10.4 INITIAL EVALUATION RESULTS OF THE ETA, NMM, GFS, SREF, AND RUC MODELS DURING THE 2003 NEW ENGLAND HIGH RESOLUTION TEMPERATURE PROGRAM

James Wilczak*, Robert Zamora, Jian-Wen Bao, Irina Djaloalova
Environmental Technology Laboratory, NOAA

Jeff McQueen, Brad Ferrier, Zavisa Janjic, Hua-Lu Pan, Ken Mitchell, Geoff DiMego, Jun Du, Binbin Zhou
Environmental Modeling Center/NCEP/NOAA

Stan Benjamin
Forecast Systems Laboratory, NOAA

1. INTRODUCTION

During the summer of 2003 several NOAA Research Laboratories (Environmental Technology Laboratory, Forecast Systems Laboratory, and National Severe Storms Laboratory) collaborated with the National Weather Service (NWS) including the National Centers for Environmental Prediction (NCEP) on a program to improve surface temperature forecasts. This program used a regional observational and modeling testbed approach, focusing on the New England region. The motivation for this project was the recognition that the annual cost of electricity on a nationwide basis could decrease by $1 billion per year if the accuracy of surface temperature forecasts improved by 0.6 degree Celsius (NOAA/NWS Science and Technology Infusion report, 2003). The specific objectives of the New England High Resolution Temperature Program (NEHRTP) program are:

1) To quantify improvements in the forecasting of temperature in the New England region which result from new and augmented observations and modeling.
2) To assess the benefits of better predictive capabilities to the energy sector.
3) To provide a pathway to operational high-resolution temperature forecasting.

Because temperature at ground level is strongly influenced by the structure of the atmospheric boundary layer (ABL) and observations of the ABL are very limited, ETL deployed a network of 915 MHz wind profilers with RASS (Radio Acoustic Sounding System) to measure temperature profiles, ceilometers to measure cloud base, an instrument to measure aerosol optical depth, and surface meteorology including radiation (Fig. 1). Measurements from the wind profilers/RASS and surface met instruments were compared, on an hourly basis in real-time during the experiment, against 48h predictions made by operational and research forecast models, including the Eta, GFS, RUC, and NMM models, as well as several experimental versions of the SREF (Short Range Ensemble Forecast) model that were under development at NCEP. In addition, surface temperature, dewpoint, and winds were evaluated at 8 NWS operational surface meteorological sites (Fig. 1). Meteorological forecasts at these 8 "Energy Sites" are routinely used by the energy industry to predict short-term (0-48h) energy requirements in New England. The meteorological comparisons were made during a 72 day experimental period which ran from 1 July to 10 Sept, 2003. The real-time comparisons were made available through a web-site display, and can be found at http://www.etl.noaa.gov/programs/2003/nehrtp/verification

In this paper, we present model error statistics with highlighted discussions on the relation of surface temperature errors and radiation errors in the models.

2. SURFACE TEMP AND WIND STATISTICS

Wind speed, direction, temperature and dewpoint bias and RMSE statistics have been computed over the 72 day NEHRTP experiment at each of the 8 "Energy Sites" indicated in Fig. 1, and then averaged (Fig. 2). The bias is defined as the model minus observed value, and the RMSE values are bias adjusted. These statistics have been computed using the 00 UTC cycle, 48 h forecasts of the Eta, RUC, and GFS models, and are compared against the forecast NGM-MOS values.

Wind speed biases are found to follow the same diurnal pattern for all 4 models, with the maximum bias at night and minimum or negative biases during the daytime hours (Fig. 2). The NGM-MOS has a small positive mean bias, while the Eta and GFS are near zero. The RUC has the largest bias, under-predicting the mean wind speed by about 0.5 ms^{-1} on average. Wind direction biases show much less diurnal variation in all models. The mean

Corresponding author address: James Wilczak, NOAA/ETL, Bounder CO 80305. email: james.m.wilczak@noaa.gov

bias is largest for the Eta, but still only about 6 degrees.

The temperature bias is smallest for the NGM-MOS, and is only about -0.25 C. Although the NGM has coarser resolution and is older than the other models, MOS corrections are able to reduce its bias to small values. In contrast the RUC has a positive overall bias of approximately 0.5 C, while the Eta and the GFS have the largest mean biases of -0.75 and -1.1 C respectively. Of the 4 models, only the Eta has a pronounced diurnal variation to the temperature bias, with the model being slightly too warm during the daytime, but too cold during the nighttime hours, by as much as 1.5 C.

For dewpoint temperature, the RUC, Eta, and GFS all show pronounced diurnal variations. Both the Eta and GFS tend to be too dry at night and too moist during the day, while the RUC is the reverse, being too moist at night.

