

Distributed Ocean–Atmosphere Modeling and Sensitivity to the Coupling Flux Precision: The CATHODE Project

C. CASSOU,* P. NOYRET,⁺ E. SEVAULT,[#] O. THUAL,[@] L. Terray,* D. BEAUCOURT,⁺ AND M. IMBARD[&]

**Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique, Toulouse, France*

⁺*Electricité de France/Département Etude et Recherche, Clamart, France*

[#]*Service Central d'Exploitations Météorologiques, Météo-France, Toulouse, France*

[@]*Institut National Polytechnique de Toulouse, Toulouse, France*

[&]*Laboratoire d'Océanographie Dynamique et de Climatologie, Université Paris, Paris, France*

(Manuscript received 29 May 1996, in final form 10 July 1997)

ABSTRACT

The authors present the distribution of a coupled ocean–atmosphere global circulation model. The atmospheric (ARPEGE) and the oceanic (OPA) components run separately at different sites; the coupling is achieved through the exchanges of fluxes via the coupler (OASIS) and the three independent programs communicate together through the 2-Mbit RENATER network. The coupling and distributing procedure is based on the PVM software and is validated by 1-yr simulations. Performances and difficulties raised by the distributed environment are also presented.

An additional study deals with the sensitivity to the precision in exchanged data in coupled mode. This question is addressed by introducing small artificial perturbations into the forcings of each component. The transient growth of these perturbations is first followed over 10 days on specific grid points. A global spatio-temporal analysis is then performed over the planet for 1-yr simulations.

During the first 10 days of the experiments, the “error” dynamics is amplified by the atmosphere with a doubling time of the order of 5 days, while the upper ocean simply relaxes toward equilibrium. For long time ranges of simulation, errors tend to saturate and oscillate around a plateau, following the seasonal cycle. Spatio-temporal studies prove that the most sensitive areas to the precision in exchanged forcings are related to the regions where the variability is the most pronounced. These analyses are integrated into the general studies of predictability in coupled ocean–atmosphere models.

1. Introduction

The climate system is composed of the ocean, the atmosphere, the cryosphere, and the biosphere. Nowadays, it is widely recognized that the highly nonlinear interactions between these subsystems control the climate of the whole planet. Modeling the feedback between the two main components, namely the ocean and the atmosphere, is therefore a key issue for realistic climate simulations. This feedback can be assessed with coupled general circulation models (CGCMs), which are the most powerful tools to represent the main features of the climate, its variability, and its global change. They include numerous physical processes and also provide a consistent basis to analyze the physics and the thermodynamics in each fluid. This is a crucial point with respect to the experiments and the understanding of the impact of environmental changes such as the anthropogenic increase of greenhouse gases.

Climate simulations with CGCMs require huge computer resources. With the present technology of vectorial supercomputers, several months are necessary to integrate long simulations. For instance, a 25-yr coupled simulation (2.8° resolution in the atmosphere, about 2° resolution in the oceans) requires roughly 600 CRAY C90 hours of computing resources and about three months of elapsed time. One solution to decrease the total duration of climate experiments is to parallelize and distribute the models.

With improved analyses of global climate processes and computationally efficient GCMs, Mechoso et al. (1993) tested the distribution of a climate model based on domains, tasks, and I/O decomposition. They demonstrated the enhanced performance of such a decomposition, which leads to a significant decrease of the experiment duration. These results open the way to increased resolution and more complex physical parameterizations, which are necessary to simulate small-scale phenomena of greatest importance (oceanic eddies). The present work follows this pioneer experiment in a different environment, with different models, and with different coupling techniques.

A first goal of the so-called CATHODE project

Corresponding author address: O. Thual, CERFACS, 42 Av. Coriolis, F-31057 Cedex, Toulouse, France.
E-mail: thual@cerfacs.fr

(Couplage, Atmosphère Océan distribué) is to realize a global coupled simulation, distributed on two remote supercomputers, to analyze its performance and to study the difficulties raised by distributed computing between different research centers, with their own operational constraints. The coupling is achieved through the exchange of fluxes via a high-speed network. This study is a two-year cooperative project between the CER-FACS (Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique), EDF/DER (Electricité De France/Direction des Etudes et Recherches), Météo-France, and the LODYC (Laboratoire d'Océanographie DYnamique et de Climatologie).

