Building an Advanced Climate Model

Program Plan for the CHAMMP Climate Modeling Program

December 1990

U.S. Department of Energy
Office of Energy Research
Office of Health and Environmental Research
Atmospheric and Climate Research Division
Washington, D.C. 20585

*Computer Hardware,
Advanced Mathematics,
and Model Physics* **MASTER**

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Foreword

The issue of global warming and related climatic changes from increasing concentrations of greenhouse gases in the atmosphere has received prominent attention during the past few years. Calls for action to curb emissions of the greenhouse gases can be heard at both the national and international levels, while the scientific debate continues regarding the reality and severity of the predicted climatic change.

In 1978, when only a small number of scientists and science managers were concerned about the greenhouse effect, the Department of Energy (DOE) started the Carbon Dioxide Research Program to estimate the climatic change from the increasing concentration of carbon dioxide in the atmosphere. Since then the Program has studied many facets of the issue, from the carbon cycle to climate diagnostics and from effects on vegetation to impacts on resources. Currently, the Carbon Dioxide Research Program is DOE's principal contribution to the U.S. Global Change Research Program (USGCRP), which is coordinated by the Committee on Earth and Environmental Sciences (CEES) of the Office of Science and Technology Policy (OSTP).

One of the principal objectives of the USGCRP is to develop integrated conceptual and predictive models of the Earth system, including the general circulation models that are currently the main predictive tools of climatic change. The DOE research program places special emphasis on accelerating improvement of the models used to predict global climatic change. This emphasis encompasses study of the timing, magnitude, and regional characteristics of this change. The Computer Hardware, Advanced Mathematics, and Model Physics (CHAMMP) Climate Modeling Program is designed to contribute directly to this rapid improvement.

The program has two aspects that are inescapably coupled and must be conducted in an increasingly comprehensive series of coordinated projects. First, advanced climate models must be developed that dramatically increase the computational throughput of climate model simulations by combining state-of-the-science supercomputing systems with advances in numerical methods and model physics. These new generation models must be used to address three critical problems: more accurate prediction of regional climatic changes that may occur over decadal to centennial periods, determination of the limits to climate predictability imposed by the inherent variability of climate as a dynamic system, and development of realistic schedules for achieving improved predictions of climatic change.

Many individuals drawn from government, national laboratories, universities, and industry have contributed to the preparation of this plan. A draft version of this plan was distributed for critical review to the CEES agencies, to the scientific community, and to a special technical oversight panel including the JASON Group. Comments from all of these sources have been incorporated into this final version of the plan.

We believe that the CHAMMP Climate Modeling Program will serve the CEES objectives in global change research and will support the DOE mission of formulating a national energy strategy that takes into account the potential for global climatic change.

Dr. Ari Patrinos, Acting Director
Atmospheric and Climate Research Division
Office of Health and Environmental Research
Office of Energy Research, ER-76
U.S. Department of Energy
Washington, DC 20545

Contents

Foreword.....	iii
Summary.....	ix
Near-term tasks (1 to 2 years)	ix
Intermediate-term tasks (3 to 6 years).....	ix
Extended-term tasks (7 to 10 years).....	x
1. Background.....	3
2. CHAMMP Rationale.....	7
3. Goal of the CHAMMP Climate Modeling Program	17
4. Program Plan.....	23
Focus 1: Model Physics.....	23
Focus 2: Advanced Mathematics and Computational Methods.....	24
Focus 3: Computer Hardware.....	26
5. Program Management.....	29
6. Linkages.....	33
7. Resource Requirements.....	37
Computer Hardware.....	37
Advanced Mathematics and Computational Methods	39
Model Physics.....	39
Project Coordination	40
Appendices	
A. Brief History of the Department of Energy Carbon Dioxide Research Program	43
B. Prospects for Large-Scale Computing.....	47
Hardware Characteristics and Performance.....	47
Climate Model Requirements.....	48
Numerical Methods	49
Parallel Programming Tools	50
Software Portability and Performance	50
Examples of Early Applications.....	51
Numerical linear algebra	53
Finite element method.....	53
Hydrodynamics	53
Preliminary climate model results.....	53

C. Supercomputing in the Department of Energy.....	55
Computational Studies at DOE	56
Current Supercomputing Facilities at DOE	56
Research and Development on Future Computational Resources.....	56
Hypercubes.....	57
Shared-memory machines	57
Special Graphics machines	57
Summary.....	57
D. Data Storage and Visualization	61
References.....	65
Acronyms and Abbreviations.....	69

Figures and Tables

Figures

1. Schematic illustration of the processes and typical vertical gridding in the current generation of global climate models. Fully coupled versions of the atmosphere-ocean-land components are just beginning to become available. 9
2. Global distribution of GCM grid points (the boxes are outlined in dots with the grid center shaded in) over land areas using 4° latitude by 5° longitude resolution, which is typical of that used by climate modeling groups. The same grid-point density extends over the oceans. 9
3. Distribution of land grid points over the western United States (the crosses indicate the centers of grid zones) for a climate model using a typical resolution of 4° latitude by 5° longitude. A cross section of the central California grid box is shown to illustrate the limited regional resolution provided by the current generation of global climate models. 12
4. Comparison of the *peak* 64-bit floating point performance (Mflops) of supercomputers and microprocessors versus time. In comparing these numbers, it is also important to evaluate the relative success of codes in reaching toward peak performance (see Dongarra, 1989). 19
5. CHAMMP management structure 30
6. CHAMMP conceptual design. 34

Tables

1. Current versus needed capabilities of climate models. 11
2. CHAMMP resource requirements (\$M/yr). 38
3. Potential annual data storage requirements for simulations with advanced climate models. 39

Summary

Evaluating the importance of the future response of regional climates to the increasing atmospheric concentrations of greenhouse gases will require advanced climate models capable of much longer and more comprehensive simulations, having finer resolution, and incorporating significant improvements in their representations of the physical, chemical, and biological interactions of the climate system. Developing advanced climate models capable of such simulations will require use of the emerging generation of highly parallel computers. In addition, such models must use improved numerical algorithms, mathematical formulations, and software that take full advantage of extensive parallelism in order to further accelerate computational speed. These models must also be thoroughly verified against more comprehensive sets of field observations and should be used to investigate the limits of predictability.

The goal of the CHAMMP Climate Modeling Program is to develop, verify, and apply a new generation of climate models within a coordinated framework that incorporates the best available scientific and numerical approaches to represent physical, biogeochemical, and ecological processes, that fully utilizes the hardware and software capabilities of new computer architectures, that probes the limits of climate predictability, and finally that can be used to address the challenging problem of understanding the greenhouse climate issue through the ability of the models to simulate time-dependent climatic changes over extended times and with regional resolution.

The CHAMMP Climate Modeling Program will develop the advanced climate system models within a framework that takes full advantage of the hardware and software capabilities offered by emerging computer architectures, which are projected to permit an increase in the computational capabilities of climate models by a factor of 10^4 during the 1990s. The necessary tasks for the CHAMMP Program to achieve its goal include:

Near-term tasks (1 to 2 years)

- To demonstrate the scalable performance of present atmospheric and oceanic models by transferring these models to newly-available parallel computers.
- To explore the limits of the predictability of climate and to determine the potential for predicting regional climatic behavior and the potential for climatic change.
- To evaluate the efficiency of numerical algorithms, load balancing, data structure and retrieval, visualization, and other issues controlling the performance and analysis of models and component algorithms on a diverse set of high-performance computers as the basis for developing design criteria for the framework of the advanced climate model.

Intermediate-term tasks (3 to 6 years)

- To develop scalable versions of climate system models that incorporate improved representations of atmosphere, ocean, and land processes and that are capable of achieving performance in excess of 100 gigaflops.¹

¹One gigaflops (Gflops) equals 1 billion (10^9) floating point operations per second.

CHAMMP Program Plan

- To conduct experiments on climate predictability and sensitivity using models capable of distinguishing the separate contributions of internal and external causes of climate variability.
- To develop and implement improved numerical-solution, software, visualization, and analysis techniques for new high-performance computers.

Extended-term tasks (7 to 10 years)

- To assemble an optimized advanced climate system model having finer resolution and better verified process representations that is capable of achieving performance in excess of 10 teraflops² for extended simulations.
- To conduct simulations of the time-dependent greenhouse effect with a climate system model that couples atmosphere, ocean, and land surface systems and has a horizontal resolution of 100 km or less.
- To generate, display, and provide for analysis of regional estimates of the potential changes in climate resulting from time-dependent greenhouse forcing.

The CHAMMP Climate Modeling Program will collaborate and coordinate with other national and international initiatives on climate and computer-related research. These will include particularly the U.S. Global Change Research Program, organized by the interagency Committee on Earth and Environmental Sciences, and the Federal High Performance Computing Program, organized by the Office of Science and Technology Policy. The Federal High Performance Computing Program has identified climate as one of its "grand challenge" scientific problems.

CHAMMP will require increasing resources to complete its initial, model development, and model implementation phases over the next ten years. These resources must be in addition to resources required by other programs for development of process algorithms and study of climate feedbacks.

Appendix A provides a brief history of the DOE Climate Research Program. Appendix B describes prospects for future large-scale computing. Appendix C provides a history of supercomputing within DOE, which lays the basis for the CHAMMP Climate Modeling Program. Appendix D describes the data storage and visualization needs of CHAMMP.

²One teraflops (Tflops) equals 1 trillion (10^{12}) floating point operations per second.

Computer Hardware,
Advanced Mathematics
and Model Physics

Background

1. Background

The Department of Energy (DOE) has statutory responsibility for developing a National Energy Strategy that fully considers the environmental impacts and implications of energy-related activities. The potential climatic and environmental changes that may result from the continuing and projected emissions of carbon dioxide, methane, and other greenhouse gases will be a central focus in development of this strategy (see DOE, 1990).

Over the past 12 years, the research programs of DOE and other agencies have made significant progress towards advancing understanding and better defining the degree of uncertainty and the consequences associated with the potential for global climatic change (see Ap-

pendix A). The rising concentrations of the greenhouse gases have been well-documented. In recognition of the importance of these changes, innovative new research programs are being developed to ensure accurate representation of the relative roles of human activities and natural processes of the land, biosphere, and oceans in these cycles. Models of the global climate system have advanced to include realistic geography, the annual cycle of the seasons, varying cloud cover, and, very recently and embryonically, coupling of the ocean-atmosphere system. Results suggest that the projected greenhouse gas emissions may lead to a several degree increase in the global average temperature and to significant redistribution of precipitation during the next century (DOE, 1990).

Research during the 1980s also demonstrated that there remains considerable uncertainty in these estimates, particularly concerning changes in the regional climate and the rate of future climatic change. Furthermore, although the 1980s have been especially warm, the extent of global warming over the past century may be as much as a factor of three less than the upper estimates of the warming projected by currently available models. Further, when the results of similar models constructed by different groups are compared, there exist substantial differences in their estimates of changes in temperature and, especially, precipitation. As a consequence, although "significant change" seems a plausible conclusion, we do not know with sufficient accuracy how large the climatic changes will be, how rapidly the changes will occur, or how the changes will be distributed spatially. We also know virtually nothing about the potential changes in the frequency of extreme climatic events that are crucial to assessing the impacts of these changes. Indeed, we do not even know the extent to which fundamental limits to the predictability of climatic conditions (e.g., means, variances, cross-correlation statistics) might prevent models from giving detailed answers to these questions. These questions become even more uncertain as the horizontal resolution of models is reduced in the attempt to generate greater regional resolution.

To accelerate the identification, evaluation, and implementation of the long-range energy policies that may be needed to improve the global standard of living while ensuring environmental quality, the research effort into understanding climatic change and reducing uncertainties must be significantly intensified. As its contribution to the intensified interagency effort called for by the Committee on Earth and Environmental Sciences in *Our Changing Planet: The FY 1991 Research*

When the results of similar models constructed by different groups are compared, there exist substantial differences in their estimates of changes in temperature and, especially, precipitation. Although "significant change" seems a plausible conclusion, we do not know with sufficient accuracy how large the climatic changes will be, how rapidly the changes will occur, or how the changes will be distributed spatially. We also know virtually nothing about the potential changes in the frequency of extreme climatic events. We do not even know the extent to which fundamental limits to the predictability of climatic conditions might prevent models from giving detailed answers to these questions

in the regional climate and the rate of future climatic change. Furthermore, although the 1980s have been especially warm, the extent of global warming over the past century may be as much as a factor of three less than the upper estimates of the warming projected by currently available models. Further, when the results of similar models constructed by different groups are compared, there exist substantial differences in their estimates of changes in temperature and, especially, precipitation. As a consequence, although "significant change" seems a plausible conclusion, we do not know with sufficient accuracy how large the climatic changes will be, how rapidly the changes will occur, or how the changes will be distributed spatially. We also know virtually nothing about the potential changes in the frequency of extreme climatic events that are crucial to assessing the impacts of these changes. Indeed, we do not even know the extent to which fundamental limits to the predictability of climatic conditions (e.g., means, variances, cross-correlation statistics) might prevent models from giving detailed answers to these questions. These questions become even more uncertain as the horizontal resolution of models is reduced in the attempt to generate greater regional resolution.