Bias adjusted RMS errors averaged for the 8 Energy Sites are shown in Fig. 3. For wind speed and direction all 4 models have very similar values. For temperature and dewpoint the NGM-MOS has the smallest RMSE, closely followed by the Eta and GFS, while the RUC has a significantly larger RMSE. For all models and variables little diurnal variation is seen, except for temperature where the RMSE is maximum in the late afternoon (~19UTC, 15EST). The RUC also shows some diurnal variation in dewpoint RMSE, which again is largest in the late afternoon hours.

3. RADIATION EVALUATION

Four-stream radiation measurements (short-wave up and down, long-wave up and down) were taken during NEHRTP at the Concord, MA wind profiler site, and in addition, short-wave down and net radiation were measured at each of the wind profiler sites. At the Concord, MA site the global solar irradiance was determined as the sum of the diffuse downward solar irradiance measured using a shaded black and white pyranometer and the direct downward solar component measured using an Eppley normal incidence pyroheliometer. The upward and downward longwave components were measured using conditioned Eppley precision pyrgeometers. The Concord site pyranometers and pyrgeometers were calibrated at the NOAA/CMDL Solar Calibration facility, which is a WMO World Region IV Center.

The four-stream radiation measurements were available for 48 consecutive days between 25 July and 10 Sept. Figure 4 shows the biases for each of the 4 radiation components for the Eta, GFS, and RUC models. Downward fluxes are defined to be positive and upward fluxes from the surface are negative.

For the Eta model, the largest radiation error is for the short-wave down (incoming solar irradiance, or SW-down). For the first day of the forecast period the error averages about 50 Wm^{-2} during the daytime hours (EST = UTC - 4h), increasing to 75-80 Wm^{-2} for day 2. The over-prediction of the solar irradiance in the Eta is part of the reason for the large diurnal variation in the temperature bias of the Eta model that was shown in Fig. 2. We note that as the SW-down bias increases from day 1 to day 2, so does the warm temperature bias.

The over-prediction of SW-down irradiance in the Eta model was previously known at NCEP, and is likely caused by several parameterization weaknesses. The most important of these is that parameterized shallow (non-precipitating) convection is invisible to the radiation scheme. A simple fix to partially address this problem was incorporated in the operational Eta model at the start of the NEHRTP field program. This fix was to assume a low-level cloud fraction of 10% which was parameterized with simple assumptions for the cloud properties of shallow cumuli. In contrast, during the previous summer, in a field program that preceded NEHRTP, a mean bias of almost 200 Wm^{-2} was found in the Eta. So the simple fix implemented in the Eta model has reduced the solar radiation bias in the model during the summer of 2003 to about half of magnitude found during the summer of 2002.

The short-wave up (SW-up) flux bias is also large, about 50 Wm^{-2} on both days of the forecast cycle (Fig. 4). A small part of this (~15 Wm^{-2}) is just the reflection of the too large downward short-wave, while most of it must be due to having a different local albedo from the grid-point albedo assumed in the model.

For the IR components, the long-wave down (LW-down) flux is under-predicted by the Eta during the nighttime hours by as much as 20 Wm^{-2}. This is also consistent with the hypothesis that the model under-predicts the effects of clouds on radiation, and it is consistent with the model being too cold and too dry at night. However, in this regard, we note that the long-wave down flux bias tends to slowly increase during the course of the night and is greatest near sunrise (9 UTC), which suggests that the LW-down bias may as much be the result of the cold (and dry) model biases as the cause of the cold bias.

Finally for the Eta model we find that the LW-up flux is too small (a positive bias). This is consistent with the overall cold bias of the Eta, assuming that the skin temperature and 2m temperature biases are of the same sign. However, no obvious diurnal variation is present in the LW-up flux, while there is in the 2m temperature.

Similar to the Eta, the GFS is found to significantly over-predict the SW-down flux, by about 40 Wm^{-2} on day 1, increasing to 80 Wm^{-2} on day 2 (Fig. 4). However, in contrast to the Eta the SW-up flux is too small (a positive bias), which again must be due to differences in the local and model albedo.

The GFS LW-down bias is negative throughout the 48 h forecasts period, and is larger during the day (-25 Wm^{-2}) than at night. As for the Eta, the negative LW-down bias is consistent with the observed cold bias of the GFS (Fig. 2), but it is not possible to say if this the LW-down bias is the cause of the cold bias or only a result of the cold bias.

The GFS LW-up flux is significantly too small during the day (a positive bias), with the bias reaching almost 50 Wm^{-2}, and is slightly too large at night. Since the 2m temperature is consistently cold through the 48h forecast period, and since the LW-up flux is proportional to $sT_s^{\ 4}$, where T_s is the skin temperature, this means that the biases in the skin temperature and the 2m temperature diverge greatly over the diurnal cycle and become opposite sign from one another during the nighttime hours.

The RUC radiation biases are significantly different from both the Eta and the GFS (Fig. 4). First, the SW-down flux is too small, especially in day 2, with the bias reaching -75 to -100 Wm^{-2}. Second, although the LW-down bias is negative during the daytime (as it was for the GFS), the 2m temperature bias is warm, rather than cold. The LW-down bias in the RUC (too small of a flux) therefore cannot be the cause of the warm RUC bias, nor can the temperature bias (too warm) be the cause of the LW-down bias.

