EVALUATION OF A SURFACE DATA ASSIMILATION TECHNIQUE USING THE MM5

Kiran Alapaty1*, Nelson Seaman2, Devdutta Niyogi3, Madhavi Alapaty4, Glenn Hunter2, and David Stauffer2

1MCNC–Environmental Programs, Research Triangle Park, North Carolina
2Pennsylvania State University, University Park, Pennsylvania
3North Carolina State University, Raleigh, North Carolina
4M.A.S.C., Apex, North Carolina

1. INTRODUCTION

Large errors in atmospheric boundary layer (ABL) simulations can be caused by inaccuracies in the inputs, assumptions in and simplifications of physical formulations, and other modeling deficiencies. For certain applications, such as air quality studies, these errors can have significant effects. To alleviate such modeling errors, surface observations can be used to improve the accuracy of the simulated ABL. Ruggiero et al. (1996) studied the effects of frequent assimilation of surface observations using an objective analysis in an intermittent technique. However, this was not a continuous assimilation technique. In attempting to perform a direct and continuous assimilation of surface temperature observations, Stauffer et al. (1991) found that serious errors arose in the ABL structure because the sign of the surface buoyancy flux changed unrealistically as new data were assimilated, even in midday conditions. Hence, there is a need for a robust surface data assimilation methodology that can correct modeling errors in the ABL.

Alapaty et al. (2001) have developed a technique that allows continuous assimilation of surface observations to improve surface-layer simulations. In this technique, analyzed surface data are assimilated in a model’s lowest layer simultaneously with an indirect assimilation of ground/skin temperature, thereby maintaining greater consistency between the ground temperature and the surface-layer mass-field variables. This approach helps to eliminate the spurious changes in the sign of the surface buoyancy flux noted by Stauffer et al. (1991). Using this approach, Alapaty et al. (2001) showed that simulation errors in the ABL can be reduced, while minimizing the disruption of the model’s physical processes within the ABL. Note that in some cases errors in the simulated surface-layer variables may be due solely to errors in the soil and vegetation parameters, and in other cases these errors may be due to spurious clouds or advection. Therefore, Alapaty et al. (2001) did not insist that the surface fluxes match observed fluxes, since the goal was to obtain the correct atmospheric structure. To that end, we have implemented the surface data assimilation technique of Alapaty et al. (2001) in the MM5V3.4.

2. DESCRIPTION OF THE TECHNIQUE

Following Stauffer et al. (1991) and Alapaty et al. (2001), the surface data assimilation (SDA) equation for a near-surface variable, α, for use in the MM5 can be written as

$$\frac{\partial \alpha^*}{\partial t} = F(\alpha, x, y, t) + G_{\alpha} W_{\alpha} \varepsilon_{\alpha} p^*(\alpha^* - \alpha)$$

where $p^*$ is the difference between base state pressures at the surface and model top; $t$ is time; $F$ is a forcing term representing all other physical processes affecting $\alpha$ in the model’s lowest layer; $x$ and $y$ are the horizontal spatial coordinates; $G_{\alpha}$ is a nudging factor for $\alpha$; $W_{\alpha}$ is a weighting function that determines the horizontal, vertical, and time weighting applied to the analysis; $\varepsilon_{\alpha}$ is an analysis quality factor ranging between 0 and 1; and $\hat{\alpha}$ is the analyzed (gridded) value obtained from observations for $\alpha$. Substituting temperature ($T_L$) and water vapor mixing ratio ($q_L$) from the model’s lowest layer (closest to the surface) for $\alpha$ in the above equation, respective equations for the surface data assimilation can be written as

$$\frac{\partial p^* T_L}{\partial t} = F(T_L, x, y, t) + G_T W_T \varepsilon_T p^*(T^* - T_L)$$

$$\frac{\partial p^* q_L}{\partial t} = F(q_L, x, y, t) + G_q W_q \varepsilon_q p^*(q^* - q_L)$$