To accelerate the identification, evaluation, and implementation of the long-range energy policies that may be needed to improve the global standard of living while ensuring environmental quality, the research effort into understanding climatic change and reducing uncertainties must be significantly intensified. As its contribution to the intensified interagency effort called for by the Committee on Earth and Environmental Sciences in *Our Changing Planet: The FY 1991 Research*

Plan of the U.S. Global Change Research Program (CEES, 1990), DOE is augmenting the research efforts within its Carbon Dioxide Research Program by intensifying its core climate modeling program and by adding two new scientific programs for the 1990s:

- a. ARM—the Atmospheric Radiation Measurement Program will improve understanding of cloud-radiation interactions, which have been identified as the most important cause of the uncertainty in estimating the magnitude of the potential global warming. ARM will provide a comprehensive data set for use in developing better algorithms for calculating atmospheric radiation, cloud amount, optical properties, and for representing convection and precipitation.
- b. CHAMMP—the Computer Hardware, Advanced Mathematics, and Model Physics Climate Modeling Program will develop advanced climate models capable of increasing the throughput capability of climate models by a factor of 10,000 or more within ten years. These models will be designed to carry out more realistic simulations with finer resolution and more comprehensive and detailed physical representations over longer periods than is possible with the current generation of models. Typical simulations should extend 200 years or more and include slowly changing atmospheric compositions, grid resolution of less than 100 km, and detailed treatment of ocean-atmosphere coupling.

Additional program initiatives are also in development by DOE.

To carry out these challenging programs, DOE will involve its national laboratories, universities, the private sector, and other government agencies in a concerted and coordinated effort. The success of these efforts will also depend on the success of many initiatives of other agencies, as explained in the USGCRP Plan. Programs of particular importance include the World Ocean Circulation Experiment (WOCE), which is particularly important for improving model representation of the oceans and to account for their influence on the rates of climatic change and trace gas uptake; the Tropical-Ocean Global Atmosphere (TOGA) program, which will increase understanding about interannual climate fluctuations; satellite measurement programs, which provide vital global data sets for investigating many physical and chemical processes; and expanded climate modeling programs that intensify research to improve representations of individual processes and to link climate system components. The Federal High Performance Computing Program (Executive Office of the President, 1989) will also play a vital role for CHAMMP by fostering the development and use of the much more powerful computers that are essential if climate models are to be substantially improved for use in predicting changes in climate.

CHAMMP will provide a numerical framework for integrating the results of these field, modeling, and computing programs and will contribute to the timely increase in understanding and simulation of the greenhouse issue. This effort is thus essential to justify moving from concept to initiation of modified energy policies. Accordingly, coordination and collaboration with the other agencies responsible for these initiatives will be an integral part of CHAMMP.

*Computer Hardware,
Advanced Mathematics
and Model Physics*

CHAMMP Rationale

2. CHAMMP Rationale

Simulation of the global climate is a particularly difficult challenge because of the wide range of time and space scales governing the behavior of the atmosphere, the oceans, and the surface. Although the prediction of the weather (the state of the atmosphere at a particular moment) requires knowledge of the state of the atmosphere at an earlier time, this startup (initial) condition controls atmospheric behavior for only days to weeks because the turbulent nature of atmospheric dynamics and the relatively large fluxes of solar and infrared radiation (compared to the small heat capacity of the atmosphere) overwhelm the inertial effect of the initial conditions. Hence, the limit of weather predictability is at most a few weeks. Over longer time scales, the exchanges of energy, momentum, and water vapor at the boundaries of the atmosphere (i.e., exchanges with land, oceans, and space) are the primary factors influencing atmospheric behavior. As a consequence, the time- and space-averaged behavior of the atmosphere (i.e., the climate) can be determined only by incorporating all of the components of the climate system in a coupled manner.

The coupling of the heat, momentum, and chemical-constituent budgets of the atmosphere, the ocean, the land surface, and the cryosphere greatly broadens the set of processes and the time scales over which interactions occur. Although the large thermal inertia of the ocean provides a stability that moderates the seasonal fluctuation caused by solar radiation, the climatic record suggests that, because of variations in its circulation, the ocean can also contribute strongly to variability on time scales from a few years to thousands (perhaps many thousands) of years. Because of the length of oceanic time scales, both past and present oceanic conditions can influence the future climate, requiring knowledge of both the initial and boundary conditions. Similarly for the land surface and the cryosphere, the soil and plant distributions (which influence albedo, hydrology, and biogeochemical cycles), and the amount and distribution of ice (which influence albedo, sea level, and surface elevation) that were created by past climates carry forward information that will affect future climate, as well as being interactive influences on present climate.

Analysis of past climatic behavior indicates that the state of the ocean-atmosphere system has changed substantially, sometimes in seemingly

regular cycles that are likely influenced by conditions at the boundaries (e.g., orbital variations affecting insolation) and sometimes in an apparently chaotic and nonlinear manner governed by processes internal to the system. Some of the changes have been slow and some quite rapid, making it difficult to determine whether knowledge about past climates can be scaled from millennial to decadal periods. For such systems, experience teaches that weak couplings acting over long times can have as important effects as strong couplings that may tend to balance each other. Predicting the behavior of this system, even knowing how well and on what time and space scales it can be predicted, is a major scientific challenge (e.g., Roederer, 1990).

The emissions of carbon dioxide and other trace gases are changing the atmospheric concentrations of these gases, thereby changing the climate in a grand *geophysical experiment*. Observations of neighboring planets demonstrate the important dependence of surface temperatures on atmospheric composition (e.g., Venus, which absorbs *less* solar energy than Earth because of its highly reflective cloud cover, nevertheless has a surface temperature of about 460°C because of its optically thick atmosphere of carbon dioxide and other constituents). Analyses of terrestrial geological records indicate that past climates were very different when atmospheric composition was different (US/USSR, 1990). Over the past 100 years, there is evidence that a slow greenhouse warming may have started. However, the natural variability of climate makes detection of trends and attribution of the causes difficult and extrapolation of trends highly uncertain (IPCC, 1990). As suggested by the President's Science Advisory Committee (PSAC) in 1965, numerical computer models of the climate system must be the primary tool for estimating the consequences of the changing atmospheric composition.

In the 25 years since the PSAC report, significant progress has been made in developing and testing climate models. Climate models, often referred to as general circulation models (GCMs), represent the world in three rather than one dimension and include representations of solar and terrestrial radiation; realistic geography and orography; cloud cover and convection; transport of momentum, energy, and water vapor; soil moisture; upper ocean heat capacity; and many

other processes affecting the global climate (e.g., Gates, 1985; Washington and Parkinson, 1986). In these representations, the models divide the global environment both vertically and horizontally (see Figures 1 and 2), forming grid cells that are roughly a few kilometers thick and a few hundred kilometers in horizontal extent.

For the atmosphere, these models were first applied to predict the weather, using observed oceanic temperatures as boundary conditions. As observations of the initial state of the atmosphere have improved, and as improved representations of atmospheric dynamics and thermodynamics have been introduced, the duration of useful skill in predicting the evolving state of the atmosphere has been extended from about two to five days. The important gains in the skill of weather forecasts, such as that demonstrated by the European Centre for Medium Range Weather Forecasts (ECMWF) model with its fine resolution and careful treatment of sub-grid scale physics, must now be extended to climate simulations.

ter storms, the equatorial high-precipitation and subtropical low-precipitation belts, and many other characteristics of global climate patterns and distributions (e.g., Schlesinger and Mitchell, 1985, 1987). On continental and regional scales, however, the differences between models and observations—and even the differences among ostensibly similar models—are large, ranging up to 10°C and several tens of percent for precipitation (e.g., Grotch, 1988; Grotch and MacCracken, 1990). These shortcomings limit the confidence in and usefulness of such models.

Models of the global ocean are being developed, but the 500- to 1000-year time constant of the deep ocean makes climate simulations difficult. Simulations with coupled atmosphere-ocean models have begun, but, for reasons that must still be explained, most such simulations tend to drift to a climate state different from the observed. The available models provide useful global scale insights into the sensitivity of the climate to perturbations (e.g., doubling of the carbon dioxide

Models of the climate system must be the primary tool for estimating the consequences of the changing atmospheric composition. Predicting the behavior of the coupled ocean-atmosphere system, even knowing how well and on what time and space scales it can be predicted, is a major scientific challenge.

The addition to the atmospheric models of a representation of the oceans, at least of the upper mixed layer (on average, about 100 m deep), has allowed weather-forecast-like models to be used to simulate the global climate over periods of years to decades. The results of such simulations provide statistics that can be compared to the observed climatic conditions of the much more complex real climate system. Currently available GCMs simulate many of the large-scale (global-scale) aspects of the present climate when they are driven by the observed forcing factors (e.g., solar radiation, atmospheric composition, orography) that govern energy exchange with domains external to the climate system. Thus, regardless of the observed initial state of the climate system, atmospheric models can generate the seasonal and latitudinal variations of surface temperature, the major low pressure centers that generate win-

concentration) and have been used to simulate the warmer world that existed about 9000 to 6000 years ago (COHMAP, 1988), as well as other climatically different periods of the past.

However, we are still far from having the useful regional forecasts called for in 1965 (PSAC, 1965) and needed now to estimate potential environmental impacts on hydrology, agriculture, and other societal resources. Achieving useful results on regional scales will require a significant jump from where we are now. Results are needed for impact, assessment, and adaptation studies that provide resolution on the scale of watersheds (i.e., tens of kilometers), that can be analyzed for changes in variability and in the frequencies of extreme events, and that reliably estimate the rate of change of temperature, precipitation, and other important climatic variables.

Figure 1. Schematic illustration of the processes and typical vertical gridding in the current generation of global climate models. Fully coupled versions of the atmosphere-ocean-land components are just beginning to become available.

Figure 2. Global distribution of GCM grid points (the boxes are outlined in dots with the grid center shaded in) over land areas using 4° latitude by 5° longitude resolution, which is typical of that used by climate modeling groups. The same grid-point density extends over the oceans.

The models must also be more thoroughly tested against field observations, particularly processes such as cloud-radiation interactions (the focus of the DOE ARM program) that appear to dominate estimation of climatic change (Cess et al., 1989). The ability to run such models for longer periods of time will also help greatly, as will increased resolution, particularly of topographic features. Until models can faithfully represent the present climate, recent climatic change, and past climatic changes, confidence in such models will remain limited.

In spite of model shortcomings, the urgency of the greenhouse issue has necessitated their application to study of this prospective climatic change. GCMs predict that, once climatic equilibrium is reestablished (believed to require a century or more), a doubling of the atmospheric carbon dioxide concentration (or an equivalent radiative change resulting from increases in the concentrations of carbon dioxide and other green-

house gases) would eventually lead to an increase of global surface temperatures of 1.5 to 4.5°C and to substantial changes in the amount and seasonal distribution of precipitation (and thus soil moisture).³ Preliminary studies of the impacts of a few-degree warming over the next hundred years suggest that the environmental consequences of this warming may match or exceed the extensive benefits⁴ now derived from the processes that lead to release of the greenhouse gases.

agreement amongst the various model results nor agreement between the model predictions and the observed trends over the past 100 years to better than a factor of 2 or 3 (e.g., Schlesinger and Mitchell, 1985, 1987; Hoffert and Flannery, 1985; Grotch, 1988; Grotch and MacCracken, 1990; IPCC, 1990; US/USSR, 1990). Furthermore, this disagreement exists in models that are only limited approximations to the much more complex climate system within which the changes are actually occurring.

Although much has been learned, much remains to be done. It is of fundamental importance to determine the extent to which the detailed characteristics of future regional climates (e.g., variability, frequency of extremes) can be estimated even if models were demonstrated to represent the present climate very accurately. As indicated earlier, the chaotic behavior of the atmosphere imposes a fundamental time limit on weather predictions; it is impossible even in prin-

Achieving useful results on regional scales will require a significant jump from where we are now. The models must also be more thoroughly tested against field observations. Until models can faithfully represent the present climate, recent climatic change, and past climatic changes, confidence in such models will remain limited.

house gases) would eventually lead to an increase of global surface temperatures of 1.5 to 4.5°C and to substantial changes in the amount and seasonal distribution of precipitation (and thus soil moisture).³ Preliminary studies of the impacts of a few-degree warming over the next hundred years suggest that the environmental consequences of this warming may match or exceed the extensive benefits⁴ now derived from the processes that lead to release of the greenhouse gases.

Evaluating this balance however, requires detailed information not only of changes in mean conditions on the regional scale, but also of changes in variability and in the frequency of extreme events. Analysis and comparison of the results of several climate models clearly demonstrate that, even on a global scale, there is neither

ciple to forecast for more than a few weeks into the future (Lorenz, 1963, 1984). Although we would expect continental-scale climate statistics to be more predictable than details of the local weather, the extent to which the spatial and temporal details of climate change can be predicted is not known. For example, it may prove feasible to estimate the probability that elevated greenhouse gases would lead to drier conditions in the continental interiors of the Northern Hemisphere over a 100-year period, but not possible to make a corresponding statement about conditions in Iowa over a 20-year period. Consequently, any attempt to make regional climate predictions must first address the question of regional climate *predictability*. We must determine if there are theoretical limits that will prevent models from achieving the desired degree of detail in their regional climate predictions. Only with the model throughput and diagnostic tools being sought by the CHAMMP Program can the predictability question be thoroughly investigated.

³For reference, at the last glacial maximum, the global average temperature is estimated to have been about 4 to 5°C lower; during the warm periods tens to hundreds of millions of years ago, the global surface temperature is estimated to have been 5 to 10°C higher.

⁴For example, energy generation, food production, transportation, heating and cooling, insulation, etc.

Table 1 provides an assessment of the present state of the science and contrasts it with what is needed to achieve results that are required for regional climate impact studies that are better verified, that are more closely coupled to improved understanding, and in which more confidence can be placed. Currently, climate models represent domains the size of Colorado as a single point (see Figure 3)⁵ and generally inadequately represent the deep oceans, surface hydrology and biology, and other important processes and domains. The models also suffer from oversimplified treatments of such important

aspects as soil moisture and the optical depth of clouds. Because of limited computer resources and limited data on recent and past climatic behavior, the models have not yet been used either to check their ability to simulate the 200 years since the start of industrial activities or to explore the consequences of varying many of the uncertain input parameters. Moreover, the models have been verified mainly by examining their large-scale behavior and have been run for such short periods that their inherent low-frequency (natural) variability is only beginning to be established.