Finally, the RUC has a very large (+70 Wm^{-2}) daytime LW-up bias (too small of a flux), with a smaller bias at night. The too small upward LW flux requires a cold skin temperature bias, but in Fig. 2 it has been shown that the RUC has a warm 2m temperature bias. So again, as for the GFS, the skin and 2m temperature biases must diverge significantly and be of opposite signs.

The Eta, GFS, and RUC have all been found to have significant solar radiation biases. One might naively expect the solar radiation biases to induce a positively correlated surface temperature bias. In fact the GFS has a cold bias, day and night with little diurnal variation, despite having a positive SW-down bias. Similarly, the RUC has a warm bias, day and night with little diurnal variation, despite having a negative SW-down bias. Clearly, compensating errors exist in these models that negate the solar irradiance bias. These compensating errors most likely are present in the land-surface models (including the turbulent surface fluxes), and perhaps in the model numerics.

For the Eta model we do find a positive correlation between the solar radiation bias and the diurnal variation in the temperature bias. However, the overall bias (day and night) of the Eta is still cold, and reducing the SW-down bias will likely only exacerbate the overall cold bias. Therefore, for all three models, the Eta, GFS, and RUC, we believe that reducing the SW-down biases without at the same time reducing compensating errors in other parts of the model will almost certainly only increase the surface temperature biases.

3. SREF AND NMM MODELS

Since neither the SREF nor the NMM had 00 UTC model cycles (the SREF was initialized at 06 and 18 UTC, and the NMM at 18 UTC) they were not used in the previous analysis. Instead, we show here representative examples of surface temperature behavior in both models.

Three experimental versions of the SREF were run during the NEHRTP model evaluation study (Jun Du., 2004). The first (SREF-I) was termed the "Breeding Experiment" and consisted of 5 members each of the Eta-BMJ, 5 members of Eta-KF, and 5 members of the RSM-SAS models. The 5 members of each group consisted of a control and 2 breeding pairs. The SREF-II model was termed the "Physics/Breeding Experiment", and again consisted of 5 RSM and 10 Eta runs, with additional convective parameterization and cloudy physics diversity but fewer breeding pairs (Ferrier, 2004). The SREF-III model consisted of only Eta simulations, again using a combination of physics diversity and breeding pairs. All three SREF models were run at 32 km horizontal resolution.

Figure 5 shows a representative example of the surface meteorology for the three SREF experimental models, together with the operational Eta model. Both SREF-I and SREF-II are similar, and they significantly under-predict the observed diurnal variation of the surface temperature. The warmer nighttime temperatures of the SREF-I and SREF-II are an improvement over the Eta, but the daytime temperatures are much colder than the observations and the Eta. The SREF-III, which consists only of Eta members but with a large diversity of model physics parameterizations, more closely follows the operational Eta than do either the SREF-I or SREF-II. In addition, the SREF-III gives slightly cooler daytime temperatures and warmer nighttime temperatures than the operational Eta, closer to the observations. Although the SREF-II under-predicts the daytime temperatures, it does show other benefits in terms of model spread (Jun Du, 2004) and error statistics aloft, and so a slightly modified version of the SREF-II (with increased physics diversity) will replace the current operational SREF in 2004.

The developmental NMM model in general is found to provide surface temperatures with accuracies similar to the other models. However, occasional periods occur during which the NMM accentuates the nighttime cold bias found for the Eta. An example of one such case is shown in Fig. 6. Although the forecast daytime temperatures are highly accurate, the first nights forecast temperatures are too low (and similar to the Eta) while for the second night the NMM cold bias is even larger (and colder than the Eta's).

4. SUMMARY

Several operational and experimental forecast models have been evaluated during the summer 2003 New England High Resolution Temperature Program (NEHRTP). The purpose of this program is to improve operational forecasts of surface temperature, and if successful, has the potential to significantly reduce energy costs. Because errors in the model irradiance fields can directly produce surface temperature errors, the 4 stream (SW up and down, LW up and down) radiation components were key parameters in the model evaluations. Significant SW-down (solar irradiance) biases were found to be present in the Eta, GFS, and RUC models, ranging from 50 to 100 Wm^{-2}. However, these radiation biases have the opposite sign of the 2m temperature biases in the GFS and RUC, indicating that other model errors (probably in the land-surface specification or parameterization) compensate for the irradiance errors. For the Eta model the positive SW-down bias is correlated with a slight daytime warm bias, but the model has an overall (day and night) cold bias, so correcting the SW-bias alone will likely amplify this cold bias. Therefore, for all three models, the Eta, GFS, and RUC, we believe that reducing the SW-down biases without at the same time reducing compensating errors in other parts of the model will almost certainly only increase the surface temperature biases. For the SREF model, an experimental version (SREF-II) tested during NEHRTP was selected to become the new operational SREF in late 2003, based on its improved model spread characteristics. However, initial evaluation of this model based on case studies indicates that it has a large daytime cold bias relative to an Eta-only experimental version (SREF-III).

