ATMOSPHERIC AND ENVIRONMENTAL RESEARCH, INC.

Progress Report No. 3

on

Testbed Model and Data Assimilation for ARM

A research program funded by
Office of Energy Research
U.S. Department of Energy
through Procurement Instrument
DE-FG02-90ER61065

Submitted by:
Atmospheric and Environmental Research, Inc.
For the period
1 September 1992 through 30 April 1993

April 28, 1993

Jean-François Louis
Principal Investigator

MASTER

840 MEMORIAL DRIVE, CAMBRIDGE, MA 02139 TEL:(617) 547-6207
FAX: (617) 661-6479 TELEX: 95-1417

1. Background

The ultimate objectives of this research are to further develop the ALFA (AER Local Forecast and Assimilation) model originally designed at AER for local weather prediction and apply it to several related purposes in connection with the Atmospheric Radiation Measurement (ARM) program: (a) to provide a testbed that simulates a global climate model in order to facilitate the development and testing of new cloud parametrizations and radiation models; (b) to assimilate the ARM data continuously at the scale of a climate model, using the adjoint method, thus providing the initial conditions and verification data for testing parametrizations; (c) to study the sensitivity of a radiation scheme to cloud parameters, again using the adjoint method, thus demonstrating the usefulness of the testbed model.

The data assimilation uses a variational technique that minimizes the difference between the model results and the observation during the analysis period. The adjoint model is used to compute the gradient of a measure of the model errors with respect to nudging terms that are added to the equations to force the model output closer to the data.

The radiation scheme that has been included in the basic ALFA model makes use of a generalized two-stream approximation, and is designed for vertically inhomogeneous, multiple-scattering atmospheres.

This project is designed to provide the Science Team members with the appropriate tools and modeling environment for proper testing and tuning of new radiation models and cloud parametrization schemes.

2. Progress During Previous Periods

The plan for the first year of the project was to incorporate the radiation code of Toon, et al. into the ALFA model, to write its adjoint and to modify the model to be able to use it in data assimilation. This work has been done. As part of this work, code modularity was enforced to enable easy addition or replacement of parts of the code.

During the second year of the contract we have spent a fair amount of time checking the accuracy of the radiation code by trying to simulate the evolution of the atmospheric temperature measured during the Wangara campaign[1], and comparisons with detailed line-by-line radiation computations. We found a number of problems with the code, which were eventually traced to an error in the spectral data for the water vapor continuum. After correcting this error, the code now reproduces well the line-by-line computations and simulates the Wangara data reasonably well.

We also started writing the adjoint of the Kuo convection scheme and the stratiform precipitation scheme. Both of these phenomena are characterized by threshold processes, with different behaviors when some model variables reach some critical values. Stratiform precipitation starts when the specific humidity reaches saturation, and convection requires positive moisture convergence and conditional instability. This means that the model equations are only piecewise continuous, and that their derivatives with respect to the model variables are not defined at these threshold values. This may create difficulties in the convergence of the data assimilation procedure.

3. Progress During the Reporting Period

3.1. Model developments

The first task of this period was to complete the adjoint of the moist processes. We chose two different strategies for stratiform precipitation and convection. For stratiform precipitation, it is relatively easy to write the equation as a continuous process, by replacing the instantaneous adjustment by an algorithm that removes a fraction of the existing moisture at each time step, the fraction removed being a strong function of relative humidity. This was described in our last report.

For the convection scheme, which acts on the atmosphere with a time scale comparable to the model time step, we kept the model as it was, as a discontinuous process when going from stable to unstable regime. So far it does not seem to have created any problem, but we have not yet been able to test many cases of convective situations. it is hoped that the forthcoming June 1993 IOP will provide the necessary data.

---

[1]Clarke, R. H., A. J. Dyer, R. R. Brook, D. G. Reid and A. J. Troup, 1971: The Wangara Experiment: Boundary layer data. Technical Paper No. 19. Division of Meteorological Physics, CSIRO, Australia.

