# P1.50 APPLYING PORTAL TECHNOLOGIES TO ENSEMBLE MODELING OF CONVECTION ON THE GRID

Brian F. Jewett[1], Robert B. Wilhelmson and Mohan K. Ramamurthy
Atmospheric Sciences Department & NCSA, University of Illinois

Jay Alameda, Shawn Hampton, Albert L. Rossi, and David Tcheng
National Center for Supercomputing Applications (NCSA)
University of Illinois at Urbana-Champaign

## 1. INTRODUCTION

Numerical investigation has evolved into a diverse and complicated task. Its many components include experimental design, parameter definition, job management, code execution, output interrogation, data mining, visualization and data archival. The paradigm of carrying out such studies is shifting to one which incorporates all elements into an efficient, powerful and comprehensive system. Emerging portal technologies and Grid tools under development (Alameda, 2002) facilitate such integration and make computationally demanding studies possible.

Atmospheric and computational scientists at the University of Illinois are working together to apply portal technologies to ensemble modeling studies on the Grid. Two computationally intensive atmospheric sciences projects form the scientific application testbed for this portal facility: a thunderstorm cell interaction study (Jewett et al. 2002) and a hurricane ensemble modeling effort (Ramamurthy et al. 2002). We will discuss the convective storm study in some detail, briefly outline the hurricane ensemble work, and describe our plans for integration with portal services

## 2. NUMERICAL STUDY OF STORM INTERACTION

Our initial computational and scientific challenge is part 3 of a tornado outbreak study (Jewett et al. 2000, Lee et al. 2000). During April 19, 1996, over 30 tornadoes struck Illinois and nearby states. Prior to tornado formation, many examples of thunderstorm splitting and merging were observed by radar. Our investigation into this thunderstorm cell behavior has taken the form of numerical modeling, where a pair of simulated storms form and interact over several hours. The Weather Research and Forecasting (WRF; Michalakes 2001) model is used for the simulations. The parameter space formed by the range of environmental conditions (soundings) and storm size, intensity, and relative position is large. As a result, the computational cost is significant, and the mechanics of making *hundreds* of such runs is formidable.

The work to date has taken two forms. In the first, demonstrated at the 2001 Denver Supercomputing Conference, the parameters describing the initial storm environment and properties were specified randomly. In collaboration with job management, visualization and data mining (D2K, Welge et al. 2002) specialists at NCSA, a system for problem setup (in Denver), execution and visualization (across the network at NCSA) and evaluation, display and mining (at Denver) was tested. While the details of implementation have changed, this project formed the basis for future efforts.

In the second stage of this work, we restricted the parameter space to just the relative storm location. A mesh of initial storm positions was created (Fig. 1), defining a matrix of 232 simulations. After each simulation was completed, it was evaluated in a limited fashion, resulting in a “score” of results such as

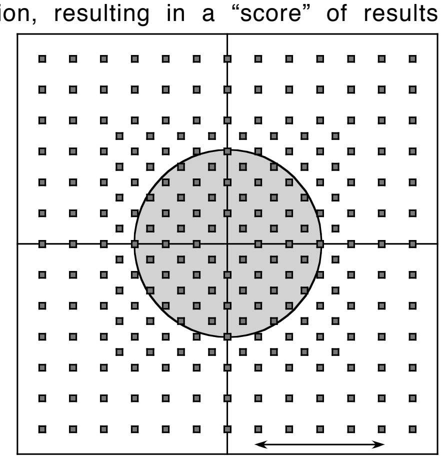

**Fig. 1**: Storm pair matrix. Primary storm cell is at middle; secondary cell at one of marked positions.

[1]Corresponding author address:
Dr. Brian F. Jewett
212 Atmospheric Sciences Bldg
105 S. Gregory St., Urbana IL 61801
email: bjewett@ncsa.uiuc.edu

peak surface vorticity, etc. When one such metric was plotted in the same parameter (horizontal position) space as Fig. 1, it revealed significant, and detailed response of storm behavior to the original storm pair orientation (Fig. 2).

