JP 2.18

A MM5-Based Four-Dimensional Variational Analysis System Developed for Distributed Memory Multiprocessor Computers

Frank H. Ruggiero
Space Vehicles Directorate, Air Force Research Laboratory, Hanscom AFB, MA

George D. Modica, Thomas Nehrkorn, Mark Cerniglia
Atmospheric and Environmental Research, Inc., Lexington, MA

John G. Michalakes
Argonne National Laboratory, Boulder CO

Xiaolei Zou
Department of Meteorology, Florida State University, Tallahassee, FL

1. INTRODUCTION

Since the equations defining atmospheric structure and motion are relatively well understood, the problem of numerically simulating weather systems is primarily one of obtaining the correct initial state. Obtaining this initial state for meso time and space scales is virtually impossible due to the coarse distribution of conventional observations as well as the complete unavailability of conventional data in many areas, particularly over the oceans. The deployment of remote sensing systems has the potential to change this situation. For example, Doppler radar and satellite sensors can provide data at very fine temporal and spatial resolutions. A significant obstacle to the use of radar and satellite data is that the quantities that can be measured by these instruments are not directly usable by the models and tend to be irregularly distributed in time and space. To be truly useful, these data must be converted into regular 4-D fields (time and 3 spatial dimensions) for meteorological quantities or be directly assimilated into the forecast model to produce the 3- and 4-D fields for atmospheric variables. This latter option is made possible in 4-D Variational (4DVAR) data assimilation through the use of a forward observation operator that converts the model variables into the observed quantities.

Corresponding author address: Dr. Frank H. Ruggiero, AFRL/VSBL, 29 Randolph Road, Hanscom AFB, MA 01731-3010
email:frank.ruggiero@hanscom.af.mil

The Air Force Weather Agency is currently running the Penn State/National Center for Atmospheric Research's Fifth Generation Mesoscale Model (MM5) in real time to support operations. In this presentation, we describe activities underway to improve the mesoscale forecasting capability there by developing for real-time use a scalable version of a 4DVAR system that also will be able to assimilate satellite observations. A vector version of a 4DVAR system compatible with MM5 already exists (Zou et al. 1998), however it is too computationally expensive for operational use. In order to achieve the goal of real-time speed for operations, it is estimated that the system will have to have a wall-clock time speed up of over an order of magnitude.

2. BACKGROUND

Variational data assimilation using the adjoint formalism was first proposed in a concise formulation by Le Dimet and Talagrand (1986) for meteorological applications and has been implemented by many others since. Of particular interest to this current project is the work on the application of 4DVAR to the MM5 NWP model (Zou et al. 1995). Zou et al. (2001) showed that the 4DVAR technique combined with MM5 produced superior forecasts of the position and central pressure of Hurricane Felix (1995 Atlantic Hurricane Season) compared with forecasts from the National Hurricane Center. 4DVAR requires small model and observational errors, good parameterizations of physical processes, and quality control of the data. These requirements are usually not restrictive. However, the principle disadvantage of 4DVAR is its large computational requirement due to the iterative nature of method and its high demand in terms of CPU and memory. This limitation has all but prohibited its use at operational weather forecasting centers and previous experiments have been limited to very small domains or relatively coarse spatial resolutions (see, for example Zou and Kuo, 1996). The European Center for Medium Range Weather Forecasts is currently the only operational center in the world that is running 4DVAR (Rabier et al. 2000), albeit on a global scale with coarse spatial resolution and simple model physics for the data assimilation model. Typically, at least 10-15 iterations are needed to minimize the cost function. The method is only practical with computational power in the tens of gigaflops, such as is available from Massively Parallel Processors.

3. METHODOLOGY

As mentioned above, a vector version of the four-dimensional variational analysis system for MM5 already exists, however it is based version 1 of the model. In addition to the model dynamics, the MM5v1 4DVAR system also includes a bulk aerodynamic formulation of the planetary boundary layer, a dry convective adjustment, grid resolvable large-scale precipitation and Kuo (1974) cumulus parameterization. The current MM5 forecast model is now up to version 3 and includes substantial coding and physics upgrades. Therefore, it was decided that instead of making the existing MM5v1 4DVAR system scalable, the project would create updated versions of the tangent-linear and adjoint model components that would be compatible with the latest release of MM5 (version 3.4). This decision was made based on feedback from potential users of the code as well as issues regarding maintainability of the code.

