EVALUATION OF SURFACE SENSIBLE WEATHER FORECASTS BY THE WRF AND THE ETA MODELS OVER THE WESTERN UNITED STATES

William Y.Y. Cheng[1] and W. James Steenburgh

Department of Meteorology and NOAA Cooperative Institute for Regional Prediction, University of Utah, Salt Lake City, Utah

1. INTRODUCTION

The Cooperative Institute for Regional Prediction (CIRP) at University of Utah ran a beta version of the WRF model (CIRP WRF) from March 2003 to September 2004 in real time. The CIRP WRF domain had an area of 4 410 000 km^2 with 169 by 169 horizontal grid points at 12.5 km grid spacing, covering the entire western United States and portions of Mexico and Canada (Fig. 1). 34 half-η levels were used.

This study aims to help model developers by evaluating the ability of the WRF model to predict warm season surface sensible weather over the western United States. CIRP WRF's performance is also compared to that of the Eta model, the operational mesoscale model developed by the National Center for Environmental Prediction (NCEP). Surface sensible weather variables (i.e., 2-m temperature, 2-m dewpoint, and 10-m wind) were evaluated because of their importance in power consumption and fire weather prediction, as well as the availability of high-density surface observations provided by the MesoWest cooperative networks (Horel et al. 2002).

Figure 1 Realtime CIRP WRF domain.

Corresponding author address: William Y.Y. Cheng, Department of Meteorology and NOAA Cooperative Institute for Regional Prediction, University of Utah, Salt Lake City, UT 84112. Email: wcheng@met.utah.edu

2. METHODOLOGY

Forecast cycles were selected for verification only when both the CIRP WRF grids and the corresponding Eta grids were available continuously from forecast hours 0 to 48. Each forecast cycle was verified at three-hourly intervals. Due to local hardware failure or missing grids, this represented 119 out of a possible 184 cycles. MesoWest stations within the CIRP WRF domain that reported at least 50% of the possible observations in the three month period provided the verification data. A total of 1875 MesoWest stations met these criteria. MesoWest obtains data from a number of networks, each designed to meet the specific needs of its operating agency. As a result, there is considerable diversity in the site characteristics, sensor types and heights, and reporting intervals of the observations. Because of a lack of station metadata, no effort was undertaken to account for variability in sensor height, and it was assumed that large contrasts between the model and observations reflected the characteristics of the model more than the verifying dataset.

To perform the verification, 2-m temperature, 2-m dewpoint, and 10-m wind forecasts from the CIRP WRF and the Eta models were first bilinearly interpolated to the MesoWest station locations. Note that forecast/observation pairs for wind were excluded for observed winds of 2.5 m s^{-1} because anemometers are not very accurate at low wind speed. The number of observation/forecast pairs for temperature, dewpoint, and wind was 3 587 164, 2 658 297, and 934 844, respectively.

3. RESULTS

3.1. 2-m temperature

The three-month cumulative temperature MAE for CIRP WRF (the Eta) was 3.3^{o}C (2.8^{o}C) for the 0000 UTC cycle and 3.1^{o}C (2.8^{o}C) for the 1200 UTC cycle (not shown). For 0000 UTC cycle, as a function of forecast hour, the CIRP WRF and the Eta temperature MAE exhibited maxima at forecast hours 12 and 36, which correspond to early morning (1200 UTC), and forecast hours 24 and 48, which correspond to late afternoon (0000 UTC) (Fig. 2). Maxima in temperature MAE for CIRP WRF were particularly pronounced in the early morning when they exceeded 4^{o}C and were 1^{o}C greater than those of the Eta.

For the 1200 UTC cycle, the CIRP WRF temperature

MAE peaked near 3.7^{o}C at forecast hours 12 and 36, which corresponds to late afternoon (0000 UTC) (Fig. 2). This contrasts with the 0000 UTC CIRP WRF which featured peak temperature MAEs in the morning. The 1200 UTC cycle Eta temperature MAE featured weak maxima in late afternoon (forecast hours 12 and 36) and early morning (forecast hours 24 and 48), similar to the 0000 UTC cycle, with some evidence of error growth with increasing forecast projection. During some forecast hours, the Eta MAE was at least 0.5^{o}C lower than that of WRF.