5. REFERENCES

Ferrier, B.S., 2004: Modifications to Two Convective Schemes used in the NCEP Eta Model. 16th Conference on Numerical Weather Prediction, Seattle, WA, Jan. 11-15, 2004.

Jun Du., 2004: Impact of model error and imperfect IC perturbation on evolution of ensemble based PDF's in NWP model. 17th Conference on Probability and Statistics in the Atmospheric Sciences, AMS (2004).

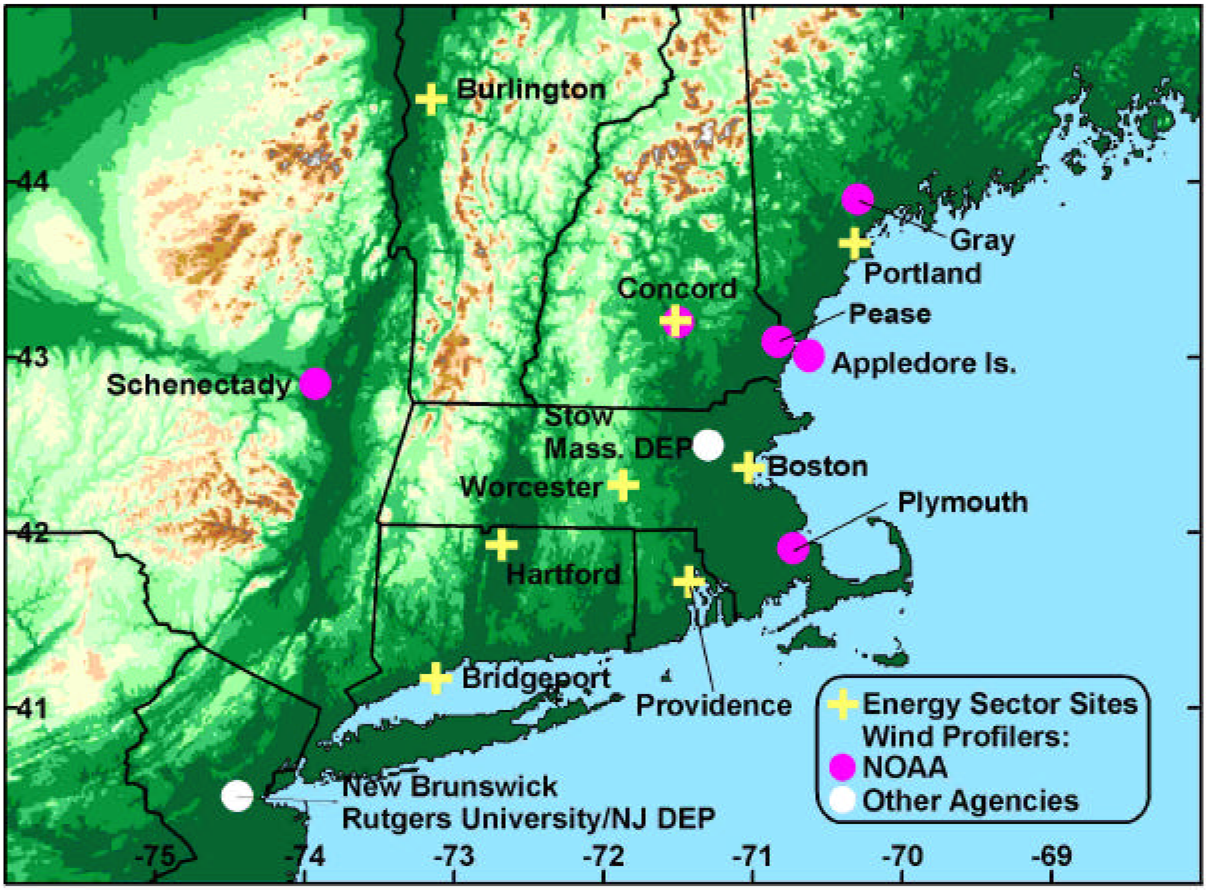

Figure 1. Base map of the New England High Resolution Temperature Program field experiment. Solid circles show locations of wind profiler/RASS sites with surface meteorology. Plus symbols indicate locations of the 8 Energy Sites, which are NWS Metar stations used by the energy industry.

