Data Assimilation for the June 1993 Intensive Observation Period at the Southern Great Plains Site

J. Dudhia
Mesoscale and Microscale Meteorology Division
National Center for Atmospheric Research
Boulder, Colorado

Introduction

An Intensive Observation Period (IOP) of the Atmospheric Radiation Measurement (ARM) Program took place at the Southern Great Plains Cloud and Radiation Testbed (CART) site from June 16-26, 1993.

The National Center for Atmospheric Research (NCAR)/Penn State Mesoscale Model (MM5) has been used to simulate this period on a 60-km domain with 20- and 6.67-km nests centered on Lamont, OK. Simulations are being run with data assimilation by the nudging technique (Kuo and Guo 1989; Stauffer and Seaman 1990) to incorporate upper-air and surface data from a variety of platforms. The model maintains dynamical consistency between the fields, while the data corrects for model biases that may occur during long-term simulations and provides boundary conditions. For the work reported here the Mesoscale Atmospheric Prediction System (MAPS) 3-hourly analyses were used to drive the 60-km domain while the inner domains were either unforced or nudged with observations. A continuous 10-day period was simulated.

Overview

One goal of the ARM program is to improve general circulation models (GCMs) by obtaining detailed meteorological information in limited areas on the order of 200 km square and comparing GCM parameterizations with the mean radiative and convective properties in such areas. Typical GCM grid boxes are 100-200 km square, but there is in reality much structure at smaller scales that is represented by their parameterizations. Meteorological observations alone cannot represent this structure, so we use a full-physics mesoscale model forced by large-scale tendencies to give as complete a picture of the sub-100 km scale structures as possible. This allows us to produce a full four-dimensional characterization of the atmosphere that, given sufficiently complete physics in the model and sufficiently good data, will provide a representation of the actual state of the atmosphere. The mean properties and time evolution of features resolved by this model will be compared to those inferred from GCM parameterizations. Furthermore, this dataset is being used to initialize cloud-resolving simulations of selected features (Dudhia and Parsons 1996).

The MM5 Model

The model features and options used in this study are as follows. Equations are for nonhydrostatic, compressible motion in terrain-following coordinates with a polar-stereographic map projection. Prognostic equations exist for wind components, vertical velocity, pressure perturbation, temperature, water vapor, ground temperature, and microphysical water and ice content variables. It has an upper radiative boundary condition, relaxation lateral boundary conditions, and interactive two-way nesting. The model includes microphysics with cloud, rain, snow/graupel, and ice processes on all domains' resolved scales. The Grell cumulus parameterization scheme is adopted only on the 20 km and coarser domains. The Blackadar high-resolution planetary boundary layer and a surface energy budget calculation are used. There is also an atmospheric longwave and shortwave radiation scheme interacting with model clouds and land surface.

Observational Nudging Tests

A series of observational system simulation experiments has been carried out to evaluate the effectiveness of observational nudging in the presence of various amounts of data. In these experiments an independent simulation at higher resolution (5 km) is used to provide the "truth," and selected profiles are taken from this simulation and assimilated into a coarser (20 km) simulation. The independence of these simulations is ensured by using different analyses and boundary conditions to initialize them. Figure 1 shows a typical set-up with a 3 x 3 network of observations (crosses) within the square 5-km domain being assimilated into the 20-km simulation. Configurations ranging from 2 x 2 to 5 x 5, corresponding to sounding spacings from 250 to 100 km were tested to evaluate the impact of network density on the assimilation's accuracy.

The case tested was the 10-11 April 1979 SESAME IOP #1, a severe tornado outbreak. The assimilation's success was measured by r.m.s. error and mean error in the wind components; temperature and moisture at 200, 500, and 800 hPa over the square "truth" domain; and the surface rainfall's threat score.

Results From SESAME Tests

Extensive tests have been carried out in earlier work for ARM, and a few are selected for presentation here. It can be seen from Figures 2a and 2b that for u-wind component and temperature increasing network density gives improved r.m.s. errors and (not shown) improved mean errors. The precipitation threat score is also improved.