Table 1. Current versus needed capabilities of climate models.

Current state	Needed capability	Required gain in computational capability
1. Horizontal resolution of hundreds of kilometers that is unable to treat regional and watershed domains	1. Horizontal resolution of tens of kilometers that permits treatment of watershed-size domains	10^2 to 10^3
2. Poorly tested and over-simplified treatments of important processes such as radiation, clouds, convection, and hydrology	2. Physically realistic and thoroughly tested treatments of all important processes, including clouds, and surface processes, snow cover, sea ice, etc.	2 to 10
3. Restricted representation of the ocean, often including only the upper mixed layer of the oceans; empirically constrained coupling of the atmosphere-ocean domains	3. Fully coupled simulation of the atmosphere-ocean system, including treatment of the deep ocean and horizontal transport by ocean currents	2 to 5
4. Separate, uncoupled, and limited treatment of atmospheric composition, chemical interaction, and aerosol effects	4. Interactive treatment of the roles and interactions of important species such as ozone and aerosols	2 to 5
5. Very limited treatment of biosphere-climate interactions affecting albedo, hydrology, and other surface processes	5. Comprehensive treatment of the interactive role of the terrestrial and marine biospheres on the climate	about 2
6. Single model simulations of tens of years with relatively coarse resolution	6. Multiple-model simulations of hundreds of years and the methodology to explore the resulting output for climate variability, time-dependent climate evolution, predictability, and sensitivity to uncertain input parameters	10 to 10^2

⁵Weather forecast models have domains of order 100 km, thus still encompassing the entire, geographically and meteorologically diverse San Francisco Bay Area as a single grid point.

Figure 3. Distribution of land grid points over the western United States (the crosses indicate the centers of grid zones) for a climate model using a typical resolution of 4° latitude by 5° longitude. A cross section of the central California grid box (shaded area) is shown to illustrate the limited regional resolution provided by the current generation of global climate models.

Moving from present capabilities, which were designed within the limits of available computational resources and capabilities, to the capabilities needed to fully address the greenhouse issue will require substantial increases in the computational throughput of models. Only with a large gain will we be able to address the challenging scientific questions with reasonable-duration simulations while representing atmospheric, oceanic, and biospheric physics, and chemistry to the full extent necessary. Remedying the most

important deficiencies will require an increase in computational throughput of at least 10^4 . Table 1 suggests that solutions of a comprehensive climate system model will ultimately require a calculational gain as high as 10^7 , and this must be the goal of continuing efforts. However, even the factor of 10^4 that is set as the goal of the CHAMMP Program during the 1990s would lead to significant improvement in model results and understanding. Weather forecasting experience, for example, has illustrated that atmospheric dynamics

and precipitation are much better represented when the spatial resolution of models is reduced from 500 km to about 100 km.

A factor as large as 10^4 cannot be achieved simply by increasing the speed of present generation supercomputers or by increasing the number of computers allocated to climate studies, although both would be very helpful. In addition, the climate models themselves, including the numerical methods used to represent processes, must be substantially modified to improve overall computational performance. Although a few research models optimally use multiprocessor machines (e.g., the ocean model of Semtner and

Chervin, 1988), the merging of multi-component models having the best physical and numerical representations of processes with fully parallel computation has not yet been achieved, even though substantial gains in computational speed have been demonstrated. Furthermore, models must be substantially restructured before they can take full advantage of emerging massively parallel capabilities.

The CHAMMP Climate Modeling Program is being designed to move systematically toward remedying these deficiencies and towards improving model capabilities and scientific understanding.

Computer Hardware,
Advanced Mathematics
and Model Physics

Goal of the CHAMMP Climate Modeling Program

3. Goal of the CHAMMP Climate Modeling Program

The current generation of climate models was developed in the 1960s and 1970s, with gradual additions and adjustments in the 1980s. In many of these models, representations of important processes such as soil moisture and convection changed little over this period. Horizontal resolution has been essentially fixed; only recently has resolution started to be refined. The emphasis has been on gradually adding representations of processes originally not treated interactively (e.g., clouds and oceans).

The models now include interactive representations for most processes believed to be important. Many of the representations, however, are well behind the state of scientific research in each

dynamics. Successes in simulating ocean eddies and currents indicate readiness to incorporate ocean models if sufficiently extended simulations can be conducted. Successes with localized ecosystem models treating leaf area index, root mass, and other integrated variables indicate readiness to include terrestrial ecosystems in global models. Improved statistical methods will enable more accurate model-to-model and model-to-field data comparisons.

However, many scientific questions and uncertainties remain to be tackled. The CHAMMP Climate Model Program will draw from the many initiatives under the U.S. Global Change Research Program, from the advances

The current generation of climate models was developed in the 1960s and 1970s, with gradual additions and adjustments in the 1980s. Many of the representations, however, are well behind the state of scientific research in each area. The present path of steady, gradual improvement with careful analysis of the incremental changes in model performance must be augmented by a major step ahead to a new generation of advanced climate models that can take advantage of both the improved process algorithms and the enhanced computing capabilities becoming available.

area. The present path of steady, gradual improvement with careful analysis of the incremental changes in model performance must be augmented by a major step ahead to a new generation of advanced climate models that can take advantage of both the improved process algorithms and the enhanced computing capabilities becoming available.

Climate models must include our best understanding if they are to become a useful tool for policy development and assessment. Recent modeling advances resulting from current programs sponsored by many agencies clearly provide the basis for improving present models. Successes in mesoscale and convective system modeling and in numerical weather prediction indicate the capability to proceed to finer resolution. Successes with chemical models of stratospheric ozone and the Antarctic "ozone hole" indicate readiness to treat the coupling of atmospheric chemistry, radiative heating, and atmospheric

being made in applied mathematics, and from the strengths of the Federal High Performance Computing Program to build an advanced climate model that can provide better estimates of regional climatic change, including changes in the frequencies of extreme and unusual events.

The goal of the CHAMMP Climate Modeling Program is to develop, verify, and apply a new generation of climate models within a coordinated framework that incorporates the best available scientific and numerical approaches to represent physical, biogeochemical, and ecological processes, that fully utilizes the hardware and software capabilities of new computer architectures, that probes the limits of climate predictability, and finally that can be used to address the challenging problem of understanding the greenhouse climate issue through the ability of the models to simulate time-dependent climatic changes

over extended times and with regional resolution.

Current GCMs have been severely limited by the capabilities of existing hardware, necessitating serious simplifying assumptions. The CHAMMP Climate Modeling Program seeks to overcome much of the hardware limitation by developing an advanced generation of climate models that make use of high-performance computers and improved algorithms. The long-term focus of the CHAMMP Program will be to develop, apply, and make available a comprehensive advanced climate model that can make the best possible application of emerging resources and understanding of how to represent behavior of the climate system. Because the best numerical and computer approach is not now evident and because no individual model will be best suited to all aspects of the issue, the actual result will most likely be a coordinated set of advanced climate models. During the 1980s, the rapid gain in per-

formance in coding and use of more efficient algorithms and numerical methods. By its completion in 10 years, CHAMMP aims to increase computational throughput by a factor of 10^4 (i.e., from a value of about 1 Gflops on the 8-processor Cray Y-MP to about 10 Tflops).

Having faster models is not enough. The judgments involved in the design, development, and implementation of advanced climate models must be "problem-driven." For energy policy development, the key scientific issues relate to regional climatic change, the rate of climatic change, changes in the frequency of rare and extreme events, and the limits of the accuracy that can be reached in predicting future climate. The CHAMMP climate modeling system will include a set of component models within a framework for interfacing the various components of the climate system. The framework will permit assembly of a specialized model ranging in comprehensiveness from the current climate models to the more en-

By including processes and scales covering a wider range than in present models, investigation can be initiated into how the variability of climate may change as the climate changes. To ensure that these models are not just creating more and more numbers, the limits of predictability and the inherent variability of the climate system must be explored.

formance of microprocessors relative to supercomputers, yet at significantly reduced cost, has laid the foundation for massively parallel supercomputers (see Figure 4). Current models of both SIMD (Single Instruction, Multiple Data) processors, such as the TMC CM-2 and MassPar MP1, and MIMD (Multiple Instruction, Multiple Data) processors, such as the NCUBE 6400, Intel iPSC/860, and BBN TC2000, are currently capable of providing gigaflops performance; and the DARPA TeraOp Initiative is targeted for the 1992-1995 timeframe. Appendix B highlights projected developments, and Appendix C describes the extensive record established by DOE for using state-of-the-art machines to solve challenging physics problems. As a consequence, massively parallel supercomputers costing about the same amount as present-day supercomputers, but with as much as a factor of 100 improvement in speed, are within reach. Over the next decade, a further factor of 10 improvement in microprocessor speed is anticipated. A further factor of 10 increase in model speed will be sought by im-

compassing models needed to answer the detailed questions concerning environmental impacts that arise in addressing the regional aspects of the greenhouse issue. The complexity of the physical processes and subsystem component models, that will be developed will be consistent with the requirements of time scales (tens to hundreds of years) and spatial scales (less than 100 km) for the intended applications (e.g., regional climatic response to time-varying greenhouse gas concentrations). A fully interactive representation of the global ocean and detailed representations of the biosphere and cryosphere must be included to allow study of the rate of climatic change over periods of decades to centuries. By including processes and scales covering a wider range than in present models, investigation can be initiated into how the variability of climate may change as the climate changes. To ensure that these models are not just creating more and more numbers, the limits of predictability and the inherent variability of the climate system must be explored.

Figure 4. Comparison of the *peak* 64-bit floating point performance (Mflops) of supercomputers and microprocessors versus time. In comparing these numbers, it is also important to evaluate the relative success of codes in reaching toward peak performance (see Dongarra, 1989).

In pursuing these objectives, the advanced climate models being developed under this program will seek to extend scientific capabilities in several specific ways:

- By providing increased horizontal and vertical resolution in a flexible and adaptable manner that permits accurate representation of processes and simulation of climate on regional scales as fine as 50 to 100 km.
- By incorporating improved representations of important climate processes to the degree needed to ensure that the models represent and retain accuracy and validity as resolution is scaled and varied.
- By providing for both partial and full coupling of the atmospheric, oceanic, cryospheric, land surface, and biospheric domains to permit investigation of a full range of climate system interactions.

- By accelerating model speed to allow longer simulations that treat time-dependent evolution of atmospheric composition and climate over decades to centuries.
- By generating and analyzing ensembles of simulations to determine the level of inherent variability and the limits of climatic predictability.
- By validating the performance of the new models against observational data sets drawn from recent, historical, and paleoclimatic periods.

The process representations to be included in the advanced climate models will evolve as our understanding of the climate system increases and as our ability to harness emerging high-performance computer technology develops. A balance will be maintained between the quest for physical completeness and the need for computational efficiency. The advanced generation of climate models must also be designed to be widely usable by multiple research groups. To the maximum

extent feasible, the new generation of advanced climate models must be:

- Modular, so that representations of particular processes can be easily interchanged.
- Selectable, so that process representations of different degrees of comprehensiveness can be coupled together to form specialized models (e.g., biogeochemical cycles, variability) that focus on the processes of most interest. (The advanced climate models should include sets of possible representations of various processes or solution methods from which options can be selected.)
- Portable, so that users can readily bring the advanced climate models onto the expanding set of computer systems of various architectures while the model retains the capability to be run on "traditional" supercomputers.
- Analyzable (including the availability of diagnostics and analysis capabilities), so that model results can be easily organized, tested, and compared.
- Evolutionary (building from current modeling capabilities), so that what we have already learned about climate dynamics forms the basis for continuing advances.

Balancing the use of faster computer hardware, more detailed formulations of climate processes, and greater flexibility and usability in an optimal manner will require that the advanced climate models be viewed comprehensively in a way that includes the many associated and accompanying set-up programs and performance and output analysis routines. Modeling and computer hardware capabilities are both poised for the major development effort that the CHAMMP Climate Modeling Program will pursue. Preliminary efforts have demonstrated significant success with multitasking of oceanic and atmo-

Climate has been recognized as a "grand challenge" issue; it is time to treat it as such and to make the commitment to the next-generation climate model.

- Scalable, so that such models can readily and efficiently take advantage of the increasing number of processors to be available in advanced computer systems.
- Accessible, so that users can gain remote access to the models through widely available computer networks.
- Documented on-line, so that users can easily understand the numerics and physics of the representations of climate system processes.

spheric codes and with other types of codes on massively parallel machines. Software and mathematics techniques are advancing rapidly, and national and international research programs are being enhanced to focus on improving understanding of critical processes. Climate has been recognized as a "grand challenge" issue; it is time to treat it as such and to make the commitment to the next-generation climate model.

Computer Hardware,
Advanced Mathematics
and Model Physics

Program Plan

4. Program Plan

The CHAMMP Climate Modeling Program will emphasize the construction of climate models that take full advantage of advanced computers and computing techniques and include the most comprehensive algorithms for representing physical, chemical, and biological processes. Constructing a new generation of advanced climate models will require a coordinated effort; building and testing of such models must be done in close collaboration with the broader scientific community.

To accomplish the CHAMMP goal, a three-phase program over ten years will be needed.

- The first phase will assemble and strengthen the scientific and computational resource base needed to construct and test advanced climate models. In so doing, this phase will advance climate modeling by transferring the existing generation of climate system models to the new generation of high-performance computers.
- The second phase of CHAMMP will develop and couple new representations of the atmosphere, ocean, biosphere, and other components of the Earth system. These new component models will include the most comprehensive and suitable processes and mathematical algorithms and achieve throughput gains on highly parallel computers of a factor of at least 100 over present models. Having such component models will substantially expand the range of scientific questions that can be posed, thereby significantly contributing to climate modeling.
- The third phase of CHAMMP will emphasize combining model components on the most advanced computers and conducting studies that test and confirm model performance on problems requiring substantially-enhanced computer capabilities (e.g., regional climatic change). This phase of CHAMMP will utilize the new models in simulations sufficiently long to explore questions of predictability, changes in the frequency of rare and extreme events, and coupling of the full climate system.