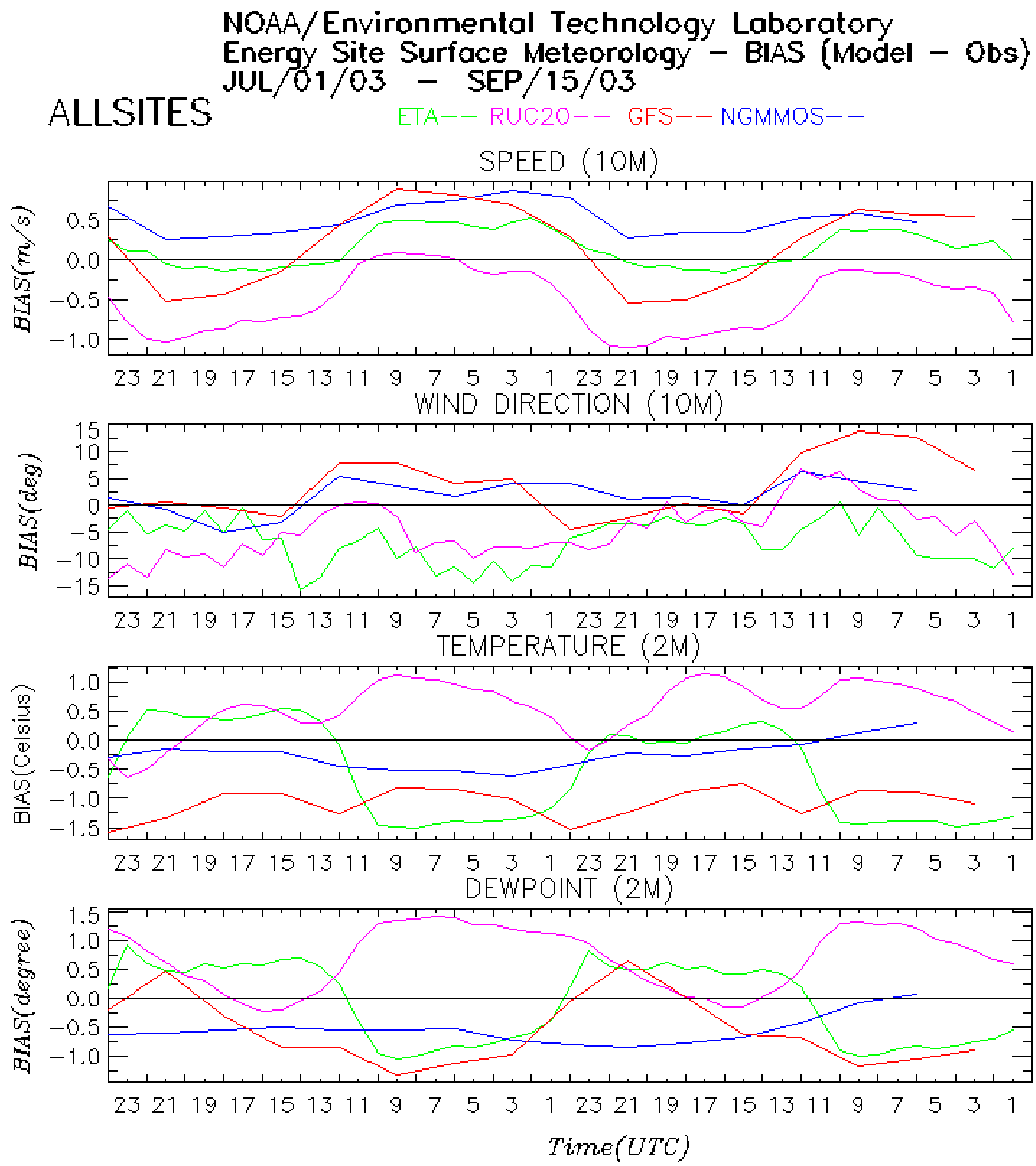

Figure 2. 10-m wind speed and direction and 2-m temperature and dewpoint biases for the Eta, RUC, GFS, and NMG-MOS models averaged at the 8 Energy Sites over the 72 day field study.