The parameter $G_{\alpha} = G_T = G_q = 9.0 \times 10^{-4}$ s$^{-1}$ is the nudging factor that determines the magnitude of the data assimilation term in the above equations. Note that the inverse of the nudging factor gives a characteristic assimilation time scale (Stauffer and Seaman, 1990). Here, $G_{\alpha}$ for the surface data is chosen to be three times greater than that used for the upper-air sounding data (see Alapaty et al., 2001) because the adjustment rate of the surface fluxes to changes in forcing is quite rapid compared to the time scale of the inertia-gravity waves.
typically responsible for adjustments in the free atmosphere. Adjustment of the ground/skin temperature due to the assimilation of the surface data is as follows. First, we can rewrite the last term in Eq. (2) as \( \frac{\partial T^f_i}{\partial t} \), the rate of change of the surface-layer temperature due to the direct nudging. Since we have chosen to let all of the effect due to the data assimilation occur at the surface, then the nudging adjustment to the turbulent sensible heat flux, \( H^f_s \) (Wm\(^{-2}\)), can be written as

\[
H^f_s = \rho C^s (\frac{\partial T^f_i}{\partial t} \cdot \Delta t) = \rho C^s (\frac{\partial T^f_i}{\partial t}) \Delta z
\]

Similarly, if \( \frac{\partial q^l_i}{\partial t} \) is the rate of change of the surface-layer water vapor mixing ratio due to direct nudging, then the adjustment to the turbulent latent heat flux, \( H^f_l \) (Wm\(^{-2}\)), can be written as

\[
H^f_l = \rho L (\frac{\partial q^l_i}{\partial t} \cdot \Delta t) = \rho L (\frac{\partial q^l_i}{\partial t}) \Delta z
\]

where \( L \) is the latent heat due to condensation. Thus, the adjustment to the ground/skin temperature due to indirect assimilation of surface-layer temperature and moisture data over the interval \( \Delta t \), \( \Delta T^f_i \), can be written in the form of the surface energy budget equation as

\[
\Delta T^f_i = (\frac{\partial T^f_i}{\partial t}) \Delta t = (H^f_s - H^f_l) \frac{\Delta t}{C^s}
\]

(4)

where \( C^s \) is the thermal capacity of the uppermost soil slab per unit area. The ground/skin temperature increment from Eq. (4) is applied at the subsequent time step to be consistent with numerical requirements. Note that we do not assume that all errors in \( T_i \) and \( q^l_i \) are due to errors in the surface fluxes, and that these adjustments should not be considered as “corrections” to those fluxes. We merely recognize that altering the ground temperature through an adjustment to the surface fluxes based on known errors in \( T_i \) and \( q^l_i \) yields a physically convenient indirect way to correct for these errors, regardless of their source.

While this simultaneous direct and indirect assimilation approach does not ensure that the model’s fluxes will always be nudged toward the real fluxes (which may or may not be observed), it does adjust the ground temperature at the surface layer and humidity must converge toward the observations of those variables. In summary, our continuous data assimilation approach inserts corrections smoothly at each advection time step.

3. NUMERICAL SIMULATIONS AND RESULTS

In the present study, the standard 3 and 6 hourly surface observations were used to study improvements in model simulations utilizing the SDA technique. Note that the nudging was performed for every advection time step. The hourly surface data, available from the Techniques Development Laboratory (see http://dss.ucar.edu/datasets/ds472.0/), was used only in the statistical evaluation of model results. However, in near-future work, we will be using hourly surface data to perform SDA to achieve better results than those presented in this paper. The MM5V3.4 simulations were performed for six days starting from July 10, 1997. We used 26 vertical layers; there are 12 layers between the surface and ~2.5 km altitude, with the lowest half-level placed at about ~18 m AGL. We used the NMC Eta model analysis (see http://dss.ucar.edu/datasets/ds068.0/) to prepare model inputs. The standard soil moisture availability scheme was used to estimate the surface latent heat fluxes and the Grell scheme was used to account for subgrid-scale cumulus convection. A nonlocal closure scheme suggested by Blackadar (HIRPBL) was used to represent convective turbulent mixing in the ABL. The four-dimensional data assimilation option was used in the free atmosphere. In the ABL the winds were nudged using the surface data. These are the options used in the base case simulations. In the second set of simulations, in addition to the above-described options, temperature and water vapor mixing ratio are nudged only in the model’s lowest layer using the SDA technique. Thus, temperature and moisture are not assimilated between the second layer from the surface and the layer containing the top of the mixed layer. The horizontal resolution used in these simulations was 36 km. The modeled domain of 121X101 points was centered at 40° N, 90° W and included over 80% of the continental United States (not shown). However, we present the evaluation results for an eastern U.S. subdomain of 1800X1836 km, which is being used for air quality model simulations.

