Technical Report to the U. S. Department of Energy
Office of Energy Research

Special Research Grant: CHAMMP

Grant # DE-FG05-91-ER61219

EXPLOITATION OF PARALLELISM

IN CLIMATE MODELS

Principal Investigators

Ferdinand Baer

Joseph J. Tribbia

David L. Williamson

Period: 1 March 1992 - 28 February 1993

Introduction

The U. S. Department of Energy (DOE) through its CHAMMP initiative, hopes to develop the capability to make meaningful regional climate forecasts on time scales exceeding a decade, such capability to be based on numerical prediction type models. We propose research to contribute to each of the specific items enumerated in the CHAMMP announcement (Notice 91-3); i.e., to consider theoretical limits to prediction of climate and climate change on appropriate time scales, to develop new mathematical techniques to utilize massively parallel processors (MPP), to actually utilize MPP's as a research tool, and to develop improved representations of some processes essential to climate prediction.

In particular, our goals are to:

- Reconfigure the prediction equations such that the time iteration process can be compressed by use of MMP architecture, and to develop appropriate algorithms.

- Develop local subgrid scale models which can provide time and space dependent parameterization for a state-of-the-art climate model to minimize the scale resolution necessary for a climate model, and to utilize MPP capability to simultaneously integrate those subgrid models and their statistics.

- Capitalize on the MPP architecture to study the inherent ensemble nature of the climate problem. By careful choice of initial states, many realizations of the climate system can be determined concurrently and more realistic assessments of the climate prediction can be made in a realistic time frame.

To explore these initiatives, we will exploit all available computing technology, and in particular MPP machines. We anticipate that significant improvements in modeling of climate on the decadal and longer time scales for regional space scales will result from our efforts.

This report summarizes the activities of our group during the first one and one-half years' effort to meet the objectives stated in our proposal. We will comment on the three research foci, time compression studies, subgrid scale model studies, and distributed climate ensemble studies and additional significant technical matters.

Time Compression Studies

The techniques necessary to utilize massively parallel processors for solving the climate prediction problem exploiting *established* numerical processes and state-of-the-art models seem now to be well understood and have been explored to a significant extent. Unfortunately, even if one optimizes current and near future MPPs, the need for adequate computer processing time may not be met. This limitation is based on the fact that currently models are solved as a marching problem in time, and no matter how fast we can compute results in a given time step, we are limited in time by the number of steps needed to march into the future; and climate change requirements indicate that the prediction time must be increased to the order of tens of years. One possible solution to this bottleneck is to reconstruct the forecast system so that many time levels may be computed concurrently, i.e., in a given machine cycle provided sufficient processors are available.

Our purpose is then to use as few time cycles as possible to complete a model integration. How can the time cycles be compressed? Let us define the "computing cycle" to be that time required to do once all the calculations which must systematically be repeated to complete the

entire calculation. On a true serial machine which can handle just one computation at a time (non-vector), the computing cycle would be the time for that operation. On a massively parallel processor with unlimited processors, the computing cycle would include all the calculations which would not need repetition by their dependence on previous calculations. For conventional marching problems, the minimum computing cycle could be one complete time step. It is this computing cycle that we attempted to approach as a limit for an integration.

We have approached this problem by first considering the barotropic vorticity equation (BVE) in both spectral and interaction coefficient form. This simplification is based on the desire to *uncomplicate* the problem of nonlinearity so that the basic procedure can be identified. The relevant equation which is listed as equation (1) in our proposal is as follows:

$$\frac{d\psi_i}{dt} = \sum_{j,k} \psi_j \psi_k I_{i,j,k}$$

Note that one can differentiate this equation as many times as desired and by iteration can represent the highest derivative by a series of the coefficients (ψ_i). Now consider an expansion of any coefficient by Taylor's series,

$$\psi_i^{(n\Delta t)} = \psi_i(o) + \frac{d\psi_i}{dt}(o) \, n\Delta t + \frac{d^2\psi_i}{dt^2}(o)\frac{(n\Delta t)^2}{2!} + \cdots$$

