The Effect of Network Resolution on Data Assimilation in a Mesoscale Model

J. Dudhia
Mesoscale and Microscale Meteorology Division
National Center for Atmospheric Research
Boulder, CO 80307-3000

Introduction

One goal of the Atmospheric Radiation Measurement (ARM) Program is to characterize meteorological fields over wide areas (200-km square) in order to better parameterize sub-grid-scale variability in general circulation models used for climate studies. Such a detailed knowledge over these areas is impossible with current observational methods alone, but the synthesis of a dataset by combining observations with a mesoscale numerical model is feasible. Current data assimilation techniques allow observed data to be incorporated while a model is running, thus constraining the model to fit the data as well as the data to be dynamically consistent with the model atmosphere. This interaction may therefore be regarded as a *dynamical analysis* technique.

The technique used for data assimilation here will be the nudging method (Stauffer and Seaman 1990, Kuo and Guo 1989). Specifically, observational nudging where data at observational sites are gradually forced in the model without the need for a gridded analysis. This method is particularly appropriate for asynoptic data covering meso-ß-scales, such as will be available at the Cloud and Radiation Testbed (CART) sites. The method makes it possible to incorporate the wide variety of data coming from these sites.

A question that arises in the program's planning is, What are the minimum data necessary for a good characterization of the meteorological fields over a typical CART site? It is important to have a cost-effective balance that provides enough data for a four-dimensional data assimilation (FDDA) system to be accurate. This requires a unified view of data collection and assimilation.

The concept of an Integrated Data Assimilation and Sounding System (IDASS) ensures that the needs of data collection are partly determined by the requirements of an assimilating mesoscale model. Hence, the sounding strategy is geared towards allowing the model to do the best possible job in representing the atmosphere over CART sites, for example.

It is not clear a priori what density of coverage or types of data are required for a good simulation. In this work, we address the problem of determining the impact of varying the density of coverage of an ideal network by purely numerical experimentation. We use one model run to provide data and another independent run to assimilate it. The results of such tests are important to the design of the CART site observational network and the IDASS.

Method

The model used was the Penn State/National Center for Atmospheric Research (NCAR) Mesoscale Model, which has been applied in both hydrostatic and nonhydrostatic modes for this study. The nonhydrostatic version has been recently developed (Dudhia, in press) and requires little change to the data assimilation techniques developed for the hydrostatic model, except that it is now preferable to have the data at specified heights rather than pressures, as used formerly.

Numerical modeling studies were carried out to investigate the effect of data resolution on the accuracy of model predictions. The method of four-dimensional data assimilation was used to incorporate data from one model simulation into an independent simulation of the same event with degraded initial and boundary conditions.

These observation system simulation experiments (OSSES) were carried out on 1) a strong winter-time cold front case over Colorado (13 February 1990), and 2) a spring-time severe convective outbreak in Oklahoma (10 April 1979). Here we will concentrate mostly on presenting results from the spring-time case.

In both the Oklahoma and Colorado cases, nonhydrostatic 5-km grid simulations were run to generate the "true" data that were assimilated into 20-km simulations by the observation nudging technique. This method spreads observational influences over an area and time period surrounding each observation, typically with a 100- to 200-km radius of influence. The data were in the form of profiles of temperature, humidity and wind taken hourly at several individual columns of the 5-km simulation. The influence of changing network density was studied by using uniform square grids of 4, 9, 16 and 25 profiles covering the 500-km square domain of the 5-km simulation. A control simulation was also run without data assimilation. Figure 1 shows the 3 x 3 ideal network in the Oklahoma case's domain.

For the Oklahoma case, the 5-km simulation was 18 h and the 20-km simulations were 24 h long starting 6 h before the 5-km run that provided the data. Both were nonhydrostatic. For the Colorado case, the 5-km and 20-km simulations were 12 h long, the latter being hydrostatic.

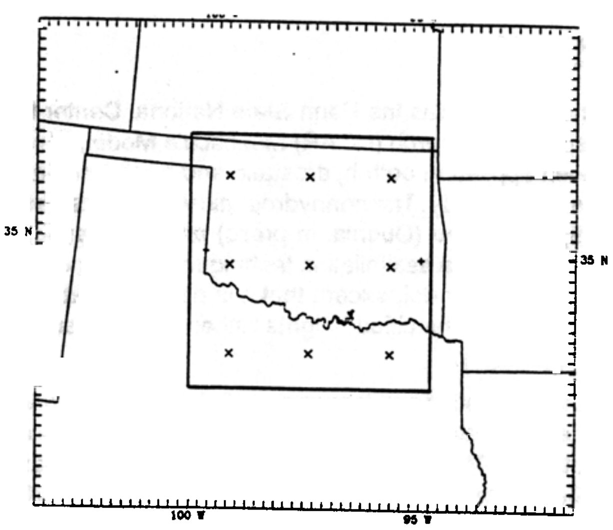

Figure 1. The 20-km domain with the 5-km domain outlined. The 3 x 3 network positions are marked by crosses.