Another goal of the project is to investigate the sensitivity of the coupled model to the precision in exchanged data (i.e., the precision in the surface boundary conditions of each GCM). This analysis belongs to the category of studies on sensitivity to the finite precision of fields due to discretization, round-off, and truncation errors in a numerical experiment. The distributed computing addresses that question by the potential differences in the type of remote machines and in the precision of exchanged fluxes. The purpose of this analysis is to evaluate if this systematic imprecision that is introduced into the model boundary conditions can induce changes on the modeled climate and its variability.

A detailed description of the different components of the coupled model is provided in section 2. A special interest is dedicated to the coupled simulations performed on a single CRAY machine validating the global climate model (Guilyardi et al. 1995; Pontaud et al. 1995). Here the emphasis is on the model robustness and stability to simulate the general features of the climatic system. Section 3 is devoted to the computing aspects. A general presentation of the CATHODE project is made whereas a more detailed description of software, methodologies, technical challenges, and performance for such a distributed application are discussed in the appendix. Sections 4–6 are devoted to the sensitivity to the precision in exchanged data. The different types of artificial perturbations introduced into the model to assess this issue are presented in section 4. Section 5 deals with the transient growth of a perturbation over ten days. In section 6, the sensitivity of the model is investigated over one year in terms of spatial patterns in order to identify the most sensitive climatic areas. The main conclusions are drawn in section 7.

2. Description of the ocean–atmosphere coupling

The OPA–ARPEGE model is used in this study. This coupled model is being developed as part of the French climate community effort to study climate variability and potential anthropogenic effects. Flux corrections have been used by several modeling groups to reduce climatic drift in long-term simulations. But the effect of such artificial flux correction on the model's variability is unclear (Meehl 1995). In this work, the coupled

model is kept free from any adjustment at the sea–air interface and within the ocean.

a. Numerical models

1) THE ARPEGE–CLIMAT ATMOSPHERIC GLOBAL CIRCULATION MODEL (AGCM)

The ARPEGE–Climat (Version 1) AGCM, from CNRM/Météo-France, is a state-of-the-art spectral atmosphere model developed from the ARPEGE/IFS weather forecast model (Déqué et al. 1994). The model has 30 vertical levels extending up to 70 km (0.02 hPa). A hybrid sigma–pressure vertical coordinate is used with 10 levels in the troposphere, 15 levels in the stratosphere, and 5 levels in the mesosphere. The standard horizontal resolutions used in this model are T21, T42, and T79.

The radiation scheme is both an extension and a simplification of the methods described in Geleyn and Hollingsworth (1979) with optical properties synthesized in a few coefficients following methods similar to those of Ritter and Geleyn (1992). The scheme is fast enough so that it can be called at each time step of the model. Both deep and shallow convection are parameterized. The deep convection uses a mass-flux scheme with detrainment described by Bougeault (1985) while the shallow convection is parameterized with a modified Richardson number scheme described by Geleyn (1987). The exchange and drag coefficients for heat and momentum are computed according to Louis et al. (1982). Convective and stratiform cloudiness are calculated using the precipitation rates and the vertical humidity profile. A two-layer prognostic soil temperature scheme is included in the model. A fourth-order horizontal diffusion is applied to all the prognostic variables and the semi-implicit time integration scheme allows for a 15-min time step for the T42 resolution.

This model is part of the AMIP (Atmospheric Model Intercomparison Project) intercomparison (Gates 1992) and has been used for sensitivity studies including the removal of Arctic sea ice (Royer et al. 1990) and time slice experiments (Mahfouf et al. 1994). CATHODE's experiments are performed with the T42 truncation, corresponding to a 2.8° grid size.