In each of these phases, CHAMMP projects will focus on three components that are equally essential to achieving success:

- Focus 1: Model Physics will assemble and test the optimal representations of physical, chemical, and biological processes suitable for use in coupled models with varying spatial resolution;
- Focus 2: Advanced Mathematics and Computational Methods will select and test the best mathematical and computational techniques for solving the complex set of coupled, interdependent equations on new generations of highly parallel computers; and
- Focus 3: Computer Hardware will take full advantage of changing computer architectures to achieve maximum computational throughput.

Focus 1: Model Physics

The Model Physics component of the CHAMMP Climate Modeling Program focuses on optimally representing climatic processes to permit accurate calculation of the many interactions that determine regional climate and climatic change. A three-element research program is needed to make regional climate projections more accurate and useful.

The first element will build on existing capabilities by incorporating and extending representations of climate processes. Of most importance will be an improvement in model resolution, which by itself will lead to a significant advance in simulation capability. Many algorithms used in available GCMs are incomplete, highly parameterized, have been tuned to present model characteristics (e.g., grid size), or rely upon unrealistic simplifying assumptions (e.g., diurnal average solar radiation). In many cases, more accurate process representations are available in limited-domain models, but have not been incorporated in global models due to the limitations of computer resources. A comprehensive review and evaluation of model algorithms is essential; identifying dependencies on resolution, determining better approaches, and ensuring independence of horizontal resolution and other model-specific assumptions can be achieved. Coordinated model-

ing activities such as DOE's Program for Climate Model Diagnosis and Intercomparison (PCMDI) and other model comparison and evaluation activities will play an important role in this effort.

Field programs such as the World Ocean Circulation Experiment (WOCE) and the First ISCCP Regional Experiment (FIRE), which are already under way, and the Atmospheric Radiation Measurement (ARM) Program, which is preparing for field monitoring and enhanced cloud-radiation modeling studies, are examples of efforts that will be essential to providing necessary data for model and algorithm validation. The CHAMMP Program will build upon these efforts by emphasizing testing and implementation of algorithms in a manner compatible with the advanced computing architectures being developed.

The second element of the Model Physics effort requires improved understanding of the factors that determine the predictability of the global and regional climate. The nonlinear interactions of atmosphere-ocean coupling can only be studied

coupled climate system cannot be met simply by an increase in throughput of 10^4 , but insightful allocation of these resources can significantly accelerate understanding. Research will be pursued that considers the relative benefits of allocating resources to finer resolution and other methods to improve accuracy, to improved model physics, to longer simulations, to coupling of climate system components, to more detailed analysis of output, and to other opportunities for enhanced understanding of climatic behavior. This research will include advanced statistical techniques capable of identifying and quantifying the tradeoffs involved.

Focus 2: Advanced Mathematics and Computational Methods

The second component of the CHAMMP Program will focus on demonstrating that climate models can take advantage of emerging high-performance computing technologies. These technologies are expected to emphasize massively parallel processing to achieve speeds of at least

Focus 1 will extend process representations, explore the limits of climate predictability, and design the advanced climate model.

with models capable of accurate simulations over long time periods. On the regional scale, the coupling of better models with model-output-statistics techniques appears to demonstrate that regional weather can be forecast using determinable and stable empirical relationships (at least for the present climate). However, there has been no comparable demonstration that regional climate statistics are predictable. Analyses and extended simulations using models with high resolution (e.g., mesoscale models) and statistical diagnostics are needed to determine the extent to which the mean climate and higher-order statistics (e.g., rare and extreme events) are predictable, especially because it is these events that are of most importance for many impact assessments.

The third element of the Model Physics component of CHAMMP is determination of the optimal choices for utilizing the increased computational power, the more accurate representations of model physics, and the improved understanding of factors determining regional climate predictability. The challenging scientific questions posed when investigating perturbations to the

several Tflops, and the feasibility of using such systems to simulate climate system behavior must be confirmed.

Focus 2 will emphasize development and utilization of algorithms, interfaces, and computing techniques. Choices will be based on minimizing the resource cost for a comparably accurate solution.

Scattered efforts to vectorize and parallelize atmospheric and oceanic models have demonstrated significant gains on computer systems with up to 16 processors. At current rates of development, "classical" supercomputers with up to 64 processors (with pipes, tailgating, many registers, etc.) will extend these capabilities in the 1990s, allowing gains in throughput by factors of 5 to 10 from hardware improvements and factors of 5 to 10 from coding improvements (see Appendix B).

This CHAMMP component will emphasize development and utilization of algorithms, interfaces, and computing techniques capable of scaling up computational performance for computers

having well beyond 64 processors. Opportunities for parallelism clearly exist; for example, some processes can be calculated on a pointwise basis (e.g., cloud microphysics, chemistry), some can be separated by vertical column and/or spectral interval (e.g., convection, radiation), and some can be treated along latitude and longitude lines (e.g., advection). These opportunities provide a very large number of independent tasks for finely resolved atmospheric and oceanic models. Of particular importance will be the development of numerical methods and algorithms capable of efficiently solving the dynamical equations that often closely couple many model gridpoints (see Appendix B). CHAMMP will evaluate the potential for implementing spectral, finite difference, finite element, and other advanced approaches, both implicit and explicit and on regular, irregular, and adaptive grids to treat advection rapidly and accurately on massively parallel systems. Load balancing, which appears straightforward for the oceanic domain where the set of calculations is relatively constant, may be much more difficult for the atmospheric domain where the

rapidly, because the optimal machine for this problem is not yet certain, and because we do not know the structure of machines that will become available in the 5- to 10-year timeframe. High-level languages must be investigated as tools for addressing these problems. Automatic model-generation techniques also must be investigated for ensuring portability and easing model development and updating. Although seemingly in conflict, scientific understanding will be advanced most rapidly if modularity and portability are effectively balanced with the overall goal of high performance and ease of model use by the researcher. Links with projects sponsored by the Federal High Performance Computing Program will be essential to addressing these problems.

Well-defined interfaces must be provided to facilitate modularity and coupling of model components and to permit design and structuring of data handling and post-processing capabilities, which could become bottlenecks. Outputs flowing from modeling systems are poised to inundate the scientist. Improved storage, analysis, and display

Focus 2 will emphasize development and utilization of algorithms, interfaces, and computing techniques. Choices will be based on minimizing the resource cost for a comparably accurate solution.

calculational burden varies greatly as a result of the varying computational demands of radiative, convective, and cloud interaction processes. The choice of algorithms, which need not be the same for all processes, should be based on minimizing the resource cost for a comparably accurate solution.

As part of this second CHAMMP component, necessary software, special programming techniques, and high-level languages must also be utilized for solving all aspects of the problem (see Appendix B). Emphasis must be placed on application of optimal software techniques in order to ensure modularity and portability. Modularity is important because it allows reuse of code modules, extends the lifespan of the code, and eases specialization of the code on different machines. Modularity is required both on the component scale (e.g., atmosphere, ocean) and on the process scale (e.g., radiation, convection) to ensure the most rapid development of advanced climate models. Portability (i.e., not over-specializing the code to a particular machine) is important because machine architectures are changing

capabilities will be essential to handle the huge volumes of data from model simulations. Coordination with NASA's EosDIS⁶ program, which is being created to handle data from the expanding set of satellite programs, is essential as a means of preparing data sets for model verification studies.

In addition to taking advantage of the new developments in computational mathematics to produce better representations of existing climate models, CHAMMP will use an array of mathematical tools to investigate the predictability of such models. The study will include such aspects as the effects of uncertainty, both physical and numerical, on the evolution of nonlinear dynamical systems; noise reduction techniques in chaotic systems for improved accuracy; and predictability when coupling multiple chaotic systems (atmosphere, oceans, etc.).

⁶Earth observing system (Eos) Data Information System

Focus 3: Computer Hardware

Current climate models utilize the full capabilities of available supercomputers. The scientific objectives of this initiative will require much greater computational power, requiring computing machines well beyond what is envisioned for next generation supercomputers (e.g., Cray-3, NEC SX/X44). Advanced climate models require supercomputer technology that is fundamentally different from the traditional machines having a small number of central processing units, each designed to run at the greatest possible speed. Appendix B provides a brief review of computer development, past and future.

Because the CHAMMP Program is one of many that are providing the incentive for computer manufacturers to develop advanced computers, CHAMMP need not be focused on designing or building computer hardware. What is important for CHAMMP is access. Thus, the third focus of the CHAMMP Program will be to

architectures compatible with proposed Tflops machines) for software development and testing. Such access will permit transfer of existing climate models to new machines as part of an evolutionary process of model improvement, as well as the development of scalable climate model algorithms for future models. Coupled data archives and interactive graphics capabilities will also be developed for the analysis of simulation results.

Program activities will be initiated to ensure the longer-term availability of computers and computational infrastructure large enough to support the planned advanced climate models. The DARPA⁷ TeraOps project being conducted as part of the Federal High Performance Computing Program is one option for the development of such a supercomputer, and is scheduled to deliver a prototype platform on a time schedule consistent with the needs of the CHAMMP Program. It may be necessary to fund alternative approaches to ensure the availability of the desired computational resources. It is also important that the hardware development be accompanied by sys-

Focus 3 will ensure access for advanced climate models to the new and emerging computer technologies.

ensure access to the steadily evolving set of state-of-the-art machines and to provide storage technology and data analysis capabilities for the large data sets that will be generated from model comparisons (see Appendix D for a description of these needs). This component will be closely coordinated with the Applied Mathematical Sciences (AMS) subprogram of the Department of Energy's Office of Energy Research (see Appendix C) and with other elements of DOE.

In the first phase of this program, access will be provided to existing supercomputers for model improvement and predictability studies and to the new parallel platforms (especially those having

tems software development so that the final result is an integrated platform that is extendable and manageable, yet flexible enough to allow experimentation with models and algorithms (see Appendix B).

An additional aspect of this component will be to develop network access so that a wide range of researchers can use the models. Because even an increase in speed of 10^4 will not be sufficient, the models must be widely available so that researchers can explore how best to allocate the increased speed in treating the many physical, chemical, and biological processes and how best to test model performance.

⁷Defense Advanced Research Projects Agency.

Computer Hardware,
Advanced Mathematics
and Model Physics

Program Management

5. Program Management

Development of advanced climate models must necessarily be a centrally-directed project that maintains a clear focus on the project goal. Figure 5 depicts the planned management structure. The DOE Atmospheric and Climate Research Division will coordinate the CHAMMP Climate Modeling Program. DOE will select the Program Leader to provide direction and focus and to promote collaborative research by the participating groups.

Because of its ambitious goal, the CHAMMP Program will require the contributions and participation of groups and scientists from many agencies, laboratories, and universities. DOE will make available funding for participation by qualified groups and individuals from universities, other agencies, and the private sector. These participants not only must contribute directly to re-

Teams. These teams, which will be relatively few in number, will focus on the development and implementation of the advanced climate models. Especially to this aspect of the U.S. Global Change Research Program, DOE and its ensemble of national laboratories bring strong capabilities in the area of computer hardware, software, and geophysical modeling, with the potential for also bringing to bear the strengths of people highly experienced in related scientific areas (e.g., fluid dynamics modeling, parallel computing, numerical methods).

The model development effort will be done in parallel with a set of CHAMMP Science Projects. This effort will involve development and testing of process algorithms, including research to determine how best to scale process representations, investigation of the limits of climate predictability,

DOE is committed to achieving the goal of much-improved climate system models in the most expeditious manner possible, working with other agencies and the scientific community to most effectively make progress. A CHAMMP Science and Development Team will be organized to provide scientific direction and coordination for the project.

search advances, but must also evince a strong and substantial commitment to the focused goal of developing advanced climate models.

DOE is committed to achieving the goal of much-improved climate system models in the most expeditious manner possible, working with other agencies and the scientific community to most effectively make progress. The CEES Task Group on Modeling and other groups will be asked to review the progress of the program on a regular basis. The JASONs, a scientific panel convened by the MITRE Corporation for DOE and other agencies, will serve as a Technical Oversight Panel to review plans and to provide scientific guidance.

Much of the model and algorithm development will be undertaken by Model Development

and studies to find improved numerical and computational algorithms. These studies will be conducted by scientists who are drawn from the laboratories and the broader atmospheric sciences community on the basis of peer-reviewed proposals. Computer access and time will be provided in a coordinated way, through both the National Energy Research Supercomputer Center (NERSC) and on massively parallel computers available within the DOE national laboratory system (see Appendix C).

A CHAMMP Science and Development Team will be organized to provide scientific direction and coordination for the project. Members will be drawn both from the Model Development Centers and from the ensemble of scientists participating in the Science Projects.

Figure 5. CHAMMP management structure

An important aspect of the CHAMMP Program will be the coordination with related programs. This will be the responsibility of all participants and will be strongly encouraged by DOE and by the CHAMMP Program Leader. DOE will

work with other agencies to coordinate support for related programs that are pursuing significant advancement of climate models in a focused and relevant way.

Linkages

6. Linkages

The Committee on Earth and Environmental Sciences report, *Our Changing Planet: The FY 1991 Research Plan* (CEES, 1990) describes a wide range of activities that will contribute to improved climate models. These activities include:

- Process studies to examine behavior of separate aspects of the climate system (e.g., clouds, oceans).
- Coupling studies to examine interactive behavior of climate system components and processes (e.g., atmosphere-ocean, cloud-radiation).
- Observations and monitoring activities to record how the system behaves.
- Model development studies to design and test algorithms and their importance in representing climate system behavior.