Other results from these experiments (presented in Parsons and Dudhia 1997, submitted to *Mon. Wea. Rev.*) have shown that if wind observations are used over a large enough area

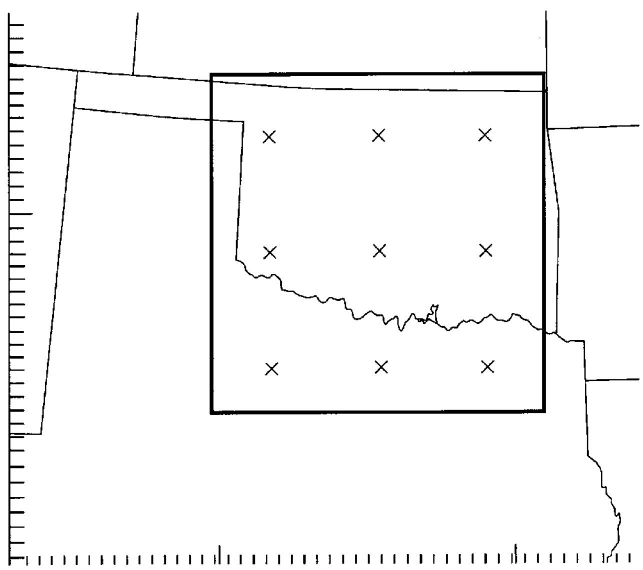

Figure 1. Data assimilation set-up, showing 20-km domain with 5-km domain (square) and assimilation columns (crosses).

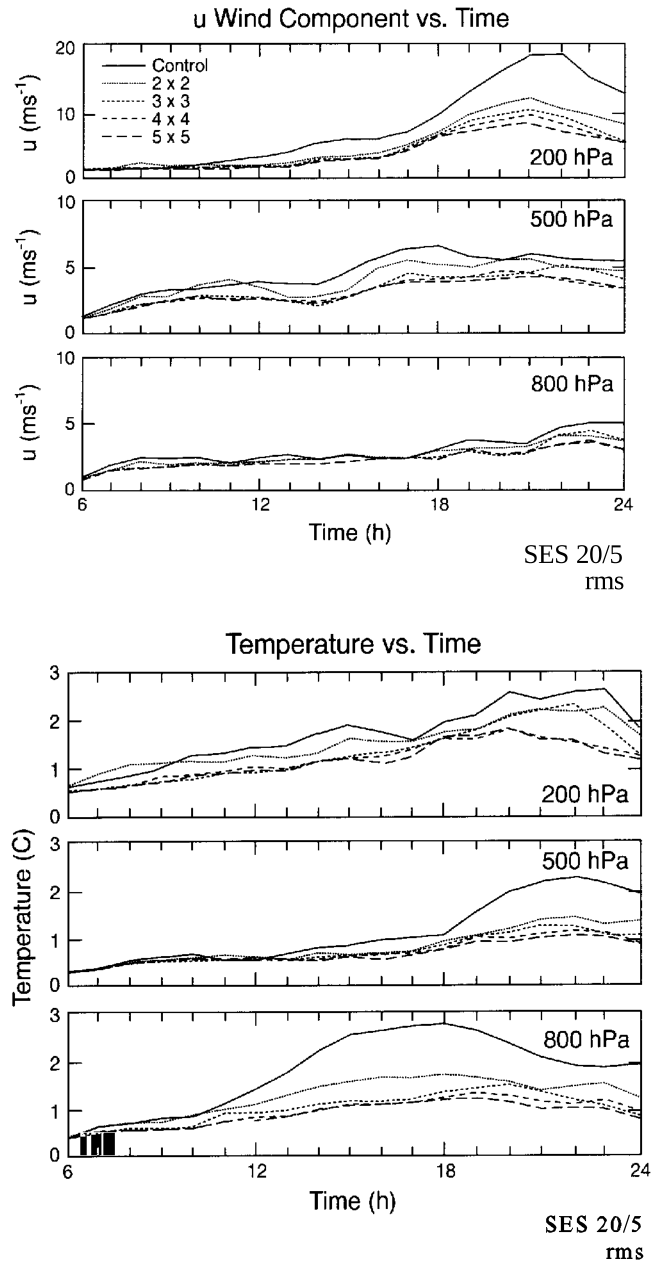

Figure 2. Time series of a) westerly wind r.m.s. error (m/s) and b) temperature versus time (hours) at three pressure levels (200, 500, and 800 hPa). Lines show different data densities as in key. Data from 5-km simulation assimilated into 20-km simulation.