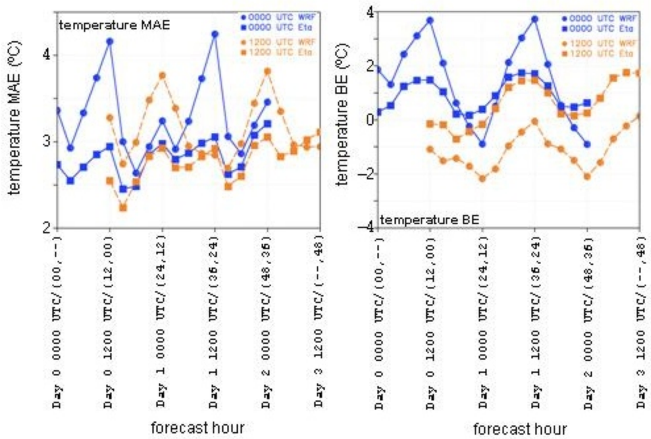

Figure 2 Average temperature MAE (^{o}C, left panel) and BE (^{o}C, right panel) as a function of forecast hour (UTC) for the 0000 UTC (blue curves) and 1200 UTC (orange curves) cycles. Circles (squares) represent results from CIRP WRF (the Eta). Note that the 1200 UTC cycle results are plotted at a 12-h lag relative to the 0000 UTC results. The first (second) number in the parentheses in the abscissa corresponds to the forecast hour of the 0000 (1200) UTC cycle forecast.

The bias analysis provides insight into the MAE characteristics of the two models. CIRP WRF three-month cumulative temperature bias varied from 1.5^{o}C for the 0000 UTC cycle to -1.2^{o}C for the 1200 UTC cycle (not shown). In contrast, the Eta model exhibited a weak warm bias in both forecast cycles (0.9^{o} and 0.6^{o}C for 0000 UTC and 1200 UTC cycles, respectively).

Both the CIRP WRF and the Eta temperature biases varied diurnally, with maximum warm bias in the morning hours [forecast hours 12 and 36 (24 and 48) for the 0000 (1200) UTC cycle] (Fig. 2). The amplitude of the diurnal variability in the bias error was particularly pronounced for the 0000 UTC CIRP WRF cycle.

The temperature bias in CIRP WRF can be explained partly by the procedure of the soil temperature initialization in the WRF slab soil model. The slab model was designed to handle global model grids with only two layers in the land surface model (Jimy Dudhia, personal communication, 2004). For example, the NCEP global reanalysis has two layers in its land surface model (0-10 cm and 10-200 cm). By default, the soil reservoir temperature from the slab model (centered at 25 cm below ground) is initialized with the top layer soil temperature (the only layer closest to the reservoir depth). The WRF developers have not yet accounted for instances where soil temperature data are available in more than two layers as with our Eta model soil temperature initialization. This caused a bias in the reservoir temperature initialization, and this is reflected in the initial warm (cold) 2-m temperature bias in the 0000 (1200) UTC cycle. Note the shift in the peak of the 2-m temperature bias in the WRF 0000 (1200) UTC cycle to the positive (negative) side of the ordinate due to the bias in the soil temperature initialization (Fig. 2). However, the shape of the bias time series is the same for both CIRP WRF and the Eta regardless of the initialization time. An erroneous soil temperature initialization can give rise to an erroneous surface sensible heat flux, leading to an erroneous 2-m temperature. We will discuss later how a better soil temperature initialization can improve the CIRP WRF forecast.

3.2. 2-m dewpoint

The three-month cumulative dewpoint MAE for CIRP WRF (the Eta) was 3.6^{o}C (3.3^{o}C) for the 0000 UTC cycle and 3.6^{o}C (3.4^{o}C) for the 1200 UTC cycle (not shown). As a function of forecast hour, the dewpoint MAE followed a diurnal pattern for both models (Fig. 3). For CIRP WRF, excluding the large initial MAE in the 1200 UTC cycle, the dewpoint MAE maxima (3.8-4.1^{o}C) generally occurred in the nighttime to morning period from 0300 to 1200 UTC (in the morning at 1200 UTC) for the 0000 (1200) UTC cycle, with MAE minima (3-3.5^{o}C) in the morning to late morning period from 1500 to 1800 UTC for both the 0000 and 1200 UTC cycles. The large initial dewpoint MAE in the 1200 UTC cycle for CIRP WRF will be explained later. As for the Eta model, the dewpoint MAE maxima generally occurred during the 0000 to 1200 UTC window for both forecast cycles, and the minima occurred during the 1500 to 1800 UTC window (at 1500 UTC) for the 0000 (1200) UTC cycle.