The current MM5v3 forecast model will be the non-linear model in the new 4DVAR system. It is already coded to run optimally on distributed memory multiprocessor machines (Michalakes 1998). For the initial work in this project, it will be run with the same physics used in the previous 4DVAR system described above. For the tangent-linear and adjoint models, the version 3 based codes will be created in a two step procedure. The first step is to create a vector version of each model. This is accomplished with help from the automated Tangent-linear and Adjoint Compiler (TAMC) developed by Giering and Kaminski (1998) and by hand checking the results with visual comparisons the MM5v1 versions. More details on this aspect of the project can be found in a companion paper presented at this conference by Nehrkorn et al. (2001). Each version 3 subroutine is individually processed though the tangent-linear and adjoint model code development procedure. Next, each tangent-linear and adjoint routine is then unit tested following procedures outlined in Zou et al (1997).

Once complete, the full tangent-linear and adjoint models are modified to run on distributed memory mult-processor machines. This is carried out using parallelization techniques developed at Argonne National Laboratory that preserve the original look and operation of the existing serial code. These techniques have been successfully demonstrated with the Same Source Parallel MM5 (Michalakes, 1998). In particular, it involves using spatial domain decomposition where the horizontal analysis domain is partitioned into smaller sub-domains or patches. A separate processor can then be dedicated to the computations for a specific sub-domain. During execution of the 4DVAR code, each patch would require only boundary information from its neighboring patches and thus shared memory can be minimized. Communication between the various patches is accomplished with the use of the standard Message-Passing Interface (MPI) Library. Adaptation of these techniques to the tangent-linear model is relatively straight forward since the tangent-linear model is simply the linearized version of the MM5v3 model to which these techniques have already been adapted. Appying the parallelization techniques to the adjoint model however will raise new issues. In particular, it will be necessary to modify loops and index arithmetic to re-establish “Owner Computes” and to eliminate instances of false recursion introduced with the creation of the adjoint model. In addition there will be issues to contend with related to the disk I/O for the trajectory information.

All new code developed will conform to FORTRAN 77 standards. The software development approach conforms to Level 2 software engineering processes in accordance with the Software Engineering Institute Capability Maturity Model (Paulk et al. 1991).

4. CURRENT STATUS

The scalable MM5v3 tangent-linear model has been completed and tested for accuracy. It was successfully tested on DEC Alpha Cluster, SGI Origin 2000, and an IBM SP3 machines. As of the writing of this paper, scaling tests are on-going and should be available to be presented at the conference. The MM5v3 compatable adjoint model is still under development with although it is expected to be completed by the time of the conference.

5. FUTURE PLANS

Table 1 shows the milestones of the project as well as their planned completion dates. In addition to the development of the scalable MM5v3 adjoint model, in the next year work will be carried on other efforts. One effort will be to develop an incremental driver for the 4DVAR system (see Courtier et al. 1994). The benefit to adapting the incremental approach is that it will allow potential users of the 4DVAR system to use it in conjunction with other forecast models. For example, users of the Weather and Research Forecast (WRF) Model will be able to use the MM5v3 4DVAR system in conjuction with the WRF model to produce optimal intial conditions. It is envisioned that this will be a common scenario during the time from when the WRF is available until its own 4DVAR system is ready. Another important aspect of the project will be the development and integration of satellite observation operators, since one of the motivations of this project is to help produce accurate analyses in regions where conventional observations are limited. In addition, after the initial system has been developed, additional physics packages will be made available in the 4DVAR system. including the Blackadar (1979) planetary boundary layer, Dudhia cloud microphysics, and Grell (1993) cumulus parameterizations

6. SUMMARY

The Air Force Research Laboratory is leading an effort produce a complete 4DVAR system based on the current version of MM5. The 4DVAR system will initially include relatively simple physics packages and will be structured within an incremental driver so that it can be run in conjunction with more sophisticated forecast models. Forward observation operators for a number of satellites and sensors will also be part of the system. The entire system will be optimized to run on distributed memory computers with multiple processors. It is anticipated that a sufficient degree of scaling will be achieved to make the developed system vialble for use at operational centers. The project is to be completed by January of 2003 at which time the new 4DVAR system is expected to be made available to the entire user community.