compared to a Rossby radius, this data assimilation method, employing a dynamical model, can improve the thermal field due to the model's large-scale balance. Furthermore, the results have shown that caution needs to exercised in selecting data for input to the system, because individual soundings may not be representative of a large area or long time period due to local effects, and these data are unsuitable for the nudging technique, particularly if the model is unable to resolve the local effects.

The needs for broad-scale data coverage and high resolution locally imply that a multiscale data assimilation approach is best, where a nested model assimilates both large-scale and local data.

Comparison with Analysis Technique

An alternative and simpler approach to providing four-dimensional fields in a local region, such as the ARM CART site, is to use objective analysis. This method would utilize a background field provided by NMC's operational models and analysis together with special ARM data to provide the analysis of the CART site's conditions. This analysis is of a resolution determined by the NMC's gridded product (currently 30-60 km) and the observing network, and the time resolution is similarly limited. Over a small region 100-200 km across it is possible to have large changes between synoptic observations, and since these are usually nonlinear, interpolation in time is not a good solution. The analysis method also lacks any estimate of cloud cover or local effects due to terrain variability, and cannot distinguish between convective-scale and large-scale ascent. However it is a useful method for determining mean properties over the region (e.g., mean divergence or temperature advection), and so a comparison has been made between objective analysis and four-dimensional data assimilation (FDDA) using the same sites as in the network studies and taking the no-FDDA 20 km control experiment as the background. The results presented in Figure 3 (right) show that analysis performs well given sufficient data density.

Figure 3 shows the 200 hPa divergence with time for the FDDA networks, the analysis with the same data densities, and the truth value. Note, for instance, how the error in the sign of the divergence around 10 and 22 hours (solid line) is corrected by both these techniques. Also note how poorly analysis performs relative to FDDA for coarser networks (shortest dashed lines).

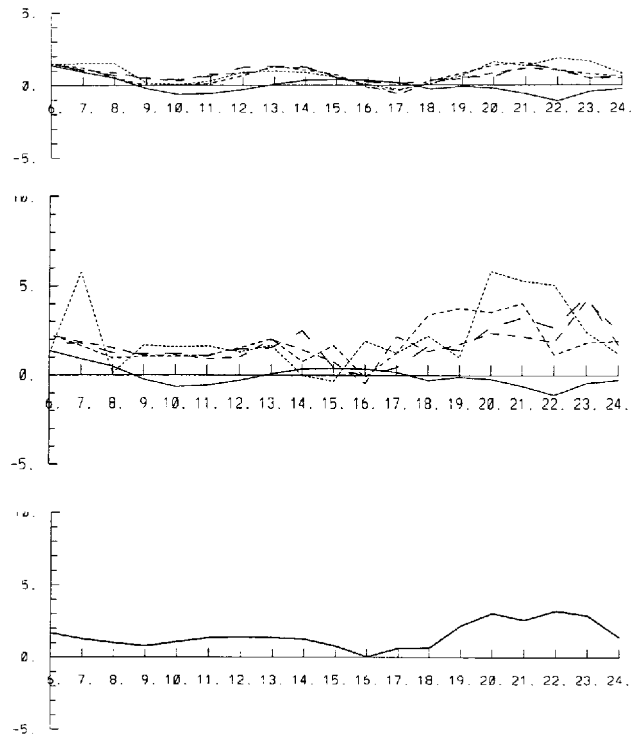

Figure 3. Divergence at 200 hPa (10-5s-1) from FDDA model output (top), analysis only (center), and "true" value (bottom) for different data densities as in Figure 2.