CIRP WRF (the Eta) featured a three-month cumulative dewpoint bias of 1.1^{o}C (-0.8^{o}C) for the 0000 UTC cycle and -0.8^{o}C (-0.9^{o}C) for the 1200 UTC cycle (not shown). Except for the CIRP WRF 0000 UTC cycle, as a function of forecast hour, both models were drier than observed, with the dry bias being more pronounced in the Eta model (Fig. 3). In particular, the CIRP WRF 1200 UTC cycle had an initial 2-m dewpoint bias of -3^{o}C, much drier than the Eta.

Consistent with the MAE analysis, the CIRP WRF and the Eta dewpoint biases also exhibited a diurnal pattern. For CIRP WRF, in terms of magnitude, dewpoint bias maxima generally occurred in the 0600 to 1200 UTC window and at 1200 UTC for the CIRP WRF 0000 and 1200 UTC cycles, respectively. As for the Eta, in terms of magnitude, dewpoint bias maxima generally occurred at 1200 UTC for the Eta for both 0000 and 1200 UTC cycles.

The large initial dewpoint BE and MAE in the CIRP WRF

1200 UTC cycle may be partly attributed to the erroneous soil temperature initialization, although it is possible for other factors to contribute to the problem such as atmospheric initialization procedures and the method of calculating the 2-m water vapor mixing ratio (Fig. 3). Note the correspondence between the initial negative 2-m temperature bias (partly a reflection of the erroneous ground temperature initialization) and the initial negative 2-m dewpoint bias in the CIRP WRF 1200 UTC cycle (Figs. 2 and 3). The correspondence between the initial positive dewpoint bias and the initial positive temperature bias can also be seen in the 0000 UTC CIRP WRF
cycle.

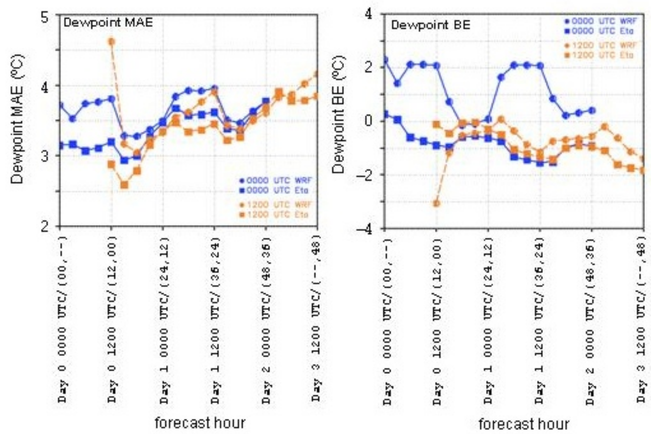

Figure 3 Same as Fig. 2 but for dewpoint.

3.3. 10-m wind

The three-month cumulative wind speed MAE for CIRP WRF (the Eta) was 1.7 (1.6) m s^{-1} for the 0000 UTC cycle and 1.8 (1.6) m s^{-1} for the 1200 UTC cycle (not shown). As a function of forecast hour, the 0000 UTC CIRP WRF wind speed MAE varied between 1.6 and 1.9 m s^{-1} with no regular diurnal variability (Fig. 4).

The 0000 UTC Eta model wind speed MAE showed a more regular diurnal pattern with maxima in the afternoon (forecast hours 24 and 48) and minima at night. This diurnal pattern also was evident for the 1200 UTC Eta cycle. Unlike the 0000 UTC cycle, the 1200 UTC CIRP WRF wind speed MAE showed some diurnal structure, with the largest errors during the late night or early morning hours (forecast hours 18-24 and 42-48).

In terms of bias, the three-month cumulative wind speed bias for CIRP WRF (the Eta) model was 0.5 (-0.4) m s^{-1} for the 0000 UTC cycle and 0.4 (-0.5) m s^{-1} for the 1200 UTC cycle (not shown). Thus, CIRP WRF tended to overpredict the wind speed, whereas the Eta model tended to underpredict the wind speed.

In both models, however, the wind speed bias varied according to the time of day (Fig. 4). CIRP WRF exhibited pronounced wind speed bias maxima at night that reached 0.8 - 1.2 m s^{-1}, but a relatively small bias (magnitude < 0.3 m s^{-1}) in the afternoon. The Eta model produced relatively small biases in the 2100 to 0000 UTC window (afternoon), but generated a large negative (often < -0.8 m s^{-1}) bias (i.e., underprediction) in the 0300 to 1500 UTC period (nighttime to early morning).