Writing the adjoint of the convection scheme raised a problem that we had not encountered in the rest of the model, and which has to do with balancing memory usage versus computation in the adjoint. In general, any step t of the model, for example the convection computation, can be written symbolically as

$$x_{t+1} = x_t + f(x_t), \tag{1}$$

where x is a model prognostic variable (temperature, moisture) and f is a nonlinear function of x, then the corresponding adjoint equation is

$$x_t^* = x_{t+1}^* + f'(x_t).x_{t+1}^*, \tag{2}$$

where x^* is the adjoint variable and f' the derivative of f with respect to x. Note that this derivative has to be estimated at time t. In practice, the value of x_t is stored at each time step during the model integration, and read at the beginning of the correct adjoint step. Since the convection scheme is written as an adjustment of the fields after all the other physical processes have acted, the value of x_t in (2) must actually be that intermediate state. In our model, we store the value of x at the beginning of each time step, and the tendencies due to all the physical processes except convection. Then, in the adjoint, the intermediate state can be recomputed before the adjoint of the convection is performed.

Having to store or reconstruct the intermediate states could become a serious problem if the model were written as a series of adjustments. In that case all the successive intermediate states would have to be stored, and read in the reverse order in the adjoint. This could place considerable stress on the needed memory. To reconstruct them would be expansive and awkward since the adjoint model performs the operations in reverse order from the forecast model. Fortunately, the ALFA model was written in such a way that all the physical schemes, except convection, act on the atmospheric structure at the beginning of the time step and all the tendencies are added together at the end. The order of computation of the physics is therefore not important.

We also modified the Toon radiation scheme to introduce the handling of partial cloudiness. At this stage we allow only one cloud layer, but it can be anywhere in the vertical and have any thickness. The radiation computation is done separately in the clear and cloudy parts, and the fluxes are combined. One example of this computation is shown on Figure 1.

In this case, the model predicts about 50% cloudiness during the first day, dissipating at night, then reappearing and reaching 100% cover at the end of the second night. Note that