ARM Data Assimilation: Domain and Simulations

Figure 4 shows the areas covered by the three MM5 domains. The 60-km domain coincides with the MAPS domain. The inner 6.67-km domain is centered on a profiler hexagon of the demonstration network around Lamont, OK, and covers a 480-km square. There is good coverage from the profiler demonstration network which provides hourly averaged winds through the troposphere. Two simulations have been conducted to evaluate the benefit of the added data within the 20-km domain: 1) A control simulation just assimilating MAPS data on the 60-km domain using the analysis nudging technique; and 2) a repeat of the first, but additionally assimilating data within the 20- and 6.67-km domains by the observational nudging technique. This data includes the 405-MHZ profiler network (Figure 4b); a 915-MHZ ISS profiler at the central CART site; 12-hourly NWS soundings plus enhanced 6-hourly soundings at the Dodge City, Norman, and Topeka sites; and NWS surface observations.

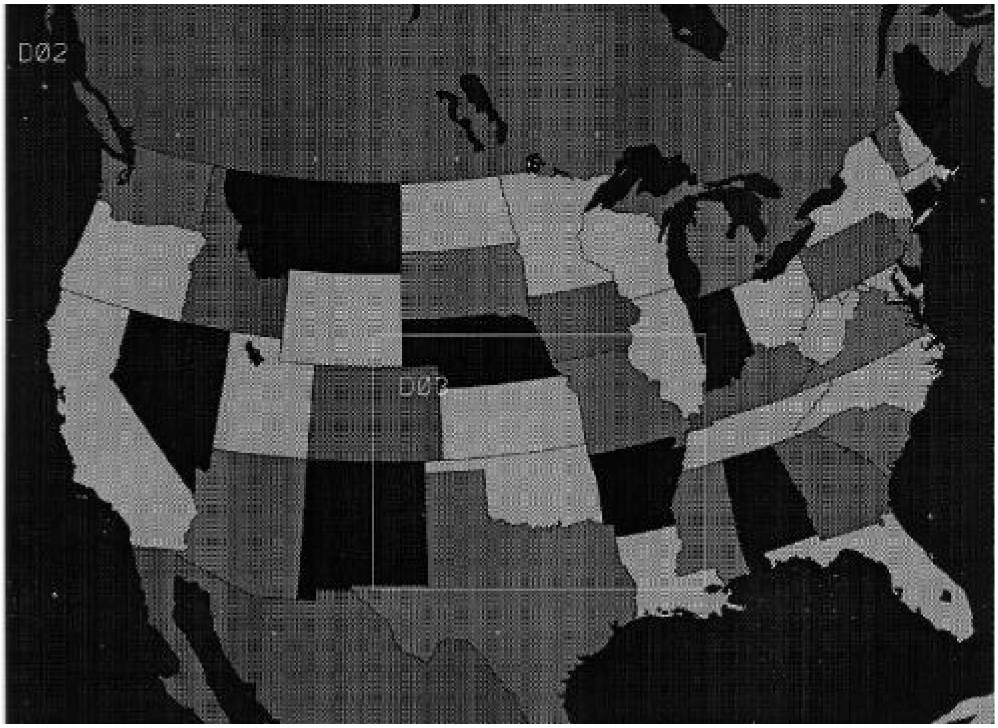

Figure 4a. 60-km domain showing 20-km domain in box.