Wind direction is perhaps one of the most difficult variables to forecast. For both initialization times, the three-month cumulative wind direction MAE for CIRP WRF (the Eta) was 61^{o} (41^{o}) (not shown). Thus, the Eta model typically produced a more accurate wind direction forecast.

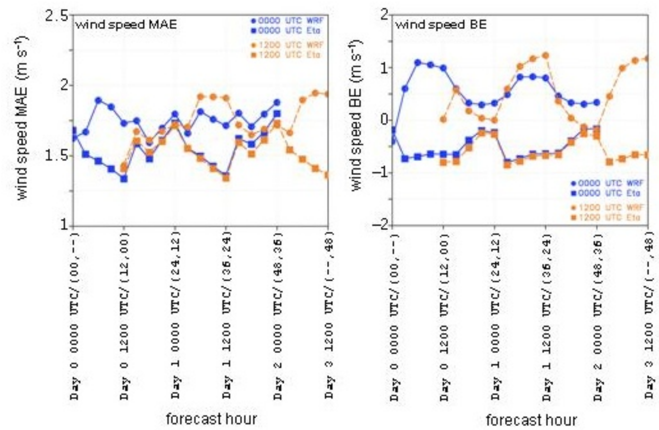

Figure 4 Same as Fig. 2 but for wind speed.

3.4. Sensitivity experiments

The CIRP WRF 2-m temperature bias can be reduced by better defining the initial temperature in the slab soil model. As mentioned previously, the default slab soil model used the top (0-10 cm) layer soil temperature from the Eta as the WRF reservoir temperature, which represents a temperature centered 23 cm below ground. This is clearly not appropriate, as the
0-10 cm soil temperature is subject to diurnal variations. In order to improve the soil temperature initialization in the slab soil model, the WRF code was reconfigured to use the Eta soil temperature at 10-40 cm below ground (i.e., centered 25 cm below ground) for initializing the CIRP WRF reservoir temperature. The 0000 UTC and 1200 UTC 1 July 2003 forecasts were re-run with this change (experiment FIXEDSLAB). Experiments were performed to examine whether using the more sophisticated Oregon State University (OSU) land surface model (LSM) would improve the WRF forecasts (experiment OSULSM). The soil moisture and temperature from the Eta model were used to initialize OSULSM. A control experiment (CTL) was also performed in which the standard configuration of the slab soil model was used.

Although FIXEDSLAB reduced the 2-m temperature MAE by 0.3^{o}C for the 0000 UTC cycle, it did not reduce the temperature MAE for the 1200 UTC cycle (Fig. 5). However, FIXEDSLAB reduced the warm (cold) bias in the 0000 (1200) UTC cycle compared with CTL from 1.3^{o}C to -0.1^{o}C (from -0.9^{o}C to -0.6^{o}C). OSULSM had larger temperature MAE than CTL, about 0.3-0.6^{o}C. In addition, OSULSM had mixed results in affecting the temperature BE. While OSULSM showed improvement in reducing the warm temperature bias in the 0000 UTC cycle (from 1.3^{o}C to -0.8^{o}C), the cold temperature bias became more negative in the 1200 UTC cycle (from -0.9^{o}C to -1.2^{o}C). The Eta model had better temperature MAE and BE than all of the WRF experiments except for the BE in FIXEDSLAB for the 0000 UTC cycle (Fig. 5).

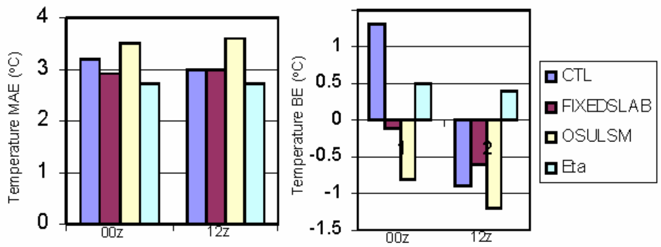

Figure 5 Cumulative temperature MAE (^{o}C, left panel) and BE (^{o}C, right panel) for 1 July 2003 for experiments CTL, FIXEDSLAB, OSULSM, and the Eta model.