Acknowledgements: This work was supported by the Department of Defense (DoD) High Performance Computing (HPC) Modernization Office's Common High Performance Computing Software Support Initiative. Additional support was provided from the DoD Major Shared Resource Center at Wright-Patterson AFB for time on the IBM SP3.

REFERENCES

Blackadar, A. K., 1979: High resolution models of the planetary boundary layer. *Advances in Environmental Science and Engineering, Vol. 1,* J. R. Pfallin and E. N. Ziegler, Eds. Gordon and Breach, 50-85.

Giering, R., and T. Kamiski, 1998: Recipes for adjoint code construction, *Trans. Math. Software,* **24**, 437-474.

Grell, G. A., 1993: Prognostic evaluation of assumptions used by cumulus parameterizations. *Mon. Wea. Rev.*, **121**, 764-787.

Kuo, H.-L., 1974: Further studies of the influence of cumulus convection on large-scale flow. *J. Atmos. Sci.,* **31**, 1232-1240.

LeDimet, F. X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Therorectical aspects. *Tellus*, **38A**, 97-110.

Michalakes, J., 1998: The Same Source Parallel MM5. Proceedings, Second International *Workshop on Software Engineering and Code Design in Parallel Meteorological and Oceanographic Applications*, Scottsdale AZ.

Nehrkorn, T., G. D. Modica, M. Cerniglia, and F. H. Ruggiero, 2001: MM5 adjoint development using TAMC: Experiences with an automatic code generator. Preprints, *14 Conf. on Numerical Weather Prediction,* Fort Lauderdale, FL.

Paulk, M.C., B. Curtis, M. B. Chrissis, and C. V. Weber, 1991: Capability maturity model for software, Software Engineering Institute, CMU/SEI-91-TR-24, ADA240603.

Rabier, F., H. Järvinen, E. Klinker, J.-F. Mahfouf and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation I: experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.*, **126(A)**, 1143-1170.

Zou, X., Q. Xiao, A. E. Lipton, and G. D. Modica, 2001: A numerical study of the effect of GOES sounder cloud-cleared brightness temperatures on the prediction of Hurricane Felix. *J. Appl. Meteor.*, **40**, 34-55.

Zou, X., F. Vandenberghe, M. Pondeca, and Y. –H. Kuo, 1997: Introduction to adjoint techniques and the MM5 adjoint system, *NCAR Technical Note,* NCAR/TN-435-STR, National Center for Atmospheric Research, Boulder, CO, 110 pp.

Zou, X. and Y.-H. Kuo, 1996: Rainfall assimilation through an optimal control of initial and boundary conditions in a limited-area mesoscale model. *Mon. Wea. Rev.* **124**, 2859-2882.

Zou, X., Y.-H. Kuo, and Y.-R. Guo, 1995: Assimilation of atmospheric radio refractivity using a non-hydrostatic adjoint model. *Mon. Wea. Rev.*, **123**, 2229-2249.

Table 1. Planned completion dates of major project milestones

Milestone	**Completion Date**
Complete vector version MM5v3 compatible tangent-linear model	Dec 00
Complete scalable version MM5v3 compatiable tangent-linear model	Mar 01
Complete vector version MM5v3 compatible adjoint model	Jun 01
Complete incremental driver for 4DVAR system	Oct 01
Complete integration of additional satellite observation operators	Jun 02
Complete scalable version of MM5v3 compatible adjoint model	Jun 02
Completion of additional physics upgrade to the system	Jun 02
Release MM5v3 compatable 4DVAR system to user community	Jan 03