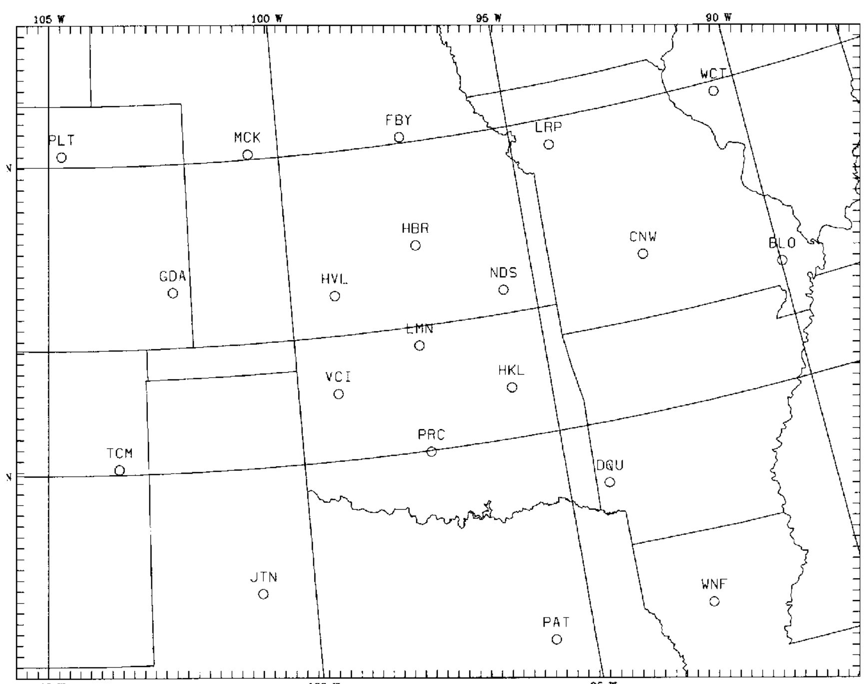

Figure 4b. 20-km domain showing wind-profiler demonstration network.