4. DISCUSSION AND CONCLUSION

Overall, CIRP WRF produced comparable MAEs to the Eta for 2-m dewpoint and 10-m wind speed. The 2-m temperature forecasts produced by CIRP WRF were, however, much worse. For the CIRP WRF 0000 UTC cycle, the 2-m temperature forecast valid in the early morning (1200 UTC) tended to be too warm partly because of the inadequate development of the nocturnal inversion layer at many locations and partly because of the initial positive 2-m temperature bias. Hart et al. (2004) reported similar problems in the inadequate development of the nocturnal inversion layer with a version of the MM5 that was run over the Intermountain West for the 2002 Olympic Winter Games and featured surface and boundary layer parameterizations similar to those of CIRP WRF. In addition, the magnitude of the morning warm bias in the 0000 UTC CIRP WRF 2-m temperature forecasts has been exacerbated further by the erroneous soil temperature initialization (Fig. 2).

In contrast, the 1200 UTC CIRP WRF cycle featured a more pronounced late afternoon (0000 UTC) cold 2-m temperature bias and little bias in the morning. This shift in the bias characteristics reflects the initial negative temperature bias (-1^{o}C) in the 1200 UTC WRF cycle as opposed to the initial positive temperature bias (2^{o}C) in the 0000 UTC WRF cycle. As a result, the 2-m temperature in the 1200 UTC CIRP WRF cycle tended to be too cool.

From the results in the sensitivity experiments, the minor change in the initialization of the slab soil model reservoir temperature appears to improve the surface temperature forecast, especially in reducing the temperature BE, albeit less so in the MAE, but without adversely affecting the BE and MAE in the dewpoint and the wind. However, using the OSU LSM (experiment OSULSM) did not result in superior performance in 2-m temperature forecast as compared to FIXEDSLAB. The results in the sensitivity experiments are similar to the findings of Zhong and Fast (2003), who found decreased forecast accuracy in mesoscale simulations using more sophisticated LSMs. Some possible reasons for this are: i) the uncertainty in defining the parameters in LSMs such as thermal conductivity of the surface; ii) not accounting for subgrid-scale heterogeneities in the land surface; iii) lack of accurate soil temperature and moisture in initializing the LSM. Therefore, it is not surprising to find that forecast accuracy was not improved using the OSU LSM.

The Noah LSM, a more advanced version of the OSU LSM has been released recently in the latest version of WRF. The Noah LSM in WRF is more tightly coupled to the radiation and PBL schemes. Although this may improve WRF's performance, the analysis of the Eta model (which uses the Noah LSM) presented in this study suggests that improvement in LSM physics alone is insufficient. Improvements in LSM initialization may be equally as important as (or perhaps more important than) improvements in LSM physics. Yang et al. (1994) found that improvement in the initialization of a global model's soil wetness can reduce the five-day MAE of the surface temperature forecast. The modeling community should consider improving the LSM initialization and parameterization of coupled land surface/boundary layer processes as a top priority in order to generate better surface sensible weather forecasts.

5. ACKNOWLEDGMENT

This research was funded by the National Weather Service through a series of grants and the NOAA Office of Global Programs under Grant GC04-281. We also thank Prof. John Horel, Drs. Jason Knievel and Jason Nachamkin, and three anonymous reviewers for comments on the manuscript submitted to Weather and Forecasting on which this extended abstract is based.

6. REFERENCES

Hart, K. A.,W. J. Steenburgh, D. J. Onton, and A. J. Siffert, 2004: An evaluation of mesoscale model-based model output statistics (MOS) during the 2002 Olympic and Paralympic Winter Games. *Wea. Forecasting*, **19**, 200–218.

Horel, J., M. Splitt, L. Dunn, J. Pechmann, B. White, C. Ciliberti, S. Lazarus, J. Slemmer, D. Zaff, and J. Burks, 2002: Mesowest: Cooperative Mesonets in the Western United States. *Bull. Amer. Meteor. Soc.*, **83**, 211–225.

Yang, R., M. J. Fennessy, and J. Shukla, 1994: The influence of initial soil wetness on medium-range surface weather forecasts. Mon. Wea. Rev., 122, 471–485.

Zhong, S. Y., and J. Fast, 2003: An evaluation of the MM5, RAMS, and Meso-Eta models at subkilometer resolution using VTMX field campaign data in the Salt Lake Valley. *Mon. Wea. Rev.*, **131**, 1301–1322.