2) THE OPA OCEANIC GENERAL CIRCULATION MODEL (OGCM)

The OPA OGCM has been developed at the Laboratoire d'Océanographie DYnamique et de Climatologie (Delecluse et al. 1993). It has been used from process studies (Madec et al. 1991) to basin- (Dandin 1993) and global-scale studies (Madec and Imbard 1995). It solves the primitive equations with a nonlinear equation of state (UNESCO 1983). Those equations are discretized on a staggered Arakawa C-grid (1972) using finite-dif-

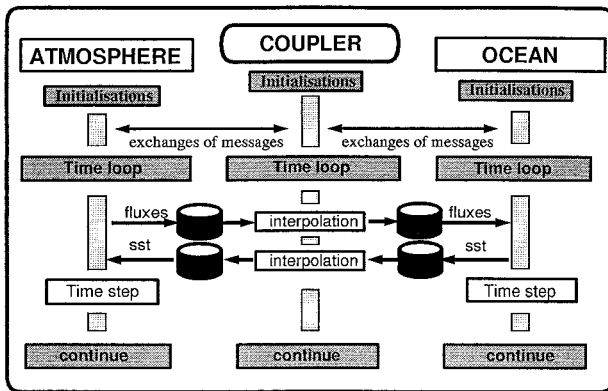


FIG. 1. Algorithm of the OASIS coupler. The coupler sends a message both to the atmospheric and the oceanic models, which start their temporal integration during a time step previously defined. Once this time limit is reached, the ocean and the atmosphere send a message to the coupler and keep the data ready to exchange. The coupler performs the grid-to-grid interpolation; then the ocean and atmosphere read the transformed data and start a new integration until the next exchange.

ference schemes. The use of vectorial operators relying on tensorial formalism ensures second-order accuracy on any curvilinear orthogonal grid. The discretization of the vorticity term in the momentum equation ensures a potential enstrophy conservation for any horizontal nondivergent flow. Time stepping for the advection, Coriolis, and pressure terms is achieved by a basic leap-frog scheme associated with an Asselin filter applied at each time step in order to avoid time splitting. The forward scheme is used for horizontal diffusive processes and an implicit one for vertical diffusion. A rigid-lid approximation is made at the surface.

The horizontal resolution of the model is roughly equivalent to a geographic mesh of $2^\circ \times 1.5^\circ$ (with a meridional resolution of 0.5° near the equator). There are 31 vertical levels with 10 levels in the upper 100 m. Vertical eddy diffusivity and viscosity are computed from a 1.5-order turbulent kinetic energy (TKE) closure scheme (Blanke and Delecluse 1993). Zero fluxes of heat and salt and no-slip conditions are applied at solid boundaries. The model does not include a sea-ice component but a simple pseudoparameterization that only involves a test on the sea surface temperature. With this resolution, the usual time step is 1 h and 40 min but it has been reduced to 1 h in the coupled experiments.

3) THE OASIS COUPLER

The ARPEGE and OPA models are coupled through the OASIS coupler developed at CERFACS (Terray 1994; Terray et al. 1995). As depicted in Fig. 1, the coupler ensures two independent tasks organized in two distinct modules: the *interpolation module* (hereafter referred as OASIS-Interp) dedicated to the spatial inter-

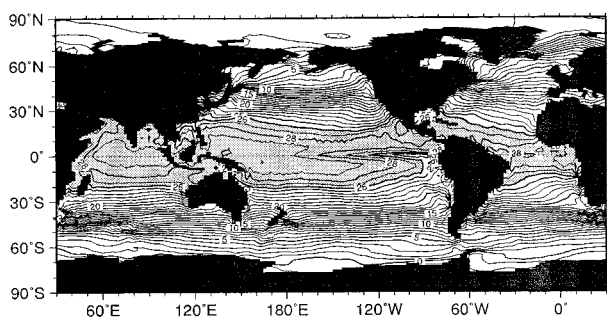
polation from one grid to another, the *time synchronization module* (hereafter referred as OASIS-Pipe) dedicated to file exchanges and time management. ARPEGE, OPA, and OASIS (OASIS-Interp/OASIS-Pipe) run in parallel and independently. They exchange coupling fields once per day, averaging out the diurnal cycle (synchronous coupling). Sea surface temperature (SST) and sea ice extent (SIE) are given to the AGCM and surface fluxes of heat, momentum, and freshwater to the OGCM. The surface fluxes are interpolated onto the OGCM grid using bicubic interpolation and a nearest-neighbor algorithm is used to interpolate the SST and SIE onto the AGCM grid.