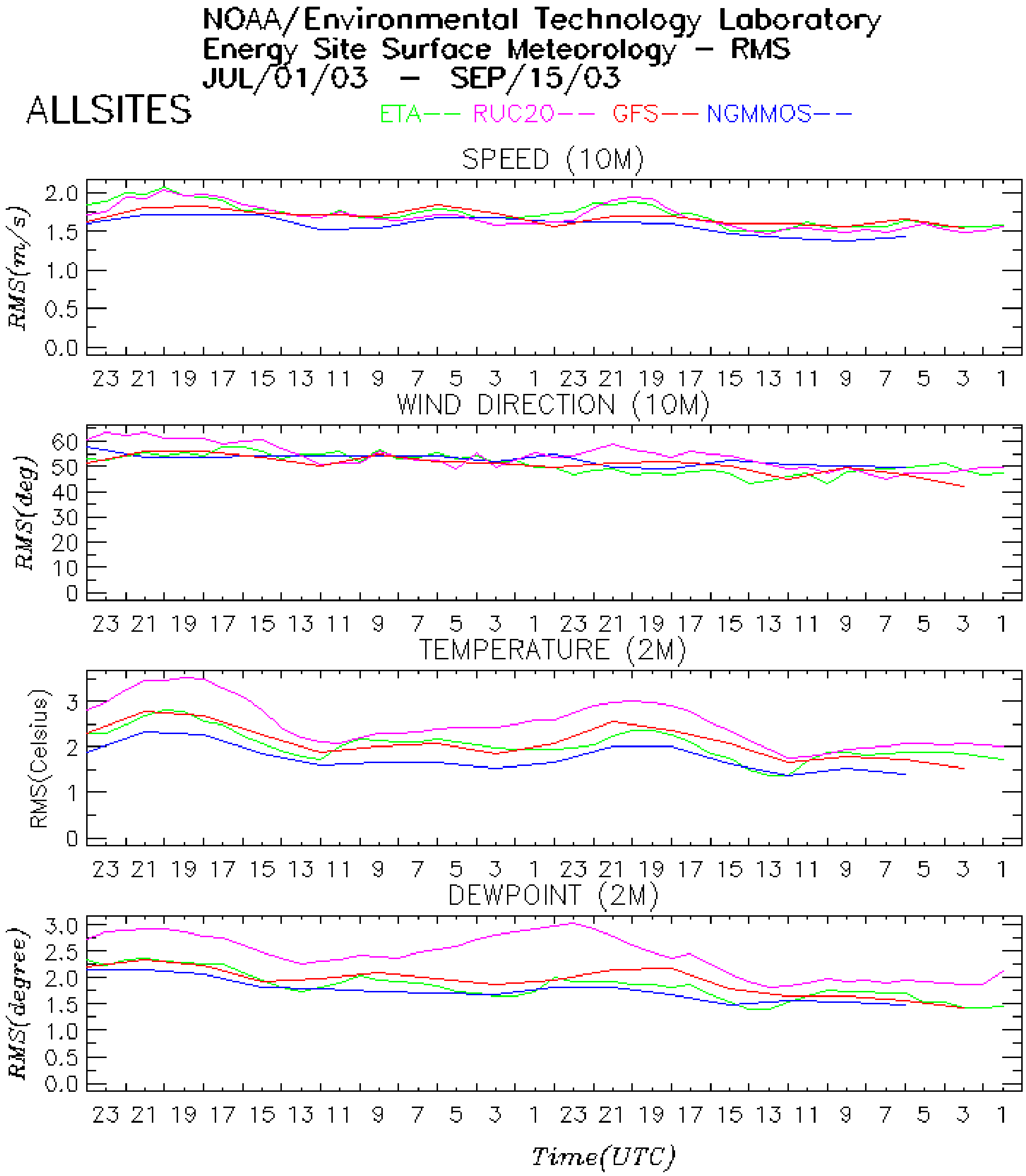

Figure 3. As in Fig. 2, except the RMS model errors (calculated after having removed the mean bias).