Rapid progress on the USGCRP, of which CHAMMP is an important part, will require the coordinated efforts of the many participants, including particularly DOE, NASA, NOAA, and NSF and their laboratories, grantees, and contractors. The interactions shown in Figure 6 indicate the linkages that must be developed.

Current modeling efforts in the core part of the DOE Carbon Dioxide Research Program (including the Program for Climate Model Diagnosis and Intercomparison), and in other agency programs supported by NOAA, NSF, and NASA are developing and applying state-of-the-science climate models on existing high-performance supercomputers. In its first phase, CHAMMP will accelerate this effort by providing a substantial resource of computer time at the National Energy Research Supercomputer Center (NERSC) for modeling groups to develop coarse-grain parallel versions of the models, conduct decadal-scale feasibility studies, carry out model intercomparison studies, and test improved process representations for inclusion in GCMs. Thus, while evaluating the optimal approaches to take in constructing the next generation of climate models, CHAMMP will assist in advancing the present generation models.

In CHAMMP's model development phase, the linkages shown in Figure 6 will become increasingly important. As programs such as ARM, FIRE, WOCE, and TOGA produce research results and better process models, CHAMMP will incorporate these improvements into the evolving structure of the advanced climate models. Similarly, as the Federal High Performance Computing Program (Executive Office of the President, 1989) and other contributing projects produce innovative computer systems and computational approaches, these advances will be incorporated into model design to optimize performance. The challenging long-term goal of better decade-to-century climate predictions at regional scales will benefit from the interactions started in the intermediate term of CHAMMP.

Collaborative efforts with climate modeling efforts currently being planned by other agencies, such as the NOAA Modeling Centers and the proposed interagency Climate System Modeling Program (CSMP), will be undertaken as appropriate. Areas of common concern include model analysis and visualization, performance improvement, algorithm design, and process parameterization. Programs such as the CSMP effort are needed to significantly enhance the rate of advancement of scientific understanding by broadening the community working to improve and test representations of physical processes and models. These efforts would be an important complement to CHAMMP, which would seek to implement these algorithms in advanced climate models and analyze resulting performance improvements.

A number of individual efforts also underway are starting to take advantage of the increasing parallelism of new supercomputers; the most successful has been the ocean GCM of Semtner and Chervin (1988), which achieves above 99% use of four- and eight-processor machines. Efforts are being initiated to move versions of these models to massively parallel computers as a means of further accelerating model speed. CHAMMP will help to focus and intensify these efforts.

Figure 6. CHAMMP conceptual design.

CHAMMP is also expected to benefit from activities conducted under the DARPA TeraOp Initiative, which has as its goal the delivery of a massively parallel computer capable of 10^{12} operations/second by the year 1995 (10^3 to 10^4 times the throughput of a single-processor Cray X-MP). The milestones of that program include delivery of a proof-of-principle machine that would be used as a building block for the full machine to be constructed by 1992. The proof-of-principle machine is expected to deliver 1% to 10% of the power of the full machine.

The CHAMMP climate models will be available for widespread use. Access to the new climate models will be provided through the networks being established under the Federal High Performance Computing Program.

These important, but diverse, linkages will be maintained through intensified communication (e.g., newsletter, meetings, advisory committees, coordinated program support) and other initiatives of the CHAMMP Program Leader. In addition, the modeling working group of the Committee on Earth and Environmental Studies will serve as an important interagency link.

Resource Requirements

7. Resource Requirements

CHAMMP is an ambitious, multi-element program that will require substantial resources, even taking advantage of the many advances in computers, software, and storage that will result from other high-technology programs. Table 2 provides an estimate of the resource requirements for CHAMMP, including broad indications of activities being undertaken in each phase of the project. An important assumption is that other agencies (e.g., NSF, NOAA, NASA) will also intensify their efforts to improve understanding of climate processes through field and model experiments and other activities.

The estimated funding requirements increase substantially over the decadal period in order to provide access to essential computer hardware, to support advancing mathematical methods and techniques, and to improvements to representations of physics in the models.

scaled to provide access to a 100-Gflops machine by years 4 and 5 and to a multi-Tflops machine by years 8 to 10. Appendix B describes prospects for such developments.

Comparable support will be required for data storage. Unlike many traditional physics problems, a very important aspect of the climate problem will be time-series analyses of extended simulations in order to examine questions of variability, frequencies of extreme and unusual events, and other factors. Animation of model output, which can provide a unique view of atmospheric behavior, and direct model output, which is needed for spatial and temporal analyses, will be required. If all relevant information is stored for analysis, which is important if such information is to be made available to the wide community of potential users, incremental increases of the order of 10^{15} bytes of data per year would be

CHAMMP is an ambitious, multi-element program that will require substantial resources, even taking advantage of the many advances in computers, software, and storage that will result from other high-technology programs.

Computer Hardware

Developing advanced climate models that run at 100 Gflops within six years and that are readily movable to the multi-Tflops machine expected by the end of the 1990s requires that a series of massively parallel machines be available. A 100-Gflops machine could be built over the next several years if the commitment were made. Although limited access to such a machine may become available through other programs, rapid development of advanced climate models will require primary access for model simulations.⁸ The proposed support requirements have been

required when full model operation begins, and would increase thereafter (see Table 3). The alternative is to output only the data needed at the time from any given simulation. Given the chaotic nature of the atmosphere, however, repeating extended runs exactly in order to investigate interesting phenomena may be difficult in that any seemingly insignificant change in a prognostic variable (whether from a change in algorithm, cosmic ray, or truncation) may rapidly lead to changes in the weather, if not in climate statistics. Funding must be included to develop the hardware for storing and providing access to these large amounts of data (see Appendix D).

⁸ It is important to recognize that, although throughput is to be increased (initially by a factor of about 100), this increase will be balanced (perhaps more than balanced) by an increase in computer requirements as a result of increased resolution and more detailed treatment of model physics; as a consequence, total run time for model experiments will continue to require full access to such a machine.

Table 2. Estimated CHAMMP resource requirements (\$M/yr).

Period	Activities highlights	Estimated annual funding ^a requirements
Startup (FY90)	<ul style="list-style-type: none"> • Program development and organization • Exploratory transfer of existing models to massively parallel computers 	\$5.8M
Initial phase (FY91 to FY93)	<ul style="list-style-type: none"> • Establishment of Model Development Teams and selection of initial set of Science Project participants • Performance studies with existing models on high-performance computers • Development of criteria and components for advanced climate models 	\$7-15M
Model development phase (FY94 to FY96)	<ul style="list-style-type: none"> • Construct advanced climate system component models incorporating improved algorithms and numerical methods and achieving computational throughput of 100 Gflops • Conduct extended simulations with coupled component models to explore climate predictability and ultimate model requirements • Develop and test advanced storage, analysis, display, and verification techniques 	\$15-25M
Model implementation phase (FY97 to FY00)	<ul style="list-style-type: none"> • Construct coupled climate system models achieving computational throughput of 10 Tflops • Conduct extended time-dependent calculations examining the limits of climate predictability • Provide climate projections of greenhouse climate effects with resolutions of finer than 100 km 	\$25-50M

^a This funding is in addition to funding for advancing research on climate system behavior as part of the DOE core climate program and to funding for modeling research by other agencies.

Table 3. Potential annual data storage requirements for simulations with advanced climate models.

Model characteristics	Storage requirements (8-bit words)
Horizontal grid points (about 40 km × 40 km grid resolution)	300,000
Vertical grid points (of order 40 in the atmosphere and 40 in the ocean)	80
Variables per grid point (including climatic and chemical variables, prognostic, and diagnostic variables)	20
Samples per day (perhaps represented as Fourier coefficients)	4
Days per year	365
Years per simulation (average)	100
Number of simulations	20
Annual storage requirements (2 bytes per variable may provide sufficient accuracy ^a)	> 10 ¹⁵

^a Because of the chaotic nature of atmospheric behavior, full-word storage will be required if model simulations are to be rerun, which is often desirable for investigating detailed aspects of model behavior. However, because such storage is required less often, it is not as demanding of storage resources.

Advanced Mathematics and Computational Methods

The primary computer demands of advanced climate models will arise from requirements related to simulation of atmospheric radiation, transport and chemistry, and to the possible need to simulate the eddies in the oceans. For the radiation and chemistry aspects, the possibility exists to decompose the global problem into many smaller problems (e.g., vertical column radiation for each of several wavelength intervals, chemistry at individual gridpoints), although it will be important to determine whether this is the optimal approach and to develop efficient solution kernels. A more thorough analysis will be needed to determine, for example, solution techniques for representing the dynamics of the atmosphere and ocean that are both accurate and computationally efficient on massively parallel computers. The implicit and spectral techniques now being used to represent transport in many atmospheric models couple the global set of grid points, whereas other approaches permit separation. Optimal vertical and horizontal gridding must be determined, and consideration must be given to questions such as whether the present vertical coordinate system in the atmosphere (using proportional mass, or sigma, as the coordinate) forces more than local coupling of grid points.

Once approaches and physics representations are selected, the software must be written and tested on the prototype machines expected to be available in the next few years, and then moved to the 100-Gflops machine to be available in about five years. Obtaining maximum efficiency will require development of a wide array of new tools for optimizing performance on these new machines.

Development of graphics and analysis packages must be simultaneous with development of advanced climate models, both to provide tools for testing new models and to ensure that the models can be effectively used once they are developed. Network connections and interfaces must be developed so that the model output is readily available for assessment and impact studies by the scientists evaluating model performance and by potential users.

Model Physics

The advanced climate models must enable scientists to assess potential impacts of the emissions of greenhouse gases. Algorithms must be specially designed and tested to perform with finer resolution. For some algorithms (e.g., clouds), model simulations with finer grids already indicate problems in scaling to finer resolu-

tion. For other algorithms (e.g., hydrology), new parameterizations must be developed and calibrated on these finer scales. The DOE/ARM program, the FIRE and ISLSCP programs, GEWEX, TOGA, and WOCE field programs, the Eos-derived data sets, and many other CEES-supported research efforts must be drawn upon to develop and test algorithms. Funding must be provided to involve CHAMMP in the community-wide effort to develop and implement improved algorithms; in this area CHAMMP must be especially well coordinated with other agency activities.

As the new generation of climate models is developed, a range of calibration tests must be performed with each model. Extensive modeling and analysis efforts will be required to establish the model control climate, to compare model results to observations, to evaluate climatic variability and predictability, etc. These efforts will also need to be coordinated with the activities at other agencies and organizations, including various model intercomparison studies and model development and testing activities within the United States and abroad. Funding must be included to significantly enhance the DOE core program ac-

tivities, as well as to bridge the CHAMMP model development activity with the multiagency program.

Project Coordination

Development of advanced climate models on a 10-year time frame will require a coordinated and directed program, as indicated in the section on management. In addition to the organizational coordination of project participants, model organizational issues must be addressed. Protocols must be established to ensure that subroutines and data can be readily interchanged, so that the model is easily transferable to new computer systems, and so that there is ready access for use by the scientific community.

To accomplish the essential coordination efforts, CHAMMP will have a Program Leader to direct its development and outreach activities. Advice on directions to take and choices to make in designing new models will be sought from the scientific community, through workshops and other communication tools.

Appendix A

Appendix A: Brief History of the Department of Energy's Carbon Dioxide Research Program

In 1977, the National Academy of Sciences challenged the scientific community on the subject of energy and climate by stating: "To reduce uncertainties and to assess the seriousness of the matter, a well-coordinated program of research that is profoundly interdisciplinary in character, and strongly international in scope, will be required."

Responding to this challenge and the growing concern about the long-range consequences of carbon dioxide (CO₂) emissions resulting from ever-increasing fossil fuel combustion, the Department of Energy undertook a thorough examination of the CO₂ problem. The first step was to convene the *Workshop on the Global Effects of Carbon Dioxide from Fossil Fuels* in 1977 at Miami Beach, Florida. Some 75 scientists discussed the state of knowledge of the CO₂ cycle and the consequences of increases in the atmospheric CO₂ concentration. The workshop identified significant gaps in understanding and recommended research to fill those knowledge gaps.

Based upon this information, the Department of Energy organized the Carbon Dioxide and Climate Research Program, with initial funding in FY 1978. The goal of this research program was the identification of possible policy options for governmental action in response to changes in the atmospheric CO₂ concentration. Achievement of this goal was seen to require increased understanding of CO₂ interactions with the atmosphere, the biosphere, the oceans, and the cryosphere and to evaluate the effects of the resulting changes on critical resources and human welfare.

In 1978, Congress enacted the National Climate Program Act to establish a comprehensive national policy for dealing with all climate-related issues. This step was largely in response to the need for a national (and international) climate research program, first called for by the U.S. GARP Committee's Panel on Climate Variation in 1975. Responsibilities under the Act involve several government agencies. In full cooperation with this program, DOE was design-

nated as the lead agency for coordinating the government's research efforts in the area of atmospheric CO₂ and the resulting climatic effects and perturbations. Within DOE, this responsibility has been carried out by the Carbon Dioxide Research Division of the Office of Basic Energy Sciences, Office of Energy Research (OER), now the Atmospheric and Climate Research Division, Office of Health and Environmental Research (OHER).

In 1979, the DOE and the American Association for the Advancement of Science cosponsored the *Workshop on the Environmental and Societal Consequences of a Possible CO₂-Induced Climate Change*. This meeting brought together 85 scholars to discuss and review 28 contributed papers. The workshop was a difficult but timely first step in identifying questions to be addressed by research on the impacts of climatic change. Prior to this workshop, there had been little discussion of secondary effects or of the potential consequences of global climatic change. One important tangible outcome was the realization that CO₂ directly affects plants and that the Earth's vegetation could change in response to increased CO₂, regardless of whether there was a climatic response. This finding led to expanded research (in cooperation with the U.S. Department of Agriculture) on the direct effects of the increasing CO₂ concentration on plants, a program that has since provided the primary data base for determining the combined effects of CO₂ and climatic change on crops and ecological systems.