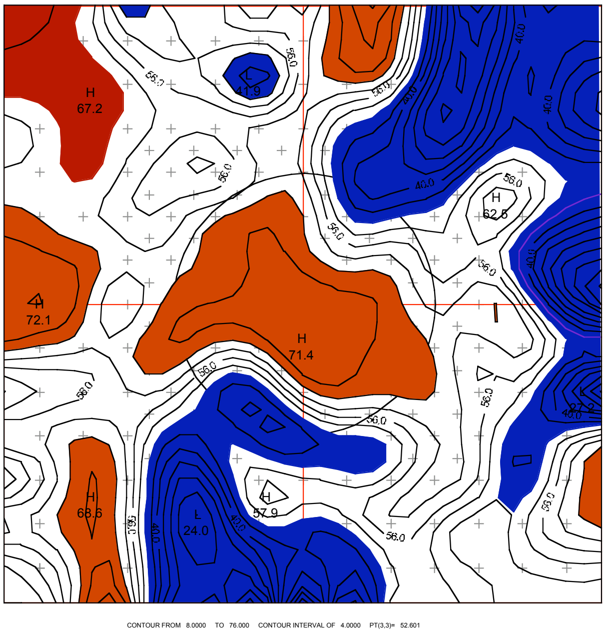


**Fig. 2**: One evaluation of all 232 storm simulations, plotted as in Fig. 1. The contours show the maximum time period (min; red is longer) for which any individual rotation center exceeded vorticity of 0.015 $s^{-1}$.

In the above work, a rich parameter space is indicated - one which we have only begun to explore (http://pampa.ncsa.uiuc.edu/~jewett/Apr19/SLS02). Conclusions can be drawn *now* about our computational methodology, however. Our job management system was quite crude. We interrogated a *very* small fraction of the simulation data of 4 GB per model run, or 880 GB for all cases. The same 0.8 TB data, while technically complete, is far from readily available. Our visualization has been similarly limited, and no extensive data mining of this data set has been done. Well-defined metadata (of input parameters or output results or conclusions) does not exist. Far more scientific work is needed to span the model parameter space, and thus the range of observed convective environments and thunderstorm cell behavior. Making the model simulations, extracting critical information, and thoroughly understanding the results is a computational challenge that is best met collaboratively.

## 3. HURRICANE ENSEMBLE MODELING

An extensive mesoscale ensemble prediction system (MEPS) is under development at the University of Illinois (Ramamurthy et al. 2002) as part of a separately funded NSF effort and related NCSA Faculty Fellow project. The MEPS system is currently being applied to intense mid-latitude winter cyclones as well as tropical cyclones. Specifying the "parameter space" in the ensemble modeling effort requires assessing, understanding, and then incorporating the underlying *uncertainties* into the ensemble design. Uncertainties lie within the initial, lower and lateral boundary conditions, model physics (e.g. boundary layer parameterization), and model numerics. In hurricane prediction, initial state uncertainties include the hurricane vortex position, intensity, and structure. The result is a wide range of modeled behavior (Fig. 3).

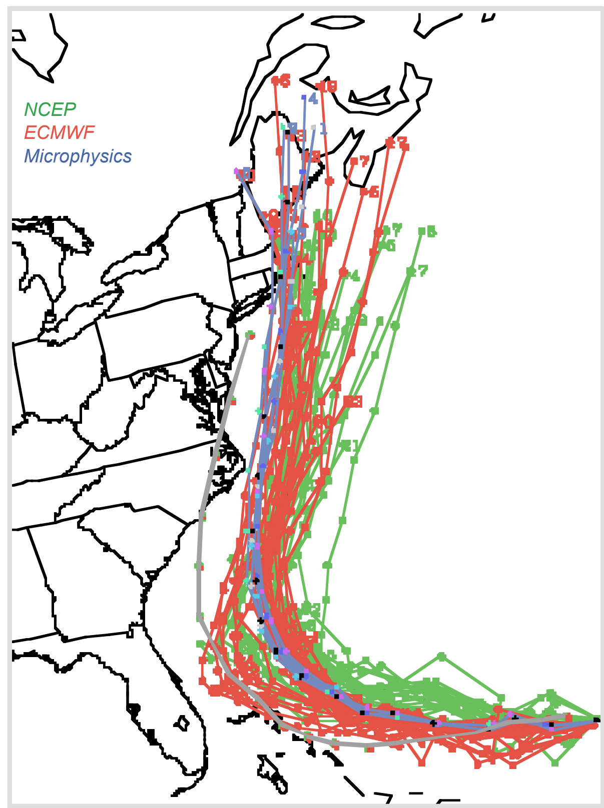


**Fig. 3**: Simulations of hurricane Floyd (1999) resulting from variations in initial conditions and model physics.

While conceptually quite different from the convective study noted earlier, many of the *workflow* issues are similar. Parameter and initial and boundary condition specification, job management for hundreds of simulations (per hurricane), visualization, data mining, and ensemble refinement are necessary components of this work, which heretofore has been carried out manually, at significant time cost for a graduate student. As in the convective effort, a new paradigm for carrying out such computationally challenging problems is necessary.