3

and take the series for n terms. In terms of initial values, $\psi_i(o)$, a series can be expressed in terms of products and powers of the initial values and the appropriate constants $I_{i,j,k}$. The particular examples we have focussed on is a low-order version of the BVE which has the form,

$$\frac{d\psi_i}{dt} = a_i\psi_j\psi_k \qquad 1 \le i,j,k \le 3.$$

This represents three equations where we permute the indices i, j and k such that j = i + 1, k = j + 1 and i + 3 = i, etc. The prediction system, using the Taylor's expansion can be given as

$$\psi_i^{n\Delta t} = \sum_{l=1}^{n+1} A_l X_l(\Delta t)^{l-1}$$

and the A_l are known constants of n, ℓ. The X_l are calculated from the initial values $\psi_1(o)$, $\psi_2(o)$, $\psi_3(o)$ in the form

$$X_\ell = \sum_{m=o}^{\ell-1} G^q[\psi_1(o),\psi_2(o),\psi_3(o)]L_m^{(\ell)} .$$

We have shown that we can perform the entire calculation in only a few machine cycles of a MPP of sufficient capacity by utilizing as many processors as possible, whereas the conventional computational procedure would take a minimum of n cycle on an MPP and many more on a serial machine. Since our last report, we have extended our calculations beyond the simple three-component system to explore the sensitivity for higher order systems. Additionally, we have

4

programmed the shallow water equations in a similar form and are experimenting with the bounded derivative method to inhibit the rapid propagation of gravity waves.

Subgrid Scale Model Studies

All nonlinear models need some form of closure on space scales to allow a computation to proceed. This is effected by subgrid scale parameterization and is done in a variety of ways, principally designed however to avoid computational difficulties. Experiments with systems in high rotation such as the atmosphere and including significant external energy input over a broad scale range to include phenomena such as boundary layer friction and convection, indicate that incorporating time independent statistics for subgrid scale forcing is not adequate to provide successful predictions, in particular of the intermediate "regional" scales.

Regional scales must be studied independently of the global scales. The principal forces which the subscale domain imposes on the global model and the time scales over which the subgrid scale can be averaged must be established. Since the subgrid scale domain is locally space dependent, domains which are sufficiently homogeneous to be included in any one model must be identified and a set of such regionally dependent models developed. Due to high frequency properties of these models, non-hydrostatic effects should be included.

(a)   *Simple Model Analysis*

To understand the impact of subgrid scale parameterization on the predicted scales, we have performed a number of calculations using the BVE. Establishing a reference triangular truncation as our "real world" (e.g., T8O), we have compared calculations of the larger scales with reduced truncations (T63, T42, etc.) and including a variety of parameterizations for the scales which have been truncated to the reference. The model has been run for pure dynamics,

5

with topography and with simulated baroclinic instability (forcing). Parameterizations used have included -3 power law energy statistics with random phase, statistics for the subgrid scales based on correlations with large scale shear, and statistics which allow the subgrid scales to decay randomly but smoothly, as observed from the reference integrations.

The results of runs with a variety of initial states suggest that none of the parameterization schemes can maintain the evolution of the large scales comparable to the reference model, and indeed, the completely truncated model without parameterization gives results quite similar to those for the same truncation but with parameterization of subgrid scales included. Since the analyses are for short term forecasts (up to ten days), the results seem to confirm Lorenz' original hypothesis on error propagation, despite the effort to compensate with parameterization.

Our continuing efforts with this experiment include extended integrations to climate scales and preparation of a three-dimensional model, since the statistics of the BVE are not realistic for the actual atmosphere. However, if the BVE cannot be parameterized successfully, more complex models may be in serious trouble.

(b)   *Global Climate Model Effects*

We have begun to conceptualize a global climate model of the atmosphere with its associated interactive models and truncated at a reasonable scale to include meaningful regional climate areas. A set of regional models is concurrently under development to encompass spatial scales not included in the global model. These models are to include the uniqueness and homogeneity of the local region they represent, and their number in sum spans the entire space of the domain of the global model.

