Development of an MM5-based Four-Dimensional Variational Analysis System for Distributed Memory Multiprocessor Computers

Frank H. Ruggiero
Air Force Research Laboratory, AFRL.VSBL
29 Randolph Road, Hanscom AFB, MA 01731-3010
Email: frank.ruggiero@hanscom.af.mil

George Modica, Thomas Nehrkorn, Mark Cerniglia
Atmospheric and Environmental Research, Inc.
131 Hartwell Avenue, Lexington, MA 02421-3126

John G. Michalakes
National Center for Atmospheric Research, NCAR/MMM
3450 Mitchell Lane, Boulder CO 80301

Xiaolei Zou
Florida State University
404 Love Building, Tallahassee, FL 32306

The Air Force Research Laboratory is leading a multi-year effort to develop an updated version of the National Center for Atmospheric Research/Penn State University Fifth Generation Mesoscale Model (MM5) four-dimensional variational analysis (4DVAR) system. The main objective of this update is to optimize the code to run on distributed memory computers with the goal of achieving substantial speedup in wall-clock time processing and allowing mesoscale 4DVAR to enter the realm of operational forecasting. The previous version of the MM5 4DVAR system (Zou et al. 1997 [4]) is coded for single processor computer architectures and its non-linear, tangent-linear, and adjoint components are based on version 1 of MM5. To achieve the project goal, two major steps need to be accomplished. First, a version of an MM5 4DVAR based on version 3 of the model has to be created. The reason for this is to be able to take advantage of the existing parallelization mechanisms (Michalakes 2000 [2]) already in place for the version 3 non-linear model. The second step is to then install the parallel code architecture into the 4DVAR system's tangent-linear and adjoint components.

The first step has been completed. A complete 4DVAR system based on MM5v3.4 has been developed and tested. The tangent-linear and adjoint components were developed using the Tangent Linear and Adjoint Model Compiler (TAMC) automatic adjoint code generator. TAMC (Giering and Kaminski 1998 [1]) is a source-to-source translator that generates FORTRAN code for the tangent-linear model or adjoint from the FORTRAN code of the non-linear model. It is possible to incorporate the TAMC as part of the model compilation process, requiring the maintenance of just the non-linear model code. However in this project, TAMC is used as a development tool only. This results in separately maintained tangent-linear model and adjoint versions of the code. This approach makes it possible to minimize changes to the MM5 non-linear model code as supported by the National Center for Atmospheric Research but does require a mixture of manual and automatic code generation. Modules of the tangent-linear model and adjoint were created and tested individually. Unit testing was done by finite differencing comparisons with the non-linear model for the tangent-linear model and direct comparisons with the tangent-linear model for the adjoint. When all modules were assembled, a final unit integration test was carried out for both the tangent-linear model and adjoint. Additional testing was carried out by comparing the results from the v3 MM5 4DVAR with those from the v1 MM5 4DVAR.

The new MM5 4DVAR system contains most of the limited set of physics packages that are in MM5v1 based 4DVAR system. In particular, there are tangent-linear model and adjoint routines for bulk boundary layer processes, convection, and explicit, grid-scale precipitation. Additional physics options planned will be added as well as the tropical cyclone bogussing scheme of Zou and Xiao (2000 [3]).

Work on parallelizing the 4DVAR code for distributed memory computers is on going. The non-linear model, which is MM5v3.4, is already structured to run on distributed memory computers. Parallel code architecture for the tangent-linear model has been installed and tested. Using a problem with 84 x 75 x 27 grid points, running on an IBM SP P3, and using 24 processors, the tangent-linear model ran over 20 times faster in terms of wall-clock time than when running on one processor. Work to improve the wall-clock speed up is continuing, with the focus being on trying to eliminate the I/O bottleneck that occurs from reading in the base state information at every time step. A parallelized version of the adjoint is currently under development and is expected to be complete by mid-2002. Release of a beta version of the fully parallelized MM5v3 based 4DVAR is scheduled for late 2002.

References

[1] R. Giering and T. Kaminski. Recipes for adjoint code construction. *ACM Trans. Math. Software,* 24(4):437-474, 1998.

[2] J. Michalakes. The same source parallel MM5. *Sci. Progr.,* 8:5-12, 2000.

[3] X. Zou and Q Xiao. Studies on the intialization and simulation of a mature hurricane using a variational bogus data assimilation scheme. *J. Atmos. Sci.* 57:836-860, 2000.

[4] X. Zou, W. Huang, and Q. Xiao. A user's guide to the MM5 adjoint modeling system. Technical Note 437+IA, NCAR, Boulder, CO, 1997.