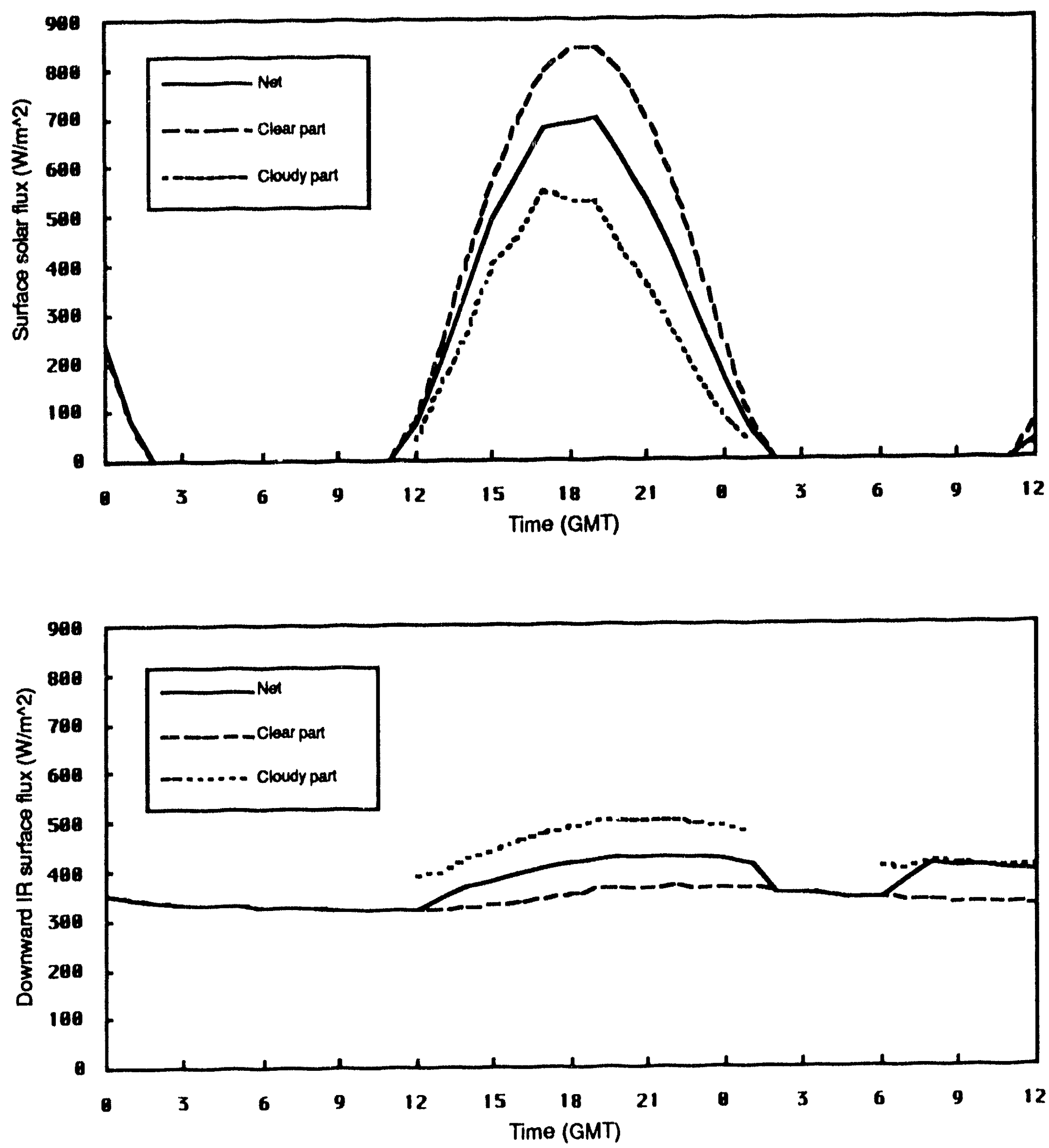

Figure 1: Example of radiation computation output during a 36 hour forecast.

this kind of output can be compared to radiometric measurements and can therefore be used in the computation of the cost function in data assimilation. That is something that no other data assimilation method can do easily. In fact, the reason why some much effort has been put into developing temperature retrieval techniques for satellite data is because the operational data assimilation schemes could not use the radiances as input data.

3.2. Data assimilation tests

Under another contract (NSF SBIR), we have accumulating one complete year of all the standard meteorological data, as well as the NMC analyses and forecasts, for North America. We have started using these data to optimize our model to the Oklahma site, and to test our data assimilation system. It was our intention to combine them with the CART data when they became available. Unfortunately, the NWS surface observations for the center of the US, including Oklahoma, have been missing from this dataset since the beginning of October.

Comparing the soundings at Oklahoma City with those made at the SGP site, about 120 km away, illustrates the need for data assimilation to eliminate small scale variations. In most case the temperatures differ by several degrees, even though both soundings are made within the CART site, only a few hours apart. This would produce differences of a few watts/m^2 in flux computations. Figure 2 shows an example, for Jan. 5 1993.