Postdoctoral Research Associate Mark Taylor has been investigating climate simulations with a simplified GCM and simple subgrid parameterization in order to examine the questions of spatial homogeneity of the forcing due to subgrid scale processes. He has examined the forcing spectra in space and time to ascertain correlation lengths and times. He has further examined the scalability of parameterizations and the influence of performing grid point parameterized physics on grids coarse enough to be resolved by the spectral resolution. This is in contrast to the common practice of performing physics computations on the de-aliasing transform grid.

Distributed Climate Ensemble Studies

The inherent ensemble nature of the climate problem suggests the use of machines with massively parallel architecture. The traditional manner in which the climate ensemble is generated in general circulation models mimics the generation of climate realizations in the physical climate system, i.e., the sequential generation of daily weather from long integrations through many seasonal cycles leading to an ensemble of realizations of particular months, seasons and years. The climatic mean of any particular time period is then the arithmetic average of the ensemble of realizations of this time period contained in the long, sequential integration. Natural climate variability is the deviation from the climate mean of the realizations of the time period within the climate ensemble. Massively parallel architecture allows us to circumvent the generation of the ensemble elements of the climate distribution by integrating in parallel a number of members of the climate ensemble. For example, if N is the total number of processors and a model of the climate system can be efficiently integrated on a subset K of the processors, then $L = N/K$ realizations of a year of climate data can be generated simultaneously from

judiciously chosen initial states. Each subsequent year of simulation results in L additional realizations of a year of the climate ensemble.

We have developed low-resolution/simple models for the integration of the climate system and begun porting these models to the parallel machines at our disposal (the CM5 at LANL and the IBM cluster at NCAR using PVM). These studies are almost always necessary to better understand the meaning and sensitivity of the results of the state-of-the-art models and for exploratory research into the interactions of the various components of the climate system. We have also begun to address the "judicious choice" of initial states. We are constructing initial states using the most important modes of climate variability (the leading EOFs). By constructing initial states in this fashion, we guarantee the independence of the ensemble elements as they evolve in time and thus sample different portions of the climate probability distribution. We are currently investigating the proper method of ensuring that climate ensemble maintain stationarity at least with respect to the mean and the covariance structure explained by the leading EOFs as a means of adaptively modifying and increasing the ensemble.

Science Team Activities

The science team for this project is now completely staffed. We hired Francois Thibaud from Rutgers University as a programmer in April, 1992. He worked part-time on the project at Rutgers for several months, acquainting himself with MPP architecture which was available to him there. In July, he moved to NCAR and began full-time employment.

In September, 1992, Mark Taylor completed his Ph.D. requirements from the Courant Institute of New York University and came to NCAR as a post-doctorate to join our group.

Two students, Bing Zhang and Eric Deweaver, are also employed as Research Assistants in the Department of Meteorology at the University of Maryland at College Park.

In July, 1992, the entire team met at the University of Maryland at College Park to discuss the project. We met again in Boulder in January, 1993. During the year, Professor Baer attended the CHAMMP Science Team Meeting in March 1993, attended the CAS93 meeting in Santa Fe, New Mexico in July, 1992, discussed the project while visiting SMHI in Norkoping, Sweden in September, 1992, and attended the CHAMMP workshop on alternative strategies to GCMs in Oakland, California in February, 1993. Dr. Tribbia attended the CHAMMP Science Team Meeting in March, 1992 and the CHAMMP workshop in Oakland in February, 1993.

Summary

The project is proceeding as anticipated in the original proposal and progressing satisfactorily. Our accomplishments to date are expanding as we study the details of the time compression problem, develop the subgrid scale models needed for parameterization, and advance the distributed climate ensemble strategy. We expect the project to continue as planned during the upcoming year and look for significant results to ensue.

END

DATE FILMED
6/3/93