model ABL depth (solid black dots). The simulations are for 12 UTC, 20 July 2002, for the Eta, WRF, RUC, and MM5 models, respectively.

5. SUMMARY

Wind profiler observations have been used to determine bias errors in four different operational or research models. Some of the models have been found to have difficulty in treating the nocturnal boundary layer's velocity structure.

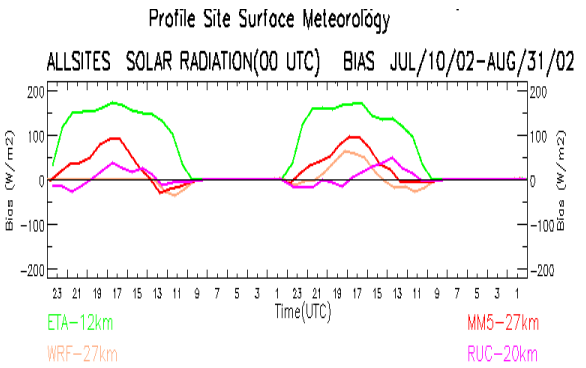

Figure 8. Shortwave radiation biases for the 00 UTC cycle of the Eta, WRF, MM5, and RUC models.

Diurnal temperature biases of varying magnitude are found in all of the models, and some have significant cooling trends as well. The ability of the models to correctly diagnose the boundary layer depth varies significantly, with three of the models under-predicting the ABL depth by as much as 35%. Solar radiation biases have also been found to be significant in several of the models.

REFERENCES

Angevine, W. M., A. B. White, and S. K. Avery, 1994: Boundary-layer depth and entrainment zone characterization with a boundary-layer profiler, *Boundary-Layer meteorol.*, **68,** 375-385.

Bianco, L., and Wilczak, J. M., 2002: Convective boundary layer depth: Improved measurement by Doppler radar wind profiler using fuzzy logic methods, *J. Atmos. Oceanic Technol.*, **19**, 1745–1758.

Weber, B.L., Wuertz, D.B., Law, D.C., Frisch, A.S., Brown, J.M.. 1992: Effects of small-scale vertical motion on radar measurements of wind and temperature profiles. *J. Atmos. Oceanic Technol.*, **9**, No. 3, pp. 193–209.

White, A. B., 1993: Mixing depth detection using 915-MHz radar reflectivity data, *Preprints, Eighth Symp. On Observations and Instrumentation*, Anaheim, Califonia, Amer. Meteor. Soc., 248-250.

Wilczak, J. M., R. G. Strauch, F. M. Ralph, B. L. Weber, D. A. Merritt, J. R. Jordan, D. E. Wolfe, L. K. Lewis, D. B. Wuertz, J. E. Gaynor, S. A. McLaughlin, R. R. Rogers, A. C. Riddle, and T. S. Dye, 1995: Contamination of wind profiler data by migrating birds: characteristics of corrupted data and potential solutions, *J. Atmos. Oceanic Technol.*, **12**, 449-467.