Results

The root mean square (r.m.s.) errors in the fields at several chosen pressure levels and threat scores for the rainfall pattern were used to verify the simulations. Despite the meteorological difference between the two cases studied, conclusions for both were quite similar. Figure 2 demonstrates an example of the results for the temperature r.m.s. at 200 (top), 500 (middle) and 800 (bottom) mb in the Oklahoma case.

These revealed that the gain in accuracy over the control simulation, which had no four-dimensional data assimilation, was significant for even only 4 profile sites in a 2 x 2 network. There was a further but much less significant gain with a 3 x 3 network, but beyond that the gain was small for 4 x 4 and 5 x 5 networks. The r.m.s. errors were typically half that of the control for these high-resolution assimilation runs.

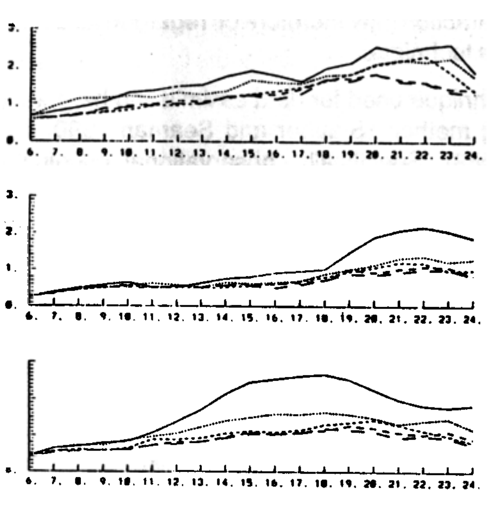

Figure 2. Root mean square error versus time in the 5-km Oklahoma domain of temperature at 200, 500 and 800 mb for five experiments. Continuous line is the Control, shortest dash to longest dash are 2 x 2 to 5 x 5 FDDA experiments, respectively. Vertical axis is in K, horizontal in hours.

For the wintertime case, the frontal position was very similar irrespective of data assimilation, so r.m.s. scores were the best measure of the gain from data assimilation.

The results from r.m.s. scores are supported by the rainfall patterns for the Oklahoma storm case shown in Figures 3a-d. These display a definite improvement with increasing assimilated data. While the control run (Figure 3b) produces weak rainfall at 21-24 h, the assimilation with 25 sites (Figure 3d) produces much stronger precipitation in approximately the correct position. However the amount is still underestimated compared with the "truth" run (Figure 3a).

Figure 3a. 3-hourly rainfall total ending 12Z 11 April 1979 for the "truth" simulation. Contours 0.1, 1, 2, 5, 10, 25 mm.

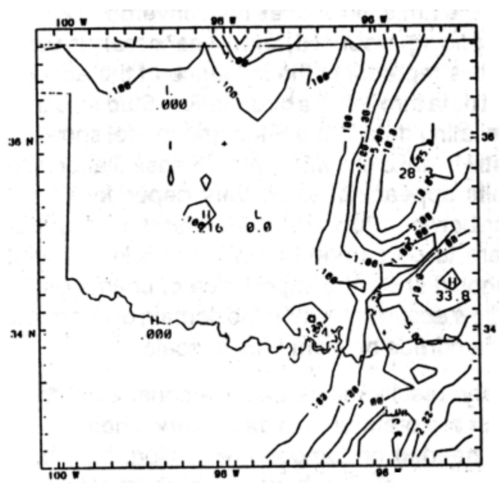

Figure 3c. As Figure 3a but for FDDA 3 x 3 run.