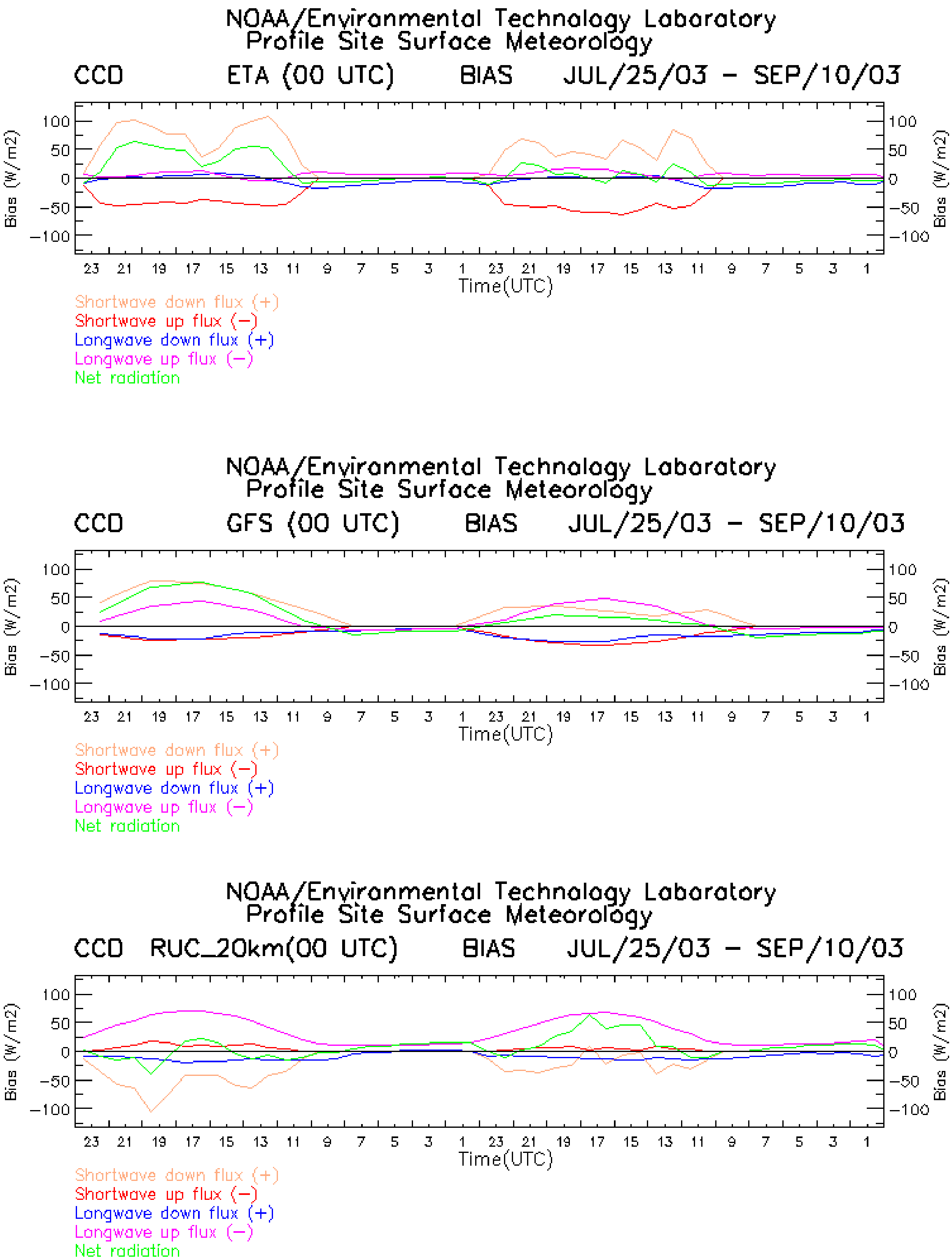

Figure 4. Radiation biases for the Eta, GFS, and RUC models at the Concord, MA site. Short-wave and long-wave downward fluxes are defined to be positive, while upward fluxes are defined to be negative in the calculation of the bias.