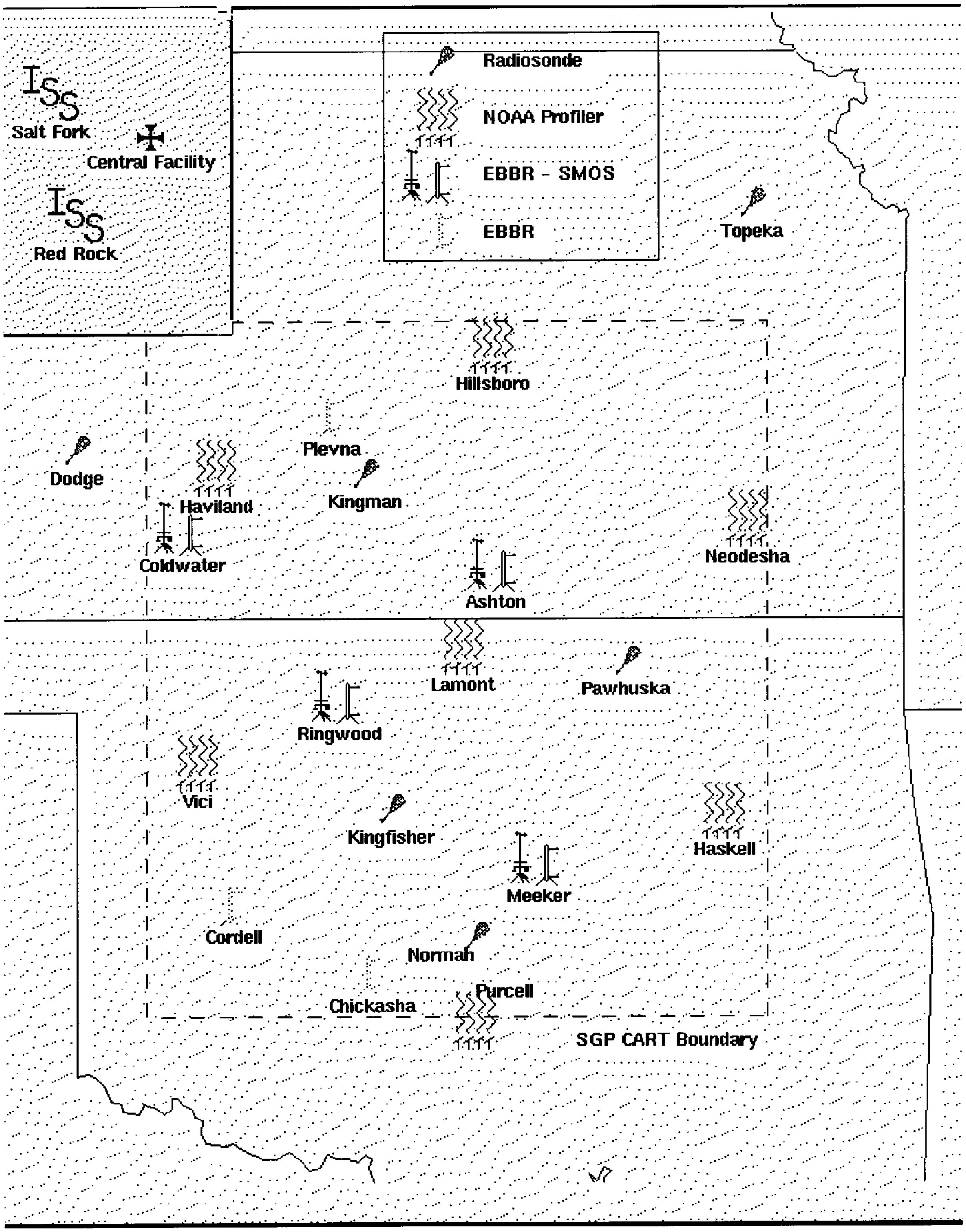

Figure 4c. ARM SGP CART site showing instrumentation.

Data Assimilation of June 1993 Case - Verification

The two simulations (with and without assimilation of data in the 20-km domain) are being compared to determine the effectiveness of the additional observational data in improving the accuracy of the assimilation. This comparison is being made using independent data (i.e., data that has been gathered during the period but not assimilated). During the June 1993 IOP special balloon soundings were taken for the purposes of verifying the data assimilation. The primary sites for these launches were Kingfisher, OK, Kingman, KS, and Pawhuska, OK (see map, Figure 4c) where 8 soundings were taken per day during the 10-day IOP. These sites are well removed from the profiler and sounding sites that were used in the assimilation, but are within the CART site and the 6.67-km domain. Extra soundings were also taken less regularly at the central facility of the CART site and at Saltfork, OK, nearby.

Soundings from the model run are directly compared with the soundings at Kingfisher on 24 June, the ninth day of assimilation, in Figure 5. This shows that the model produces realistic approximations to the actual sounding during a period that includes a nocturnal jet and convection.

Conclusions

The following conclusions were made:

- The assimilation has been effective in incorporating all the routinely available data in a model simulation of the 10-day FDDA IOP of June 1993.

- The data assimilation technique compares favorably with objective analysis in data-sparse regions where sounding spacings are 200 km or more.

- Data assimilation maintains a realistic balance between mass and wind fields and can employ wind data over a large scale to improve the thermal field. This is particularly pertinent to the CART site which is surrounded by a large-scale profiler network.

- Care needs to be taken that the data is not only clean, but also representative of the scale over which the model assimilates it. Quality control and neighbor checks are a vital step.

- The assimilation can provide realistic scales for ascent associated with mean convergence, which is information that cannot be inferred from sparse observations alone.

Given adequate microphysics the model can also provide the cloud field consistent with its assimilated state.

References

Dudhia, J., and D.B. Parsons, 1996: Cloud-resolving model simulations conducted over the SGP CART site, this volume.

Kuo, Y.-H., and Y.-R. Guo, 1989: Dynamic initialization using observations from a hypothetical network of profiles. *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.