The initial state of each GCM is built from uncoupled experiments. The AGCM ARPEGE is forced with the COLA-CAC (Center for Ocean-Land-Atmosphere-Climate Analysis Center) observed SST for the period 1979–93. The OGCM OPA is forced by the daily averaged energy momentum and water fluxes generated by the previous forced atmospheric experiment. In these forced runs, the oceanic model is run in *robust diagnostic* mode: model temperature and salinity are relaxed toward Levitus climatology below the thermocline, outside the tropics, and away from the coast. The tropical Pacific SST and the global poleward energy transport obtained from this integration are in good agreement with the observations. We use the conditions corresponding to 1 January 1986 to initialize the atmospheric and the oceanic components of the coupled model.

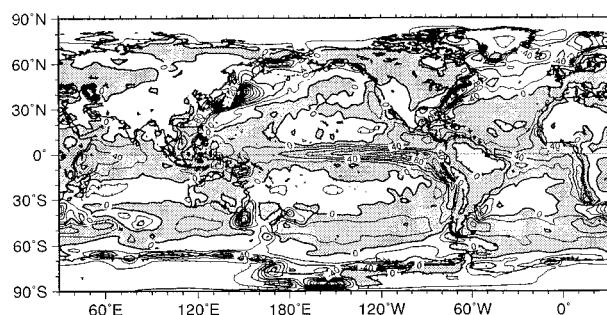
b. Undistributed coupled ocean-atmosphere simulations

Before its distribution between several remote supercomputers, the coupled model was intensively studied to identify its main qualities and deficiencies. Three simulations have been performed and used to validate the model in nondistributed mode. As a first step, the OPA model in its Pacific version has been coupled with the T42 ARPEGE AGCM for 10 years (Terray et al. 1995; Mechoso et al. 1995). Then, two global coupled experiments have been performed, with different horizontal resolution in the ARPEGE AGCM: T21 and T42 truncations are used for 50-yr and 25-yr simulations, respectively. The high-resolution experiment is the control simulation to which distributed experiments can be compared. Indeed, this experiment, referred as CG3, is the identical twin of those realized in distributed mode (same resolution and similar version of the GCMs).

Experiment CG3 exhibits a strong initial (1 year) tropical warming; however, it leads to a stable tropical mean state at the air-sea interface after 10 years. The drift of the global mean oceanic temperature is between $+0.3^\circ$ and $+0.4^\circ\text{C}$ per century. Figure 2a shows the SST averaged during year 25 of the simulation. Some usual biases of GCMs are present: a too warm SST over the tropics induces too weak trade winds in the atmo-



Contour interval = 1 from -4.06961 to 30.6665



Contour interval = 20 from -356.328 to 154.849

FIG. 2. (a) Sea surface temperature (SST) of CG3 simulation averaged during year 25. Values greater than 26°C are shaded and contour interval is 1°C. (b) Total heat flux transmitted from the atmosphere to the ocean of run CG3 averaged during year 25 of the simulation. The heat flux is positive when the atmosphere heats the ocean. Positive values are shaded and contour interval is 20 W m⁻².