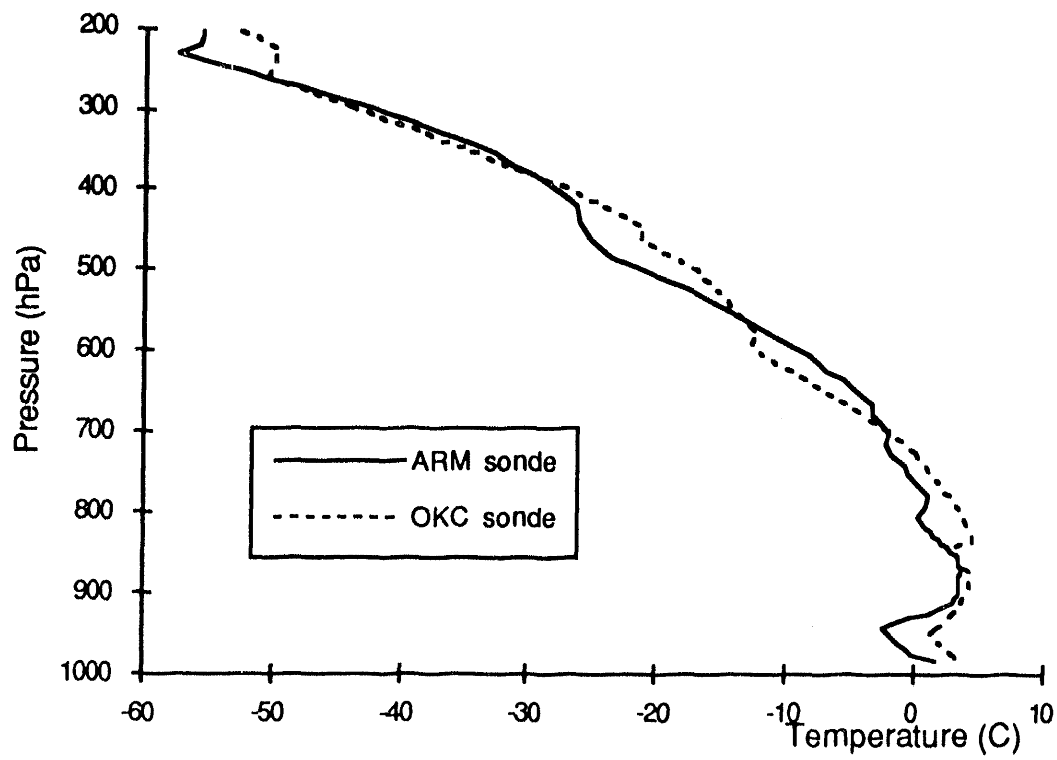

Figure 2: Soundings for Oklahoma City (19:30GMT on 1/05/93) and the ARM site (00GMT on 1/06/93)

We have been experimenting both with optimization of the model physical parameters, and with the Derber nudging[2] type of data assimilation. In Figure 3 we show the result of optimizing the ground parameters of the model (albedo, roughness length, heat capacity, heat and moisture diffusivities, field capacity for moisture).

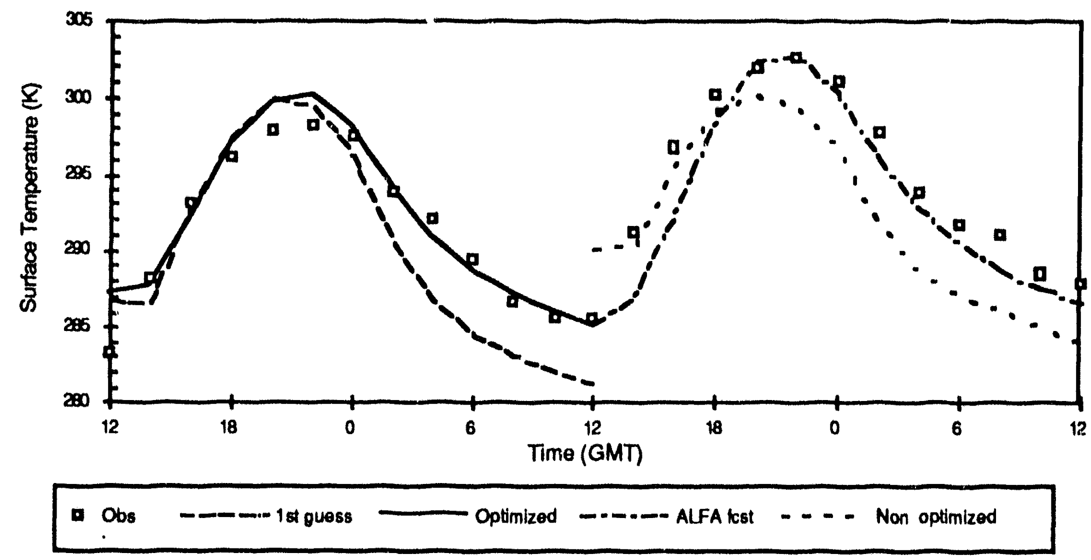

Figure 3: Example of model parameter optimization for OKC station, May 7-8, 1992

The curves labeled "1st guess" and "non-optimized" are 24 hour forecasts of the surface temperature performed with our ALFA model, using parameters that we believed to be reasonable for Oklahoma. The initial conditions are taken from the NMC analysis at 12 GMT. These forecasts are not very good. They have a tendency to drift towards lower temperatures.