Under the auspices of DOE, in-depth review of the results of global greenhouse research was conducted, culminating in 1985. This review, which involved many national and international researchers, produced four state-of-the-art reports and two companion volumes. These volumes described what is scientifically known with certainty about the greenhouse effect, what is known with less certainty, and what remains largely unknown. Findings of the state-of-the-art reports pointed to four major research needs. First, general circulation models (GCMs) must be compared to observed data and to each other to

understand where and why the results of the different models agree and disagree. Second, uncertainties in GCM simulations arising because of their treatments of clouds, oceans, and other processes must be reduced, and horizontal resolution must be refined in order to provide more reliable estimates of regional climate change. Third, whole-system data on crops and ecosystems must be gathered, and models must be developed that can provide the basis for evaluating the impacts of climatic and CO₂ change on the environment. Finally, the effects of the increasing CO₂ concentration and of climatic

change on the world's resource systems, including food and fiber, and freshwater, coastal and forest systems, must be evaluated.

The Carbon Dioxide Research Program has responded to these recommendations by refocusing and restructuring its original research program and by proposing new initiatives, including the Atmospheric Radiation Measurement (ARM) and CHAMMP Programs. The goal continues to be to provide adequate scientific knowledge to the government and other interested parties for identifying and analyzing potential responses to the greenhouse effect.

Appendix B

Appendix B: Prospects for Large-Scale Computing

Hardware Characteristics and Performance

Supercomputer speeds are typically measured relative to the Cray line of machines, which serve as production platforms at many research and industrial facilities. The approximate *single*-processor performance for a climate simulation ranges from about 50 to 100 Mflops^{B1} for either a Cray X-MP or Cray-2 to about 100 Mflops for the Cray Y-MP. Both the Cray-3 under development at Cray Computing Corporation and the C-90 under development at Cray Research project a single-processor speed of about 300 Mflops. *Aggregate* performance numbers for *multiple*-processor platforms would place a 4-processor X-MP at about 200 Mflops, an 8-processor Y-MP at about 1 Gflops^{B2} and a 16-processor Cray-3 or C-90 at about 5 Gflops. (The Semtner-Chervin parallel ocean climate model actually runs at slightly over 1.1 Gflops on an 8-processor Cray Y-MP.) The recently announced NEC SX/X vector processor is rated at an aggregate *peak* speed of about 20 Gflops using 4 processors with 8 pipes each, compared with the expected *peak* speed of about 15 Gflops for the Cray-3 and C-90. Thus, the projected performance of these computers, which are scheduled to be available within the next two years, is similar.

However, to achieve CHAMMP's goal of an increase in throughput of 10^4 compared to today's computational speed of about 1 Gflops (on an 8-processor Cray Y-MP), a system must be devised to sustain 10 Tflops. The NEC and Cray classes of computer are not likely to meet this challenge. To do so, they would have to produce systems consisting of 10,000 or more processors, because their current single processors are already near fundamental physical limits on speed. The cost of manufacturing such a computer would greatly exceed the costs of more massively parallel computers based upon microprocessor technology.

Microprocessors, represented by reduced instruction set computer (RISC) architectures such as the Motorola 88000, MIPS, and Sun Sparc, have been routinely providing about 5 Mflops for the past few years. The recently announced Intel i860 and IBM RISC/6000 microprocessors with their multiple floating-point pipes and multiple instruction lengths are now providing sustained single-processor speeds approaching about 30 Mflops. (Figure 4, on page 19, compares the peak 64-bit floating-point performance history of microprocessors to that of supercomputers.) These new chips, now no more than a factor of 10 slower than a single-processor Cray Y-MP, are some 1000 times less expensive! Thus a price/performance ratio some two orders of magnitude better than the best present-day supercomputers can be achieved by linking together a large number of these chips (furthermore, the computational speed of such a massively parallel machine could be increased without fundamental limit by increasing the number of processors if the computer application could make use of such capabilities). In other words, a massively parallel supercomputer costing about the same amount as a present-day Cray or NEC machine, but potentially with a factor of 100 improvement in speed, is within reach of current technology. Recently announced computers, such as the NCUBE 6400, Intel iPSC/860, and MasPar MP-1 Series, along with the TMC CM-2 and BBN JC2000, make excellent starts toward the desired system. A further factor of 10 improvement in microprocessor speed is anticipated during the next decade. For example, the Micro 2000 microprocessor is projected to deliver 1000 Mflops by the end of the decade (Gelsinger et al., 1989).

B1 One megaflops (Mflops) equals 1 million (10^6) floating point operations per second.

B2 One gigaflops (Gflops) equals 1 billion (10^9) floating point operation per second.

The final factor of 10 required to meet CHAMMP's goal is probably achievable through a combination of improvements and changes in numerical methods and through algorithm improvements (e.g., the 100-Mflops speed of current models on a single-processor Cray Y-MP is less than half the inverse of that machine's clock cycle time).

Climate Model Requirements

Massively parallel hardware, then, is becoming available. Can climate model programs be recoded to take advantage of it? Maximum performance from a system of N processors cannot be achieved from an application or algorithm that has fewer than N units of computation to be performed simultaneously. Because currently projected Tflops platforms will consist of 5,000 to 100,000 microprocessors, we must consider the prospects for achieving such granularity in general circulation models. (*Peak* speed of 20,000 Intel i860 processors is estimated to be 1.2 Tflops.) We can estimate the number of computational operations, and therefore the limits on granularity, involved in climate model calculations by starting with an atmospheric model. Using characteristics of present models, the horizontal resolution of 4° latitude by 5° longitude and a vertical resolution of 3 km results in a three-dimensional mesh of about 30,000 grid points.^{B3} Advanced atmospheric models will require much finer horizontal and vertical resolution, simulations extending for centuries, and more complex representations of physics, chemistry, and biology. The need for a multi-Tflops computational platform is clear.

Using these figures, placing one grid point per computational node could efficiently utilize up to 30,000 processors. Improving grid resolution by only a factor of two in each dimension could use up to 240,000 processors. Even without an increase in grid resolution, it is possible to use extra

processors to pipeline timestep calculations or to compute chemical processes. When an oceanographic model is added, the number of grid points again significantly increases. Thus, it is likely that many grid points will be assigned to each processor, thereby cutting down the communication-to-computation ratio and resulting in the potential for efficient parallel codes.

With memory costs decreasing, massively parallel computers are being produced with a substantial amount of memory, which also increases the ability of each processor to handle both a large number of grid points and relatively sophisticated representations of physical, chemical, and biological processes. Because most of the grids used in circulation models are regular, logically-connected, three-dimensional meshes, they are, at least in theory, well suited for the communications characteristics of massively parallel machines. With the exception of radiation and convection calculations, the communications required for algorithms in climate models are ideally regular "nearest-neighbor" grid communications. The use of columnar-type algorithms for radiation and convection calculations would result in a one-dimensional regular structure for the communication. In practice, the use of pseudo-spectral techniques in many climate models to represent horizontal transport does create the need for communications beyond nearest neighbors, in this case to all grid points.

A careful evaluation is needed to determine the extent to which such techniques can be parallelized, factors relating the choice of technique to the number of processes, and the potential advantages and disadvantages of introducing alternative techniques (see box on Numerical Methods). Overall, however, it is clear that sufficient parallelism exists that climate models should be able to exploit the characteristics of massively parallel machines.

^{B3} At each grid point, models solve for the longitudinal and latitudinal components of the momentum conservation equation, as well as the conservation equations for mass, thermal energy, and water vapor. This results in 5 equations, each of which may include about 10 different processes (advection, convection, surface processes, etc.). Using an estimate of 40 mathematical operations per process per equation results in about 2,000 floating-point operations per grid point. Assuming a simulation time step of 30 minutes gives an approximate computational burden of about 10^{12} floating point operations (1 Tflops) per simulated year.

Numerical Methods

A number of complementary problems related to the choices of numerical methods must be addressed in development of advanced climate models. These problems must be considered together because the optimum choice of method will likely depend on the choice of computer hardware as well as on the accuracy of the solutions that can be provided.

Many current numerical models for the global climate do not adapt the mesh to important geographic or dynamical features. With the greater computational power offered by the next generation of supercomputers, it will be possible to simultaneously increase the overall resolution and the local resolution of the models. Among the issues to consider here are the tradeoffs between the added complexity of both static and dynamic adaptivity and the simplicity of a computationally regular mesh. Further, these tradeoffs must be examined in the context of the overall climate model; choosing a method whose accuracy does not match the other components of the problem (such as model physics or boundary data) may be counterproductive.

In making use of massively parallel computers, it is important to minimize the amount of global communication or data synchronization needed by the algorithm. The cost of such communication is high enough to force a re-examination of the tradeoffs between implicit and explicit methods, particularly for problems like the global climate model, where speeds in the jet stream are of the same order of magnitude as sound-wave speeds. This re-examination must be carried out both at the level of numerical approximations to a given set of equations and in the choice of the equations themselves. In particular, a velocity potential formulation, because it introduces an elliptic equation that involves global communication in its solution, may be inappropriate for an algorithm for massively parallel computers. Taken in combination with the need to adapt to the static and dynamic features of the problem, these considerations suggest re-examining the use of explicit methods for the primitive equations that make use of irregular or locally regular meshes.

Another issue that must be considered in the case of massively parallel computing is that of load balancing. If we divide the computation into a series of tasks to be performed at each time step, we can consider the sources of load imbalance. These range from easily controlled sources, such as differing numbers of mesh points, to difficult-to-control features, such as a variable number of nonlinear iterations at each mesh point. One approach is to generate many more tasks than processors and dynamically assign tasks to processors. This incurs a potentially substantial cost if the context of a task is large (and hence requires moving a lot of data). Alternatively, the tasks may be statically assigned to processors, thereby trading the cost of moving tasks around against the cost of having idle processors. This will influence the decision whether to use a dynamically adaptive grid, and may influence the methods used to implement the model physics, because a predictable or constant amount of computational effort per processor will be needed to avoid serious losses in efficiency in the many-processor case. Intermediate approaches that consider the cost of moving a task to a different processor also must be considered. Note also that the dynamic load-balancing approach is more practical with numerical methods that have the smallest context (such as explicit differencing schemes and easily parameterized model physics computations).

Parallel Programming Tools

Advanced climate models will be large and complex, and will be used and modified by hundreds of people, on many different computers, over a period of ten years or more. Both to minimize development costs and to maximize utilization, the models must be designed to be modular, re-configurable, portable, and scalable. The CHAMMP Program will require more advanced tools, designed to support scientific programming on large parallel computers. These tools will include some or all of the following: high-level languages, performance tools, and parallel program development environments.

High-level Languages. A high-level language is one which permits algorithms to be expressed at a more abstract level than low-level languages such as Fortran. These languages provide a number of advantages, including: (1) Rapid prototyping. Simplified, throw-away prototypes can be coded quickly to prove concepts, to refine specifications, and to achieve rapid user feedback. (2) Ease of maintenance. Algorithms expressed in high-level languages are often significantly shorter and easier to understand than equivalent Fortran programs. This simplifies maintenance and modification. (3) Portability. High-level parallel languages enhance portability and scalability of parallel codes by specifying parallel execution, communication, and synchronization in a machine-independent manner.

Performance Tools. The tailoring of a program to achieve maximal performance on a particular computer is an important part of the development of a scientific program. On parallel computers, this optimization step is particularly difficult because of the need to consider issues such as granularity, mapping, and communication and synchronization costs. Both climate model developers and users working with new model configurations or on new parallel computers will require specialized tools to provide them with insights into the interactions between their model and the machine on which it is executing. These tools should emphasize the use of graphical display techniques to summarize large amounts of data.

Development Environments. Developers of parallel algorithms will require tools to aid in the debugging of their programs. Hence, any high-level parallel programming system adopted for use in CHAMMP projects must provide a convenient programming environment, including integrated debugging tools.

The most important class of climate model "programmer" will ultimately be the user of the model. Typical users of existing climate models spend a great deal of time experimenting with alternative physics, modifying parameter settings, and analyzing data. The productivity of these people can be greatly enhanced by modern graphical programming environments tailored to the particular requirements of climate modeling. Tasks to be supported by the sort of environment may include: definition of new physics modules; composition of existing modules to construct specialized models; parameter setting; running of experiments on remote computers; and statistical analysis of model output.

Software Portability and Performance

Prior large-scale program development efforts have generally been able to take advantage of the ANSI Fortran standard to achieve a reasonable compromise between portability and performance in a single maintained source code. The CHAMMP initiative does not have this luxury, because at least three parallel programming environments exist that appear to provide a path to the necessary performance.

The first parallel programming environment, often referred to as SIMD (Single Instruction, Multiple Data), is the programming model for the Connection Machine CM-2 and MasPar MP-1. The program must specify which variables are *plural* (each processor has the variable but with different data) or *singular* (accessible as a single value from the front-end processor). The user therefore still writes a monolithic program, but expresses parallelism in terms of data structures. Both the CM-2 and MP-1 offer a C language with plural variable extensions. Both machines offer

Fortran 8X, with the array constructs of Fortran 8X serving as plural declarations. There appears to be little or no performance penalty for writing the CHAMMP software in such a way that the SIMD version will port from one SIMD architecture to another and will scale with advances in that computer technology.

The second parallel environment is that of DM-MIMD (Distributed Memory, Multiple Instruction, Multiple Data) computers. This category includes the NCUBE II and Intel iPSC/860 parallel computers. In this model, the processors potentially have completely independent programs and data, cooperating by explicit passing of messages between processors. (In practice, each processor typically has a copy of a single program and its own copy of often-read global data, thereby simplifying the programming effort at the cost of additional storage requirements per node.)