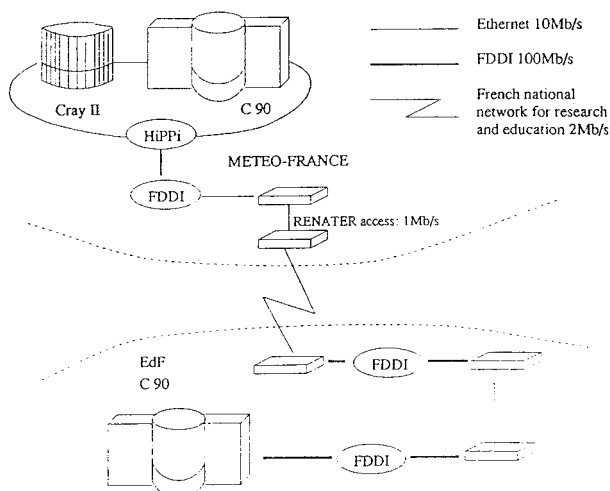


FIG. 3. Network architecture between EDF and Météo-France.

sphere. The warm pool is reasonably simulated in the western Pacific, but too spread over the basin. This eastward extension gives rise to a warm climatological mean state over the Tropics. Moreover, a double intertropical convergence zone (ITCZ) appears in the southeast equatorial Pacific during the boreal winter. At high latitudes, the ice extension is too large in the Northern Hemisphere, whereas the Southern Hemisphere shows a rapid melting. The SST interannual variability is present in the tropics (not shown) but its structure differs from the observations (too weak and too confined on the east side of the basins). Nevertheless, with no artificial flux correction, the ARPEGE-OPA coupled model realistically simulates the main features of the climate characteristics (see Fig. 2b representing the total heat flux transmitted from the atmosphere to the ocean).

3. CATHODE: Technical realizations

a. The metacomputer architecture

One objective of the CATHODE project is to pool computing resources to perform detailed numerical simulations of very complex systems such as the climate. With the emergence of the parallel version of atmospheric (Dent et al. 1994) and oceanic (Guyon et al. 1994) models as well as recent advances in distributed computing techniques, the distribution of the AGCM ARPEGE and the OGCM OPA between two supercomputer sites is now investigated. The idea is to evaluate the advantage of long-coupled simulations using “metacomputers.”

CATHODE is organized around two supercomputer centers: the center of Météo-France in Toulouse and the center of Electricité De France (EDF) in Clamart (near Paris). These centers are about 700 kilometers apart and are both connected to 2-Mbit RENATER, the French high-speed wide-area network devoted to scientific applications. Three supercomputers are used in this project: a CRAY C98 at EDF, a CRAY C98 at Météo-France, and a CRAY II at Météo-France (Fig. 3).

The communication protocol is adapted to the new distributed context. The basic OASIS-Pipe module in which the synchronization was achieved through blocking I/O on pipe files is not suitable in a distributed mode. It has been replaced by CALCIUM developed by EDF/DER, based on PVM and dedicated to synchronization and data transfer. The interpolation module (OASIS-Interp) is kept whatever the configuration. Figure 4a describes the new architecture of the distributed coupled model. ARPEGE and OPA run on different sites, the OASIS coupler being split between the two CRAYS, according to its modular concept (OASIS-Interp at EdF and CALCIUM at Météo-France).

Results obtained with this specific architecture do not show evidence of significant improvement in reducing the wall clock time. The dispersion of the coupling modules is clearly a disadvantage and leads us to build a

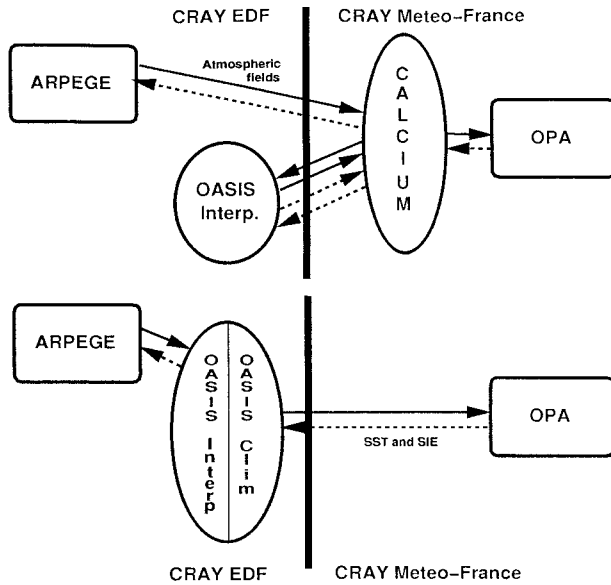


FIG. 4. Code distribution between the two supercomputer centers.