We now use the data during the first 24 hour period to optimized the model ground parameters, that is to find the set of parameters that will minimize a measure of the forecast error. We do not use only the hourly surface observations, but also the sounding data at 00 and 12 GMT. The procedure is stopped after 20 iterations, at which point the cost function no longer decreases. The resulting forecast is the curve labeled "Optimized". Except for a small overshoot at noon, the forecast temperature curve looks much better, following the observations quite closely. We now perform a forecast (labeled "ALFA fcst") starting from

---

[2]Derber, J. C., 1989: A variational continuous assimilation technique. *Mon. Wea. Rev.*, **117**, 2437-2446.

the end of the optimized run, using the new parameters. This forecast is much better than the non-optimized one.

Of course we could not claim, with one 24 hour period, to have found the best model parameters for the Oklahoma City station. In fact, when we repeat this operation for different dates, we obtain different sets of "optimal" parameters. It may be that some of these parameters, which we take to be constant, may actually change with time. For example, the albedo and the roughness length change with vegetation. Nevertheless, it is clear that when we optimize parameters over a single 24 hour period, the result is influenced rather too much by errors in the initial state. It is not yet clear how many periods should be averaged for a more representative solution. Marais and Musson-Genon[3], with a similar but simpler model, found that averaging over 10 days gave good results.

In Figure 4, we give an example of data assimilation. The figure shows the surface temperature observations for two days, our ALFA forecasts before optimization, which we also call first guess, the ALFA analysis for the first day and the resulting forecast for the

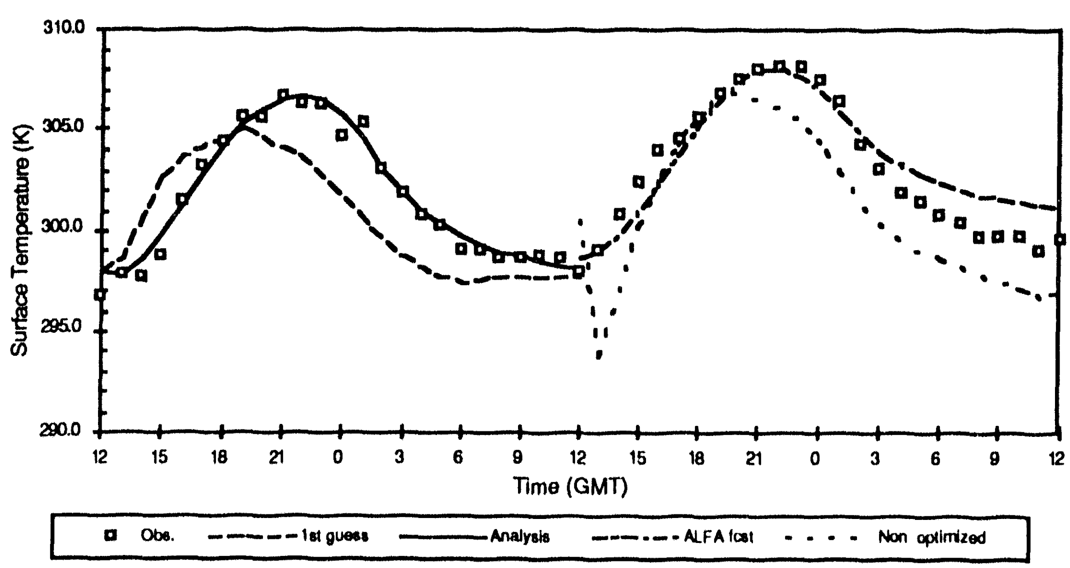

Figure 4: Example of Derber nudging data assimilation. July 21-22, 1992, OKC station.