The commercially available DM-MIMD computers all offer standard C and Fortran languages with extensions to exchange messages. The more recent offerings remove the penalty for routing to processors that are not connected directly (and hence the penalty for programming to a particular topology such as hypercube, tree, or two-dimensional mesh). It is a simple matter to develop preprocessors that convert generic message-passing constructs to ones specific to each vendor; this has been done successfully, for example, by Oak Ridge National Laboratories (Geist et al., 1990). Hence, there appears to be little performance penalty for developing a version of the CHAMMP software for a generic DM-MIMD computer. Again, the benefits of the resulting portability within the architectural class will be substantial, allowing ready transfer of codes to improving models of this type of machine.

The third programming model is that of SM-MIMD (Shared Memory, Multiple Instruction, Multiple Data). This is the model used by "slightly parallel" computers such as Iris, Cray, Sequent, and Stardent. The parallelism is usually invoked via compiler directives specific to each vendor. This approach does not appear to promise the performance increases required by CHAMMP. For this reason, CHAMMP will concentrate only on the SIMD and DM-MIMD approaches.

Can the SIMD and DM-MIMD approaches be consolidated into a single, maintainable pro-

gram? Currently, the answer appears to be that they cannot without substantial performance penalties. An experiment (by Steve Otto at Caltech) to put a SIMD environment on an NCUBE computer resulted in a tenfold performance drop. Putting a DM-MIMD program on a SIMD machine would almost certainly involve disabling many of the SIMD processors to deal with conditional branches that DM-MIMD environments handle with full processor activity. The SIMD environment is like soldiers following close-order drill, whereas the DM-MIMD environment is like a franchised business. For optimum efficiency, each requires a plan of action specific to its style of organization.

CHAMMP will, therefore, support programs that evaluate the relative strengths and weaknesses of each approach for the various challenges posed by climate system models. In developing new codes, CHAMMP will maintain a focus on the two approaches, constructing codes for the SIMD and DM-MIMD environments as appropriate. Within each environment, constructs specific to any one vendor or ensemble size will be avoided so as to allow easy porting to higher-speed computers as they become available. This will also require that CHAMMP participate in development of the many parallel programming tools necessary to achieving these objectives (see box on Parallel Programming Tools).

Examples of Early Applications

There are many initial efforts seeking to take advantage of massively parallel computing. In atmospheric sciences, Hoffmann and Maretis (1990) provide an overview of present activities. The studies they describe clearly indicate the potential gains that are needed and could be achieved.

In areas outside atmospheric sciences, there have already been successful applications. For example, the Department of Energy has extensively funded the development of parallel processing technology through its Applied Mathematical Sciences program. Over the last few years, several of the DOE laboratories involved in this program have begun to apply this technology to large, production-oriented problems. With the availability of a 1024-node NCUBE-2 at Sandia National Laboratories (SNL), a 128-node iPSC/860 at Oak Ridge National Laboratory (ORNL) and two 64K processor Connection Machines (CM-2) at Los Alamos National Laboratory

(LANL), many of these production applications have equaled or surpassed present Cray X-MP and Y-MP performance capability. Although few of these applications have been climate related, the following examples involve computational techniques that are very similar to those used in climatological calculations.^{B4}

Numerical Linear Algebra. Many techniques for the solution of sparse and full matrix formulations of linear systems of equations have been parallelized. Researchers at ORNL have attained computational speeds of 1 Gflops on the iPSC/860 for the Cholesky factorization of a dense matrix, while similar results have been achieved by SNL on the NCUBE-2 using preconditioned conjugate gradient techniques for sparse matrix problems resulting from heat conduction/mass diffusion problems. The sparse matrix techniques have also been developed for the CM-2 and have demonstrated up to 8 times the performance (32-bit calculations, SNL) (Cline et al., 1989) and 2 to 3 times the performance (64-bit calculations, LANL) of a Cray X-MP single processor.

Finite Element Method. An SNL structural mechanics code (JAC3D) has recently been ported to the NCUBE-2 and has demonstrated 5 times the speed of a single-processor Cray X-MP for moderately sized problems. The domain decomposition and load balancing techniques used in this implementation are applicable to a wide range of finite element and finite difference calculations.

Hydrodynamics. The LANL implementation of 3-D Free-Lagrange hydrodynamics code (X3D) on the CM-2 has resulted in performance equal to a full 8-processor Cray Y-MP, while the ORNL iPSC/860 version of the RSF 3-D reduced magnetohydrodynamics code achieved the same performance of a Cray-2 processor by using only 16 nodes. The Free-Lagrange method used in X3D and the pseudo-spectral method used in RSF are both applicable to the solution of the atmospheric circulation equations.

Preliminary Climate Model Results. A CM-2 implementation of the finite difference formulation of the shallow water equations by NCAR/TMC^{B5}/LANL researchers has demonstrated performance of 6 Gflops, over 6 times the performance of 8-processor Y-MP. ANL/ORNL/NCAR collaborators have implemented a spectral circulation model similar to that used in the NCAR CCM on the iPSC/860 and preliminary results indicate parallel speed-ups of 36 using 64 processors and 48 using 128 processors.

As shown by these applications, as well as others, the currently available parallel machines have been proven capable of delivering 10 times the performance of Cray single processors. Full scale machines of the current NCUBE and the next-generation Intel and TMC products will most likely provide a performance increase 100 times that of a Cray processor. Given these successes, we can expect great near-term improvements in the computational throughput of climate calculations through CHAMMP-funded research.

^{B4} Performance results were obtained from personal conversations with Harold Trease of LANL, Robert Ward and John Drake of ORNL, and Raymond E. Cline, Jr. and Sudep Dosanjh of SNL.

^{B5} TMC = Thinking Machine Corporation

Appendix C

Appendix C: Supercomputing in the Department of Energy

From the early days of scientific computing, the Department of Energy (DOE) and its predecessor agencies, the Atomic Energy Commission (AEC) and the Energy Research and Development Administration (ERDA), have played a major role in the development and application of computers. The driving force behind the AEC's desire to develop the capability to solve large computational problems was its mission to develop nuclear weapons. Designing and testing nuclear weapons present a challenging and complex set of problems that do not lend themselves to simple analytical approaches or detailed experimental measurements. Computational solutions became a necessity; consequently, much of the pioneering work on supercomputer development and application was conducted at two of the AEC's laboratories that had prime responsibility in the development of nuclear weapons—namely, the Los Alamos National Laboratory (LANL) and the Lawrence Livermore National Laboratory (LLNL).

The first digital general-purpose electronic computer, the ENIAC, was built under contract with the U.S. Army's Ballistic Research Laboratory. ENIAC has been followed by a continued stream of more powerful, state-of-the-science computers that have emerged from a productive working relationship between the DOE laboratories and computer companies. It became standard practice for either Los Alamos or Livermore to take delivery of serial number 1 or 2 of each new supercomputer. As a result, much operating system software was developed by the DOE laboratories, such as the Cray time-sharing system (CTSS). In addition, many challenging physics problems were first attempted on these machines, including the first global general circulation model constructed at LLNL by Leith in the early 1960s (Leith, 1965).

One of the first major uses of large-scale computing in a nondefense program came in the DOE's Magnetic Fusion Energy (MFE) Program. The highly nonlinear theoretical problems required sophisticated numerical approaches. Another driving force was the need to make the supercomputer capability available to a widely dispersed community of fusion researchers. The

approach taken was to establish a supercomputer center at the LLNL and to provide access to other sites through a wide-band communications network based on 56-kilobaud land lines connecting the center with major research centers.

In 1983, the user community was expanded beyond the MFE Program to include all of the research programs funded by the Department of Energy's Office of Energy Research. These additional programs include High-Energy Physics, Nuclear Physics, Basic Energy Sciences, and Health and Environmental Research. The National Magnetic Fusion Energy Computer Center (NMFECC) and its communication network successfully demonstrated that supercomputer resources could be provided to a widely dispersed set of users through a nationwide satellite and ground communications network. This operation became a model that has been emulated by other federal agencies and international organizations.

As of the summer of 1990, the complement of supercomputers at the National Energy Research Supercomputer Center (NERSC, formerly NMFECC) at LLNL include:

- an 8-processor Cray-2 with 128 Mwords of shared memory,
- a 4-processor Cray-2 with 128 Mwords of shared memory,
- a 4-processor Cray-2 with 64 Mwords of shared memory, and
- a single-processor Cray X-MP with 2 Mwords of main memory.

These machines have access to a large, common file system that provides more than 3.6 Tbytes (3.6×10^{12} bytes) of storage.

At this time, the NERSC supercomputers are accessible through the Energy Research Networks (ERNets), the backbone (ESNet) of which has a bandwidth of 1.5 Mbits/s. Using the TCP/IP and DECNet protocols, the ERNets provide communications for Energy Research scientists and their collaborators worldwide. In the United States, the ESNet is connected to the national network, providing the Department's scientists and

engineers with rapid communications to the scientific community. For example, the NSFNet and its regional affiliates or regional networks are components of Internet, making virtually all academic institutions accessible.

Computational Studies at DOE

The Department of Energy pioneered the development of scientific computing to complement experiments and analytic theory. In some areas of research, computing has become the major tool for exploration, with experiments used to check the computed results and to provide data in regimes where computer models are inadequate. There are significant areas of research where experimental measurements are extremely expensive or impossible with today's technology; computer simulation is the only practical method of developing an understanding of the nature of the processes involved. This revolution in scientific methods has taken place over the past 20 years and is still continuing as computers become more powerful and algorithms more efficient.

Most of the computational models of interest to DOE researchers are simulations of processes involving physical, chemical, or biological interactions. The following are the general categories of problems with some examples.

Energy System Design: Magnetic confinement fusion; nuclear reactors; combustion; solar collection, conversion, and distribution; atmospheric and oceanic circulation; pollutant dispersion.

Weapons System Design: Nuclear systems; warheads and packaging; inertial confinement fusion.

Basic Research: Particle accelerator design; quantum chromodynamics; structural biology; electronic structure.

Current Supercomputing Facilities at DOE

In fiscal year 1989, DOE had 32 supercomputers installed with an original purchase value of \$538 M, or 69% of the total DOE investment in automated data processing equipment (3,391 computers with an original purchase value of \$775 M). This represents approximately 14% of the total automatic data processing capability of the U.S. government.

Energy research programs account for approximately 22% of the total DOE use of computers. The mission is directed toward long-term scientific investigations in high-energy and nuclear physics, materials and chemical sciences, engineering and geosciences, biological and environmental research, and magnetic fusion energy. Supercomputer centers supported by the Office of Energy Research (OER) include NERSC and the Supercomputer Computations Research Institute at Florida State University (supported in part by OER).

Research and Development on Future Computational Resources

The Applied Mathematical Sciences (AMS) subprogram in the Office of Energy Research is the program element in the Department of Energy that is chartered to support research and development of future computational resources for the Department's fundamental research mission. The purpose of the AMS subprogram is twofold: (1) to ensure the broad range of research in the mathematical and computer sciences necessary to underpin research in all of the other sciences, and (2) to provide an international network and state-of-the-science supercomputing facilities for all of the researchers supported by the Office of Energy Research. The subprogram's research mission is subdivided into major sections on analytical and numerical methods, information analysis techniques, and advanced computing concepts.

The analytical and numerical methods activity covers the wide spectrum of activities important to mathematical modeling of physical systems. Examples include nonlinear dynamics (strange attractors, chaos) and partial differential equations (fronts, nonlinear waves). Other examples in computational mathematics include numerical algebra (systems of equations, eigenvalue problems), differential equations (partial and ordinary equations, vortex methods), and approximation and optimization (interpolation, global methods).

Topics in information analysis techniques include statistical methods, data management, and geometry and visualization.

In the advanced computing concepts category, the idea of parallel processing has firmly taken hold, and most of the techniques used in previous sequential computations must be reinvestigated for new architectural schemes de-

signed to increase the capacity and capability of computer systems for future needs of science and engineering. The approach includes both software methodology and high-performance systems. The groups supported by the activity in high-performance systems include:

- Argonne National Laboratory's Advanced Computing Research Facility (ANL/ACRF).
- California Institute of Technology's Concurrent Computation Project (CCCP).
- Los Alamos National Laboratory's Advanced Computing Laboratory (LANL/ACL).
- Oak Ridge National Laboratory's Mathematical and Sciences Section and its Advanced Computing Laboratory (ORNL/ACL).
- New York University's Ultracomputer Project (NYU/Ultra).
- Sandia National Laboratories' Parallel Processing Group (SNL/PPG).
- University of Illinois' Center for Supercomputing Research and Development (UI/CSRD).

Additional studies at these and other DOE facilities supported by other elements of DOE and by other agencies further enrich the set of activities within the DOE community.

The high-performance computing activity has helped provide an impressive array of new architectures and spawned many new activities in software and algorithm research not possible without this access. Examples include

Hypercubes:

- Intel iPSCs of various sizes at ANL/ACRF, ORNL/ACL, and LANL/ACL.
- 128-node Intel Touchstone (i860s) at ORNL/ACL.

- 1024-node NCUBE at SNL/PPG.
- 512-node NCUBE at CCCP.
- 65,000-node Thinking Machines (CM-2) at LANL and Florida State University.
- 16,000-node Thinking Machines (CM-2) at ANL/ACRF and SNL/PPG.

Shared-Memory Machines:

- 128-node BBN TC2000 (88000 RISC) at LLNL.
- 46-node BBN TC2000 at ANL.
- 32-node CEDAR-1A (Alliant processors) at UI/CSRD.

Special Graphics Machines:

- Stellars at ANL/ACRF and SNL/PPG.
- Stardent/Titan at ANL/ACRF.
- Stardent at ORNL/ACL.
- Stardent with AVS at LANL/ACL and LANL/IGPP.

Summary

The Department of Energy has been an important leader in the development and application of supercomputer technology and expects to continue its efforts with vigor in the future. The future holds many new challenges as the Department and the nation strive to maintain international leadership in high-performance computing and to advance computational science and engineering to address important problems in simulation, data retrieval, and direct computation.