Figure 1 shows the number of hourly surface observations used in model evaluation as a function of time. The zero hour on the x-axis corresponds to 1200 UTC 10 July 1997. Each of the observational sites is paired with the corresponding grid cell in the modeled domain for preparing statistics. Then, simulated temperature, mixing ratio, and winds for these representative grid cells are interpolated linearly from the model’s lowest-level altitude to the respective measurement heights to facilitate direct intercomparison.

In Figures 2 through 7, we refer to the observations as “Obs” and to model results obtained without and with SDA as “Base” and “Sda,” respectively.

The quantity \( M - O \) is used in preparing many statistics, where \( M \) represents the modeled value and \( O \) represents the observed/measured value of a variable. Thus, if \( M - O \) is positive, the model is overpredicting.
that variable, and if $M - O$ is negative, the model is underpredicting. Figure 2 shows the temporal variation of the spatially averaged observed and modeled near-surface water vapor mixing ratios. The dry bias present in “Base” is mostly absent in “Sda.”

The statistics shown in Figure 3 also indicate major improvements in the mixing ratio simulations. The improved simulation of mixing ratio in “Sda” will affect the gradients across the surface and result in changes in the surface energy budget. Nudging of the near-surface air temperature along with the mixing ratio in “Sda” avoids the possible occurrence of cooler temperatures in “Sda,” as seen in Figure 4.

In general, there is negligible deterioration in the near-surface temperatures for this case, as seen in the statistics in Figure 5. As expected, differences between “Base” and “Sda” for near-surface wind speed and direction are again negligible (not shown). As seen in Figure 6, maxima in spatially averaged ABL depths for “Sda” are either the same as or slightly higher than

Figure 1. Number of surface observational sites used in preparing hourly statistics for surface temperature, mixing ratio, and horizontal winds. The zero hour corresponds to 1200 UTC 10 July 1997.

Figure 2. Temporal variation of spatially averaged near-surface water vapor mixing ratio.

Figure 3. Temporal variation of mean error (ME), root mean square error (RMSE), and index of agreement (IA) for near-surface mixing ratio.

Figure 4. Temporal variation of spatially averaged near-surface air temperature.

Figure 5. Temporal variation of mean error (ME), root mean square error (RMSE), and index of agreement (IA) for near-surface air temperature.

Figure 6. Maxima in spatially averaged ABL depths for “Sda.”
Because the water vapor mixing ratio in “Sda” is higher than that in “Base,” this results in increased surface dew point temperature and lower altitudes of lifting condensation level (LCL) in “Sda.” Under favorable conditions in the free atmosphere (i.e., when convective instability exists), increased dew point temperature and lowered LCL will lead to increased convective precipitation. This result is apparent in Figure 7. Over certain regions we found lesser precipitation in the “Sda” than that in “Base”. However, on average, usage of SDA scheme led to an increased rainfall amounts in this case study. This is also a positive result because at coarse horizontal resolutions, a mesoscale model, in general, tends to underpredict precipitation.

4. SUMMARY

In this case study, the SDA technique dramatically reduced errors in the modeled water vapor mixing ratio in the lowest layer. This improvement also led to increased rainfall, which is considered an additional improvement.

Acknowledgments: The authors thank Donald Olerud, Andrew Holland, and Aijun Xiu for their help in this research work.

5. REFERENCES