[3]Marais, C. and L. Musson-Genon, 1992: Forecasting the surface weather elements with a local dynamical-adaptation method using a variational technique. *Mon. Wea. Rev.*, **120**, 1035-1049.

second day. The first guess forecast is performed with what we think may be reasonable physical parameters for the Oklahoma site, with the Nested Grid Model analysis of the National Weather Service as initial conditions.

We used the Derber nudging algorithm to assimilate data during the first day. We have used constant nudging: at each time step we add constant terms to all the tendency equations. These terms are different for all the variables and also depend on height. They are our control variables. They are all zero at the start of the assimilation procedure, resulting in the curve labeled "1st guess". The data assimilation procedure is iterative, each step requiring the integration of the forecast model to compute the cost function, the integration of the adjoint model to compute its gradient with respect to the control variables, and a modification of the control variables. All the data available from the NWS are used to define the cost function: hourly surface observations of temperature, humidity, and winds, and soundings at 00 and 12GMT. The iterative process is stopped when the cost function, i. e. the forecast error during the first day, is minimum (curve labeled 'Analysis"). A 24-hour forecast is then performed from the end of the assimilation period ("ALFA fcst" curve), with the final nudging terms included.

The result of the data assimilation is very good: the analyzed temperature curve follows the observations very closely, but without the small scale observational noise that we want to eliminate. The subsequent forecast, in which we continue the nudging, is also quite good. One advantage of variational data assimilation methods can also be seen: they tend to eliminate the so-called "spin-up" problem, which is a rapid adjustment of the fields at the beginning of the forecast because of imbalances in the initial fields analyzed by standard methods. Such a rapid adjustment can be seen at the beginning of the "non optimized" forecast of the second day, which uses the NMC analysis as initial state.

Our tests are not always so good. In some cases we have had problems with the convergence of the assimilation procedure. Sometimes the convergence is very slow, or the cost function grows again after reaching a minimum. This needs further investigation, and we plan to experiment with different minimum search techniques to make it more efficient and more robust. Nevertheless our first results are very encouraging, and we are confident that the variational data assimilation method, using a single-column model, will prove to be a powerful tool for data fusion and data assimilation.

A lot of work remains to be done. The Derber nudging method will require considerable tuning, especially in defining the vertical profiles of the nudging terms. Up to now we have let them adjust freely, but that creates a problem when observations are available at only a few levels. A smoothness constraint should probably be enforced.

So far, we have also chosen fairly simple situations, avoiding convective cases. It is not known yet whether the kind of thresholds involved in the convection will create convergence problems in the minimization. Finally we need to develop what might be called "observation simulators", i. e. algorithms to create output similar to the observed quantities, for as many of the ARM instruments as possible.

3.3. Sensitivity Studies

The adjoint method is a powerful tool for sensitivity studies since, with one integration of the model and one integration of the adjoint, one obtains the sensitivity of the objective function to all the model inputs and parameters. We have started doing some of this kind of work, mostly at this stage to examine the details of the data assimilation system.

Figure 5 shows an example of the kind of insight that can be obtained. We run the ALFA model for one single time step and define the cost function as the value of the downward