new OASIS version (denoted OASIS-Interp/CLIM), including a specific library (CLIM) also based on PVM and devoted to communication in a distributed context (see appendix and Fig. 4b). CALCIUM is basically built on a centralized concept whereas CLIM is built on a dispersed one.

b. Performance analysis and load balancing for both configurations

A matching CALCIUM configuration file has been built, describing the whole data flow between the two codes (ARPEGE/OPA) and OASIS-Interp as well as separate shell scripts to monitor the jobs in an automatic boost mode (including post processing of data and day-to-day storage management). Most of the development and debugging effort has been done before the real beginning of the climate experiments. The OASIS-Interp/CALCIUM coupler is then validated by 1-yr distributed simulation with the OGCM running on the CRAY II at Météo-France and the AGCM running on the CRAY C98 at EDF. A 1-yr integration has been achieved with 85 simulations of 5 days, the 12 more than expected being due to network time out, multitasking deadlocks, script rewriting, and PVM tuning. Nevertheless, the dis-

tributed computing has induced a significant decrease in the duration of the experiments because each GCM could reach batch queues with a greater priority.

Several criteria could be retained for evaluating the performance of a distributed application: a comparison with a one-host one-program system that is not always available but that one can simulate; a comparison with a one-host multiprogram system, or finally, a qualitative critique regarding what one would expect to achieve in a distributed environment: has the application taken the best of the present potentialities?

The relevant information is the total elapsed time for simulating a given *real* time (restitution time). This performance depends on the intrinsic performance of the coupling mechanism itself, the robustness of the global application and its capacities to match the constraints of operational supercomputers (available resources, jobs system queueing). An integrated ocean-atmosphere system is obviously the more efficient way of coupling but it requires so much resources (particularly in terms of memory) that it is very often overdelayed by classic queuing mechanisms such as Network Queuing System (NQS) on CRAY computers. To distribute the application (even on the same host) is undoubtedly a clear advantage from this point of view (as well as pre- and postprocessing sharing). But, by creating several processes and point-to-point communications, the distributed mode makes itself more fragile than the integrated one (remote host shutdown, network failure, process abort).

Table 1 describes the elapsed time for each time step of a 5-day run of the atmospheric model, depending on the types of couplers. It points out that the addition of the synchronization and communication time is more expensive using the CALCIUM tool due to the duplication of data transfer between the CALCIUM executable and the models and overloading of the machine by the master process.

The CALCIUM paradigm (see appendix) is to consider the application as logically centralized even if it is physically distributed (Fig. 4). The main CALCIUM executable, which reads the configuration file describing the links between connection points as well as their localization, behaves as a mailbox where all the codes send data and where they ask for incoming data. Therefore, each field is sent twice. This mechanism leads to a nonoptimal efficiency, particularly if we “cross” a long distance network (or a slow one). We gain little

TABLE 1. Elapsed time(s) for 5-day distributed runs. Times for the coupler (cpl), the AGCM (atm), and OGCM (oce) include the periods in which these programs are idle. Differences are due to the coupling technique or the computer power. 1) NQS bench class for standard version, idle machine otherwise; 2) C98 METEO idle, C98 EdF running less than five jobs; 3) both machines used at a rate of 40%.

Configuration	Pipes			CALCIUM			CLIM		
	cpl	atm	oce	cpl	atm	oce	cpl	atm	oce
1) 1C98	275	317	277	339	383	287	295	303	253
2) 2C98 Renater	—	—	—	640	1105	584	267	298	272
3) C98 + C2 HiPPI	—	—	—	—	—	—	635	642	625

benefit from the remote distribution of the codes, because the increased elapsed time due to the field transfer through the network is not offset by the input waiting time of each model. The OASIS-Interp/CLIM coupler for the ocean–atmosphere coupled model appears to be much faster and more efficient than the OASIS-Interp/CALCIUM one.