Appendix D

Appendix D: Data Storage and Visualization

Both data storage and its rapid and interactive display pose challenges that must be met in order to actively utilize the increased throughput of advanced climate models. As shown in Table 3 (see page 39), an advanced climate model would generate on the order of one terabyte (10^{12} bytes, or 8×10^{12} bits) of information for each simulated year; thus, a few dozen 100-year simulations would require storage of at least 10^{15} bytes. The best current mass storage facilities at supercomputer centers fall short of this requirement by one or more orders of magnitude. Expansion of storage facilities by the required amount, however, is achievable by a straightforward extension of present-day technology.

Supercomputer centers at the National Energy Research Supercomputer Center (NERSC) and the National Center for Atmospheric Research (NCAR) provide examples of the state-of-the-art of mass storage at operational facilities. Both centers rely on magnetic tape cartridges that store about 200 Mbytes each. At NERSC, the cartridges are stored in a half dozen cylindrical robotic retrieval facilities, or "silos." Each "silo" occupies about ten square meters of floor space and stores some 6,000 cartridges, i.e., over 10^{12} bytes. During the next five years, more advanced storage media, such as VCR-type magnetic tapes or optical storage, will probably become reliable enough to replace the tape cartridges, providing at least an order of magnitude increase in capacity. It is possible that by the end of the 1990s "silos" storing nearly 10^{15} bytes will be commonplace; if not, the required storage increment could be achieved in a brute force manner by simply purchasing a large number of "silos."

After the information generated by a model is stored, it must be retrieved and processed in order to be understood. For the sizable four-dimensional data sets produced by an advanced climate model, as well as the observational data sets to which the model's simulations must be compared, the ability to manipulate and graphically display the data rapidly and interactively—in modern parlance, to visualize it—may well become the rate-determining step in reaching a sensible interpretation of the model's behavior. Present-day technology is rather limited in this regard. Individual workstations are the natural choice for interactive graphics, and at the present time the

high-end, general-purpose models can produce either high-resolution two-dimensional images or relatively crude three-dimensional graphs almost as fast as they are requested. Production of image-quality three-dimensional graphics, including precise color and shading with light sources, is considerably slower on these machines. Furthermore, including the dimension of time by animating the graphics is sufficiently computer-intensive to require either specialized workstations or supercomputers. Nevertheless, the hardware is available at a price, and it will no doubt progress rapidly in response to commercial incentives; applications useful to climate modeling will be paced by software development, which is now under way at a number of institutions.

The database needs of the CHAMMP Initiative also require large bandwidth input/output from mass storage subsystems, i.e., fast transfer of data from the mass storage to workstations. Massively parallel processing machines may themselves be well suited for such database applications. The design of the BBN TC2000 parallel computer provides an example of a large bandwidth information processor. Because each processing element of the TC2000 can be equipped with a VME bus interface, it is possible to provide a standard SCSI disk controller at each processing node. Each SCSI controller can interface to seven disks, each with a capacity of up to 1 Gbyte of data. The current 63-node TC2000 could handle over 400 Gbytes, and future 500-node versions of the machine would permit an online capacity of 3.5 Tbytes.

For maximum utilization, it is necessary to provide for the dissemination of model results and validation data sets from the archival database to participating researchers at distributed locations. It is anticipated that the requests for information from these locations will not be required to have real-time forwarding of information. Under these circumstances, the data rate needs can be met through the use of the new regional and national networks.

Perhaps the greater concern in this area of communications is to provide a standard, flexible, and efficient interface for the analysis and graphical representation of simulation data. The database that is created will not be used to anywhere near its potential capacity unless capabili-

ties for browsing, selecting, and querying data are supported by the ability to create indices of significant pictorial features extracted from image data. A system to generate image data with visualization technology must be married to database query technology to create pattern and abstract data-indexing capabilities. With databases that are too large to scan sequentially, these indexes would provide researchers with the tools needed to locate the most valuable features to be studied. Data management systems must thus be developed to provide an integrated approach to data storage and analysis while providing a flexible and extendable user interface.

To summarize, climate modeling under the CHAMMP Program will generate an unprecedented amount of data requiring storage, access,

and analysis. The requisite technology for storage is either available now or expected to evolve without substantial investment from CHAMMP, although, of course, CHAMMP must allocate sufficient funds for acquisition of such storage equipment in support of its climate modeling activities. One pacesetter storage development is expected to be the NASA EosDIS project that is to handle satellite data sets. Because of the similar technical challenges of handling such large data sets, and because these data sets will be essential to model verification, CHAMMP will coordinate its model results storage and display program with the EosDIS project. Adequate analytical capabilities, however, must be developed with at least partial support from CHAMMP. This is particularly true with regard to statistical comparisons of output from different models and between model output and observational data.

Computer Hardware,
Advanced Mathematics
and Model Physics

References

References

- CEES (Committee on Earth and Environmental Sciences) (1990), *Our Changing Planet: The FY 1991 Research Plan of the U.S. Global Change Research Program*, Office of Science and Technology Policy, Washington, D.C.
- Cess, R.D., G.L. Potter, J.P. Blanchet, G.J. Boer, S.J. Ghan, J.T. Kiehl, H. Le Treut, Z.-X. Li, X.-Z. Liang, J.F.B. Mitchell, J.-J. Morcrette, D.A. Randall, M.R. Riches, E. Roeckner, U. Schlese, A. Slingo, K.E. Taylor, W.M. Washington, R.T. Wetherald, and I. Yagai (1989), "Interpretation of Cloud-Climate Feedback as Produced by 14 Atmospheric General Circulation Models," *Science*, **245**, 513-516.
- Cline, R.E., Jr., J.C. Meza, B.M. Boghosian, and B. Walker (1989), "Towards the Development of Engineering Production Codes for the Connection Machine," *Proceedings of the Fourth Conference on Hypercubes, Concurrent Computers, and Applications*, March 6-8, Monterey, CA.
- COHMAP members (1988), "Climatic Changes of the Last 18,000 Years: Observations and Model Simulations," *Science*, **241**, 1043-1052.
- Department of Energy (DOE), Report of the DOE Multi-Laboratory Climate Change Committee (1990), *Energy and Climate Change*, Lewis Publishers, Chelsea, Michigan, 161 pp.
- Dongarra, J.J. (1989), "Performance of Various Computers Using Standard Linear Equations Software," Report CS-89-85, University of Tennessee.
- Executive Office of the President (1989), *The Federal High Performance Computing Program*, Office of Science and Technology Policy, Washington, D.C.
- Gates, W.L. (1985), "Modeling as a Means of Studying the Climate System," 57-79 in *Projecting the Climatic Effects of Increasing Carbon Dioxide*, M.C. MacCracken and F.M. Luther (Eds.), report DOE/ER-0237, U.S. Department of Energy, Washington, D.C.
- Geist, G.A., M.T. Heath, B.W. Peyton, and P.H. Worley (1990), "PICL: A Portable Instrumented Communication Library," report ORNL/TM-11130, Oak Ridge National Laboratory, Oak Ridge, TN.
- Gelsinger, P.P., P.A. Gargini, G.H. Parker, and A.Y.C. Yu (1989), "Microprocessors CIRCA 2000," *IEEE Spectrum*, October, 43-67.
- Grotch, S.L. (1988), "Regional Intercomparisons of General Circulation Model Predictions and Historical Climate Data," report DOE/NBB-0084, U.S. Department of Energy, Washington, D.C.
- Grotch, S.L., and M.C. MacCracken (1990), "The Use of General Circulation Models to Predict Regional Climatic Change," *Journal of Climate* (in press).
- Hoffert, M.I., and B.P. Flannery (1985), "Model Projections of the Time-Dependent Response to Increasing Carbon Dioxide," 149-190 in *Projecting the Climatic Effects of Increasing Carbon Dioxide*, M.C. MacCracken and F.M. Luther (Eds.), report DOE/ER-0237, U.S. Department of Energy, Washington, D.C.
- Hoffmann, G.-R., and D.K. Marettis, (Eds.), (1990), *The Dawn of Massively Parallel Processing in Meteorology, Proceedings of the 3rd Workshop on Use of Parallel Processors in Meteorology*, Springer Verlag, Berlin, 376 pp.
- IPCC (Intergovernmental Panel on Climate Change) (1990), *Scientific Assessment of Climate Change (Report of IPCC Working Group I)*, Cambridge University Press, Cambridge, UK.
- Leith, C.E. (1965), "Numerical Simulation of the Earth's Atmosphere," 1-28 in *Methods of Computational Physics*, Vol. 4, B. Alder (Ed.), Academic Press, New York.
- Lorenz, E.N. (1963), "Deterministic Nonperiodic Flow," *Journal of the Atmospheric Sciences*, **20**, 130-141.

- Lorenz, E.N.** (1984), "Irregularity: A Fundamental Property of the Atmosphere," *Tellus*, **36A**, 98-110.
- PSAC** (President's Science Advisory Council) (1965), "Appendix Y4: Atmospheric Carbon Dioxide," 111-133 in *Restoring the Quality of Our Environment*, Report of the Environmental Pollution Panel, The White House, Washington, D.C.
- Roederer, J.G.** (1990), "The Challenge of Global Change," *EOS*, **71**, 1085.
- Schlesinger, M.E.**, and J.F.B. Mitchell (1985), "Model Projections of the Equilibrium Climatic Response to Increased Carbon Dioxide," 81-147 in *Projecting the Climatic Effects of Increasing Carbon Dioxide*, M.C. MacCracken and F.M. Luther (Eds.), report DOE/ER-0237, U.S. Department of Energy, Washington, D.C.
- Schlesinger, M.E.**, and J.F.B. Mitchell (1987), "Climate Model Simulations of the Equilibrium Climatic Response to Increased Carbon Dioxide," *Reviews of Geophysics*, **25**, 760-798.
- Semtner, A.J.**, and R.M. Chervin (1988), "A Simulation of the Global Ocean Circulation With Resolved Eddies," *Journal of Geophysical Research*, **93**, 15,502-15,522.
- US/USSR** (1990), *Prospects for Future Climate*, M. C. MacCracken, M. I. Budyko, A.D. Hecht, and Y. Izrael (Eds.), Lewis Publishers, Chelsea, MI, 270 pp.
- Washington, W.M.**, and C.L. Parkinson (1986), *An Introduction to Three-Dimensional Climate Modeling*, University Science Books, Mill Valley, CA.

Computer Hardware,
Advanced Mathematics
and Model Physics

Acronyms and Abbreviations

Acronyms and Abbreviations

AEC	Atomic Energy Commission
AMS	Applied Mathematical Sciences (DOE)
ANL/ACRF	Argonne National Laboratory's Advanced Computing Research Facility
ARM	Atmospheric Radiation Measurement (program)
CCCP	California Institute of Technology Concurrent Computation Project
CCM	Community Climate Model
CEES	Committee on Earth and Environmental Sciences
CHAMMP	Computer Hardware, Advanced Mathematics, and Model Physics (Climate Modeling Program)
COHMAP	Cooperative Holocene Mapping Project
CSMP	Climate System Modeling Program
CTSS	Cray time-sharing system
DARPA	U.S. Defense Advanced Research Projects Agency
DECNet	Digital Equipment Corporation Network
DM-MIMD	Distributed Memory, Multiple Instruction, Multiple Data (computer)
DOE	Department of Energy
ECMWF	European Centre for Medium Range Weather Forecasts
Eos	Earth observing system
EosDIS	Earth observing system Data Information System
ERDA	Energy Research and Development Administration (predecessor to DOE)
ERNets	Energy Research networks
FIRE	First ISCCP Regional Experiment
GARP	Global Atmospheric Research Program
GCM	General circulation model
GEWEX	Global Energy and Water Cycle Experiment
Gflops	One gigaflops equals 1 billion (10^9) floating point operations per second
IPCC	Intergovernmental Panel on Climate Change
ISCCP	International Satellite Cloud Climatology Program
ISLSCP	International Satellite Land Surface Climatology Program
JASON	A scientific panel convened by the MITRE Corporation
LANL	Los Alamos National Laboratory
LANL/ACL	Los Alamos National Laboratory's Advanced Computing Laboratory
LLNL	Lawrence Livermore National Laboratory

MFE	Magnetic Fusion Energy Program
Mflops	One megaflops equals 1 million (10^6) floating point operations per second
MIMD	Multiple Instruction, Multiple Data (processor)
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
TMC	Thinking Machine Corporation
NERSC	National Energy Research Supercomputer Center
NMFECC	National Magnetic Fusion Energy Computer Center
NOAA	National Oceanic and Atmospheric Administration
NSFNet	National Science Foundation Network
NYU/Ultra	New York University's Ultracomputer (Project)
OER	Office of Energy Research (DOE)
OHER	Office of Health and Environmental Research (DOE)
ORNL	Oak Ridge National Laboratory
ORNL/ACL	Oak Ridge National Laboratory's Advanced Computing Laboratory
OSTP	Office of Science and Technology Policy
PCMDI	Program for Climate Model Diagnosis and Intercomparison
PSAC	President's Science Advisory Committee
RISC	reduced instruction set computer
RSF	a LANL computer code
SIMD	Single Instruction, Multiple Data (processor)
SM-MIMD	Shared Memory, Multiple Instruction, Multiple Data (computer)
SNL	Sandia National Laboratories
SNL/PPG	Sandia National Laboratories, Parallel Processing Group
Tflops	One teraflops equals 1 trillion (10^{12}) floating point operations per second
TMC	Thinking Machine Corporation
TOGA	Tropical-Ocean Global Atmosphere (research program)
UI/CSRD	University of Illinois Center for Supercomputing Research and Development
USGCRP	U.S. Global Change Research Program
WOCE	World Ocean Circulation Experiment

- END -

DATE FILMED

12 / 28 / 90