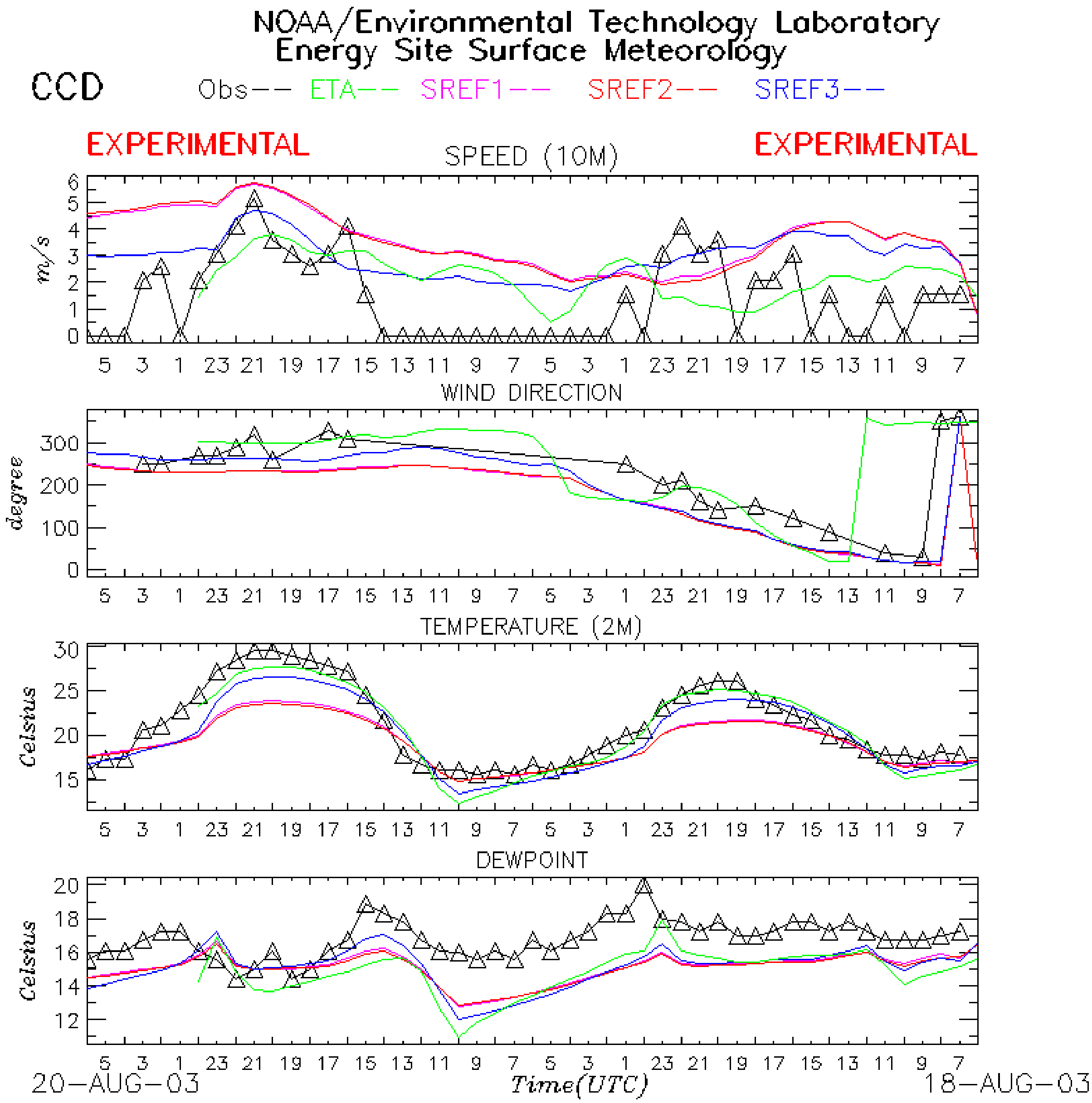

Figure 5. A representative example of surface forecasts for the three SREF models. SREF-I and SREF-II contain RSM model members, SREF-III contains only Eta model members.

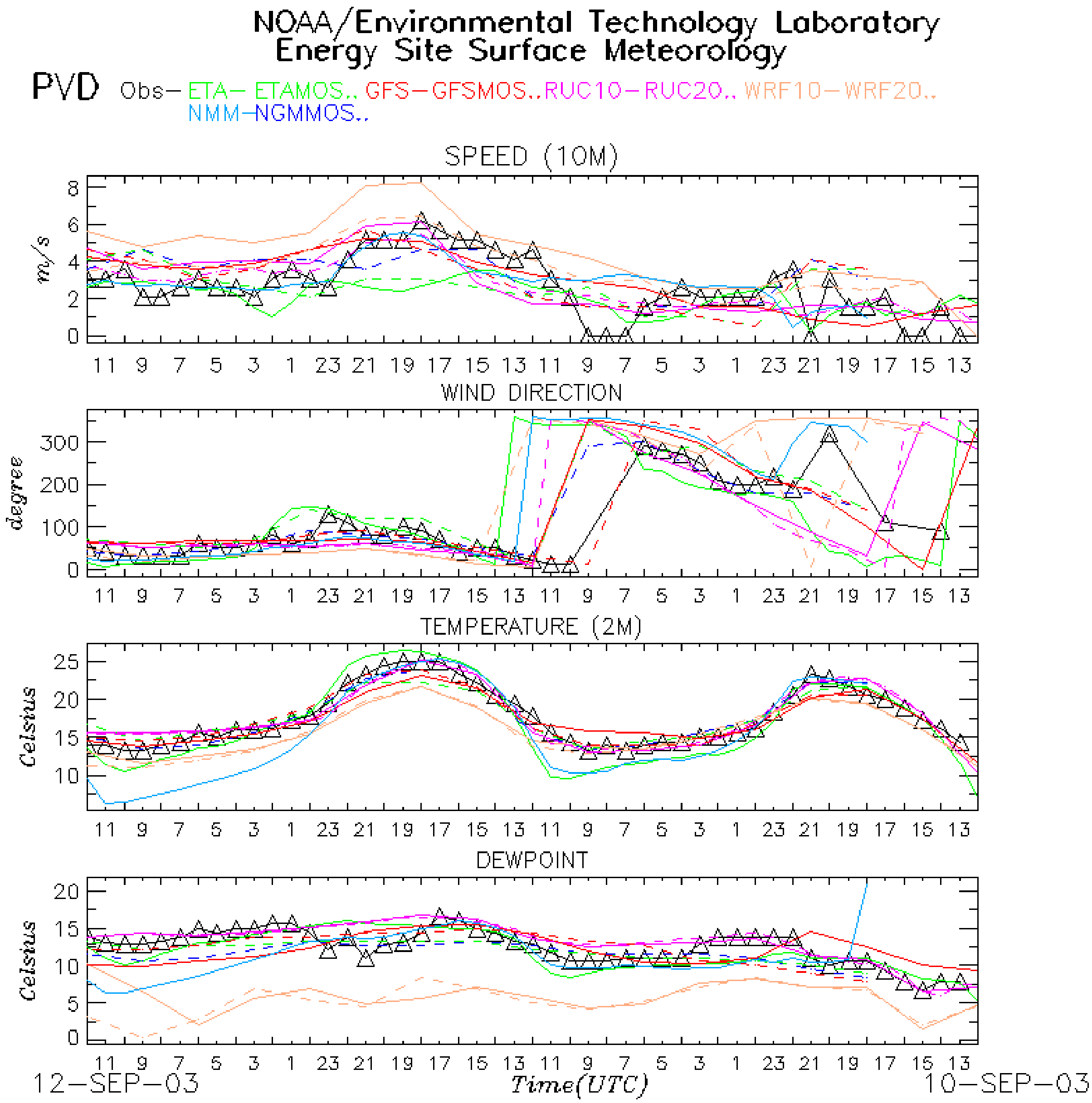

Figure 6. An example of the NMM model (light blue solid line) predicting too cold of 2-m temperatures.