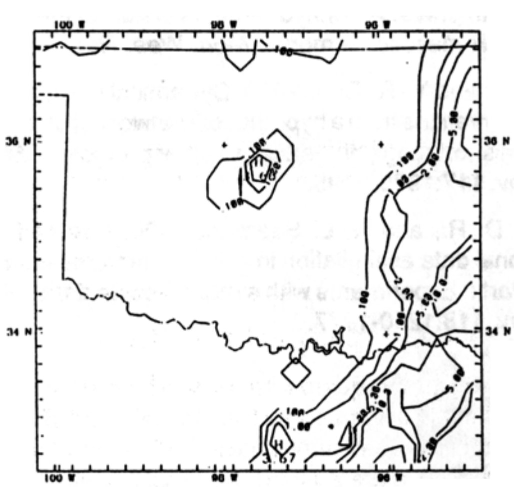

Figure 3b. As Figure 3a but for no-FDDA run.

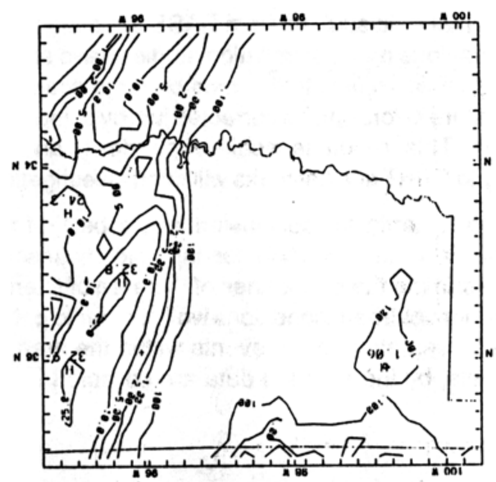

Figure 3d. As Figure 3a but for FDDA 5 x 5 run.

Discussion and Further Work

This implies that much of the benefit of assimilation is gained by even coarse networks probably because most of the meteorological fields' variabilities are at large scales even in strong meso-γ-scale events. This fact can be verified by Fourier analysis of the length scales associated with the "true" simulation where larger scales have larger amplitudes.

However the r.m.s. error does not converge to zero as the network spacing is reduced, and it is necessary to determine whether this represents the influence of the advection of imperfect data through the boundaries. Studies carried out on assimilating data into a 5-km grid model show that the rainfall still is deficient with even 25 assimilation sites, so the results appear not to be very dependent on model resolution between 20 and 5 km. Further studies of boundary effects are to be carried out in these 5-km simulations. These should show the importance of characterizing the exterior flow accurately when the domain of interest is only of order hundreds of kilometers in scale.

Preliminary results have also demonstrated that data every hour are preferable to data every 3 hours, but more tests are needed on temporal resolution. It is, however, unlikely that increasing beyond hourly resolution helps because the time scale of typical data assimilation nudging constants is of this order.

Further, studies would be required also to show how useful wind data alone are, as is easily achievable by profilers, or whether thermal data such as from RASS or balloons are essential. Previous data assimilation studies have shown that the mass field adjusts to the wind on large scales so that temperature errors can be corrected just by using wind assimilation. This result is scale-dependent, and its applicability to CART site networks will need investigation.

At this stage, a tentative conclusion would be that the primary goal in data collection for input to mesoscale models is to sample the boundaries of the area of interest and the exterior upstream conditions well in order to get an accurate representation of the events within the area. In some respects, by the time the data are collected in the site, it is too late to force the model except to improve the accuracy downstream of the observation or to verify the model against. A prospective upper-air network for CART sites probably should not be concentrated within the site, but should be spread around its boundaries to give mesoscale models the best chance to represent the meteorology within the site accurately. As shown in this study, even a few profiling sites are useful because of the inherently large scales of meteorological fields (with the exception, perhaps, of moisture).

With a suitable network of upper-air and surface observations, the data assimilation technique can be applied at CART sites to produce four-dimensional meteorological fields that are representative of the conditions over the whole site. We hope that future applications of the model will include near-real-time running, possibly on a work station at the site, to provide continuous analyses over long periods.

Acknowledgments

This work has been funded by the Department of Energy ARM Project under grant DEA105-90ER61070. Computing was carried out on the NCAR CRAY-YMP supported by the National Science Foundation.

References

Dudhia, J. In press. A nonhydrostatic version of the Penn State/NCAR mesoscale model. *Mon. Wea. Rev.* **121**.

Kuo, Y.-H., and Y.-R. Guo. 1989. Dynamical initialization using observations from a hypothetical network of profilers: Impact on short range numerical weather prediction. *Mon. Wea. Rev.* **117**:1975-1998.

Stauffer, D. R., and N. L. Seaman. 1990. Use of four-dimensional data assimilation in a limited area mesoscale model. Part I: Experiments with synoptic-scale data. *Mon. Wea. Rev.* **118**:1250-1277.