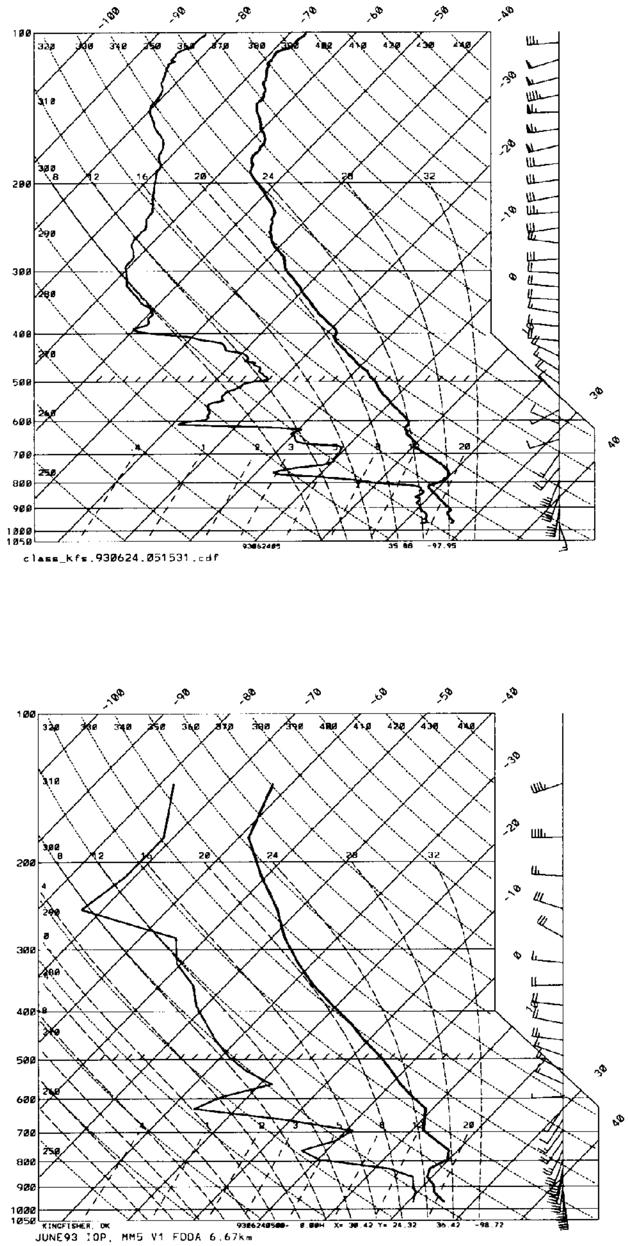

Figure 5a. Sounding at Kingfisher, OK, at 0500 UTC 24th June 1993 observed (top) and model-simulated (bottom).

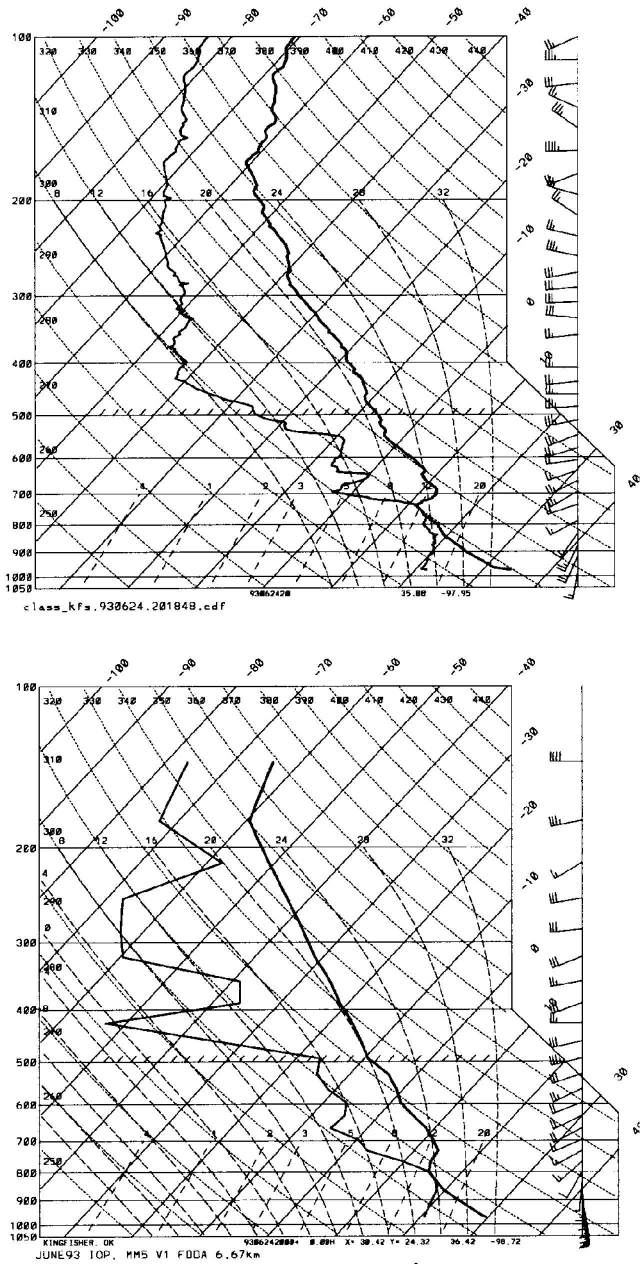

Figure 5b. As Figure 5a but at 2000 UTC 24th June 1993.