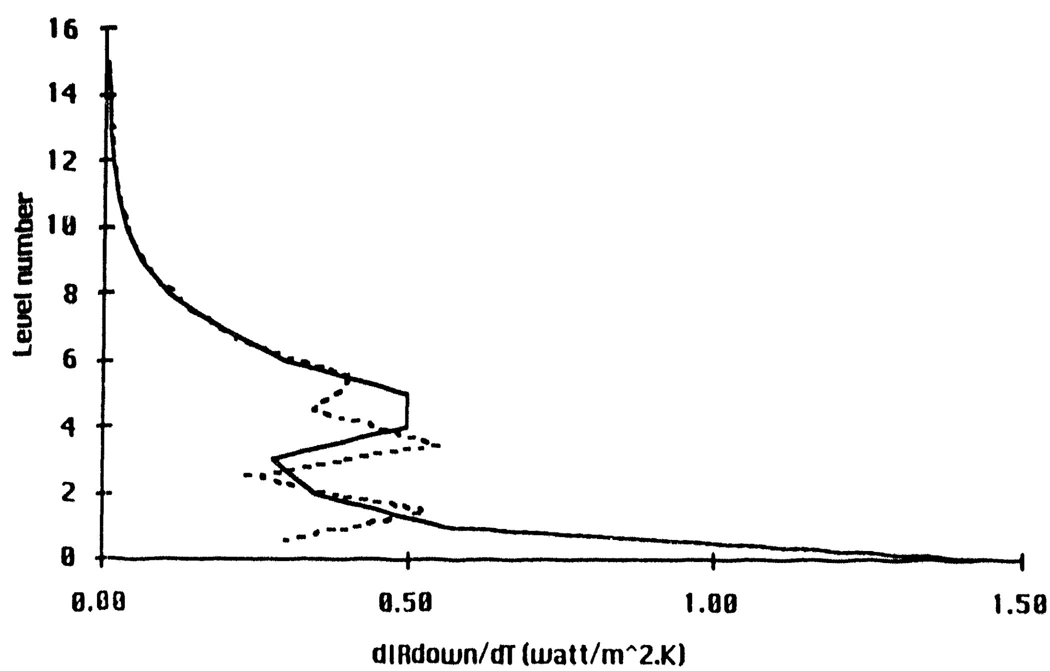

Figure 5: Sensitivity of the surface downward IR flux to the atmospheric temperature (OKC station, 9/2/92). Solid: sensitivity to the layer interface temperature. Dashed: sensitivity to the mean layer temperature.

infrared radiative flux at the surface. The adjoint is then run, computing the derivative of this flux with respect to all the model variables and constants. Here we show the derivatives with respect to the atmospheric temperature.

The ALFA model defines temperature as the mean over a layer, but the radiation scheme uses temperatures defined at the layer interfaces, which are interpolated from the layer mean temperatures. Figure 5 shows that, while the sensitivity of the surface flux to the radiation scheme temperature has a smooth vertical distribution, the interpolation results in a $2\Delta z$ noise in the sensitivity of the flux to the prognostic model temperature! This kind of analysis is extremely easy to perform with the adjoint. It is a powerful tool to discover model weaknesses that may have repercussions on the convergence properties of the data assimilation procedure. We plan to make a systematic sensitivity study of the various model output quantities.

4. Travel

Travel since the 1992 progress report :

- October 28 - 29, 1992: Workshop of the Data Assimilation and Single Column Model science team sub-groups in Richland, WA. (J-F. Louis)
- March 1-5, 1993: ARM Science Team meeting in Norman, OK. (J-F. Louis and Marina Živković

Forthcoming travel:

- May 3-7, 1993: Annual meeting of the European Geophysical Society in Wiesbaden, Germany. (J-F. Louis)

5. Plans for the Rest of the Period

The remainder of the current contract will be devoted mainly to testing the data assimilation system. We will do it in a variety of weather situations, both cloudy and clear and, if possible in some convective cases. We will determine the best minimum search technique for this purpose. We will also start making tests of continuous data assimilation, in which the state at the end of one assimilation period becomes the initial state for the first guess of the next period.

6. Publications and Presentations Since the Last Report

A review paper, entitled "Review of the use of the adjoint, variational methods and Kalman filter in meteorology", including a complete bibliography, by Philippe Courtier, John Derber, Ron Errico, Jean-François Louis and Tomislava Vukičević was submitted to Tellus, for their planned special issue covering the workshop that we organized in Monterey, CA in August 1992.

Jean-François Louis will present a paper entitled "Variational data assimilation at a single site for climate model testing" at the EGS meeting in Wiesbaden, Germany.

END

DATE FILMED

1 / 13 / 94