## 4. FUTURE: WORKFLOW ON THE GRID

Our computational problems have grown, but our computational methodology has not kept pace. The research infrastructure being developed will allow us to efficiently design and carry out our studies of severe convection and mesoscale ensemble prediction and ultimately refine our research approach and parameter space. In association with the Modeling Environment for Atmospheric Discovery project (MEAD; Wilhelmson et al. 2002), a complete workflow environment is being created. Use of Grid resources will allow us to make hundreds of simulations, transfer data to the desired repositories and carry out post-processing and data mining on a variety of platforms, as best fits the computational needs and available resources. A comprehensive job creation, management, execution and post-processing system is under development at NCSA and collaborating institutions. A metadata "Project Repository" is integrated into the system, allowing storage and identification of project design and input as well as output data, visualization and analysis results. These facilities will not only allow us to carry out our research, but will facilitate a much greater accessibility to, and effective interrogation of, our data. The design and use of portal technologies in our research will be discussed at the Conference.

## 5. ACKNOWLEDGEMENTS

The convective storms work is supported by NSF under grant ATM-9986672. Ensemble modeling research (Ramamurthy, P.I.) is funded by NSF (ATM-9730385). Computing and other support was provided by the National Center for Supercomputing Applications. The authors thank their NCSA collaborators not mentioned elsewhere in this article: Loretta Auvil, Randy Heiland, and Thomas L. Redman.

## 6. REFERENCES

Alameda, J., 2002: Alliance Science Portals. Information: http://portals.ncsa.uiuc.edu/

Jewett, B. F., R. B. Wilhelmson, and B. D. Lee, 2002: Numerical simulation of cell interaction. *Preprints, 21st Conf. on Severe Local Storms*, San Antonio, Amer. Meteor. Soc., 353-356. Preprint and other materials: http://pampa.ncsa.uiuc.edu/~jewett/Apr19/SLS02

Jewett, B. F., B. D. Lee and R. B. Wilhelmson, 2000: Initiation and evolution of severe convection in the 19 April 1996 Illinois Tornado Outbreak. *Preprints, 20th Conf. on Severe Local Storms*, Orlando, 74-77.

Lee, B. D., B. F. Jewett and R. B. Wilhelmson, 2000: Supercell differentiation and organization for the 19 April 1996 Illinois Tornado Outbreak. *Preprints, 20th Conf. on Severe Local Storms*, Orlando, 222-225.

Michalakes, J., S. Chen, J. Dudhia, L. Hart, J. Klemp, J. Middlecoff and W. Skamarock, 2001: "Development of a next generation regional weather research and forecast model" in *Developments in Teracomputing*: Proceedings of the 9th ECMWF Workshop on the Use of High Performance Computing in Meteorology. (www.mmm.ucar.edu/mm5/mpp/ecmwf01.htm)

Ramamurthy, M. K., B. F. Jewett and H. Liu, 2002: Ensemble methods applied to hurricane track and intensity forecasting. *Preprints, 25th Conf. on Hurricanes and Tropical Meteorology*, San Diego, Amer. Meteor. Soc., 216-217. Preprint and other material: http://cyclone.atmos.uiuc.edu/ensemble

Welge, M., Team leader - NCSA Automated Learning Group: Data to Knowledge program. Information: http://www.ncsa.uiuc.edu/Divisions/DMV/ALG/d2k

Wilhelmson, R. B., K. Droegemeier, S. Graves, M. Ramamurthy, D. Haidvogel, B. Jewett, J. Alameda, and D. Gannon, 2002: Modeling environment for atmospheric discovery. *Preprints, 19th Conf. on Interactive Information Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, Baltimore, Amer. Meteor. Soc. **5.9**, this volume.