The main advice we can give is to build an architecture as simple as possible, limiting exchanges and communications through the network. It is usually worth adapting the configuration of each code even if it is heavier in terms of computation than adopting a wide structure as we did with CALCIUM. The crucial point is obviously the network, on which the whole system depends. So, the distribution is feasible but requires a very high performance high-speed network combined with a well-adapted structure of codes. The above-described experiments are considered benchmarks; future improvements are expected with a fully parallelized model and a dedicated network.

4. Sensitivity to the precision in exchanged data

Coupled global ocean–atmosphere circulation models are demonstrated to be highly sensitive to the precision in coupling fields as well as initial conditions (Stockdale et al. 1994). Different precisions were used in the experiments of validation for OASIS-Interp/CALCIUM and OASIS-Interp/CLIM. In the former case, fields were written and read through ASCII files between the oceanic GCM and the coupler and through GRIB format files between the atmospheric GCM and the coupler. In the latter, full precision is kept for all the coupling fields. The divergence between the two experiments highlights the sensitivity to the precision in exchanged data and leads us to focus on the so-called “butterfly effect” in coupled mode (Lorenz 1969). For instance, this butterfly effect can arise from the change of compiler and mathematical libraries and therefore appears in distributed mode.

To assess this issue, perturbations are intentionally introduced into the model. They consist in degrading the surface boundary conditions (forcings) of each separated component of the coupled model, namely the SST seen by the AGCM and the surface fluxes of heat, momentum, and freshwater imposed on the OGCM. These perturbations are applied at different stages of the coupling. Two types can be distinguished:

- Perturbations on initial forcing fluxes (IFFs), where surface boundary conditions are modified for the first time step (1 day) of the coupling. The general idea is to test if the finite precision of the initialization data has a significant influence on the climate modelling.
- Perturbations on coupled forcing fluxes (CFFs), where surface boundary conditions are modified at each time step of the coupling.

As a simplifying procedure, we have chosen to perform simulations on a single CRAY: the CRAY C98 of

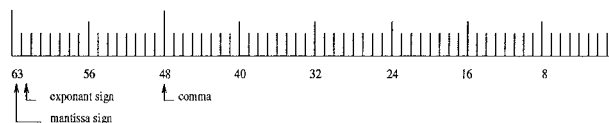


FIG. 5. CRAY representation of a real constant.

Météo-France. This consolidation allows us to get rid of the constraints of the distributed context. These constraints are obviously worth overcoming for long simulations but in the present case, sensitivity tests are analyzed on 10-day simulations up to 1-yr simulations. It is therefore more efficient to perform them on a single machine. A special simulation, for which full precision is retained, is also carried out on the CRAY II of Météo-France, in order to evaluate the effect of a supercomputer change.

a. Type of perturbations

Before describing the different types of perturbation, it is essential to keep in mind the CRAY representation of a real constant. A basic real constant is written on 8 bytes of storage, that is, 64 bits, (schematically represented in Fig. 5). The CRAY word is split into three parts: bits 0 to 47 represent the mantissa of the real constant, bits 48 to 61 are used for the exponent, and bits 62 and 63 are attached to the exponent sign and the mantissa sign, respectively.

We have perturbed this particular representation by changing the last 20 bits of the mantissa. This choice is arbitrary and consists of replacing by 0, bit 0 to bit 19. The motivation for this scheme is to bring in the model artificial perturbations, which are roughly equivalent, to convert a simple-precision CRAY real into a simple-precision real on a workstation. The advantage of such a perturbation relies also on its simplicity. But it is worth noting that the resulting error is quite difficult to evaluate. It is impossible to assess how many significant digits are represented on these 20 perturbed bits. For instance, 6 is coded on 3 bits whereas 9 requires 4 bits. Besides, this artificial perturbation depends on the sign of the exponent. For example, the binary representation of 10 is not altered by the perturbation whereas 0.1 is turned into 0.09999996. So, this method induces a nonuniform perturbation, which introduces an error at most equal to 10^{-7} for a number ranging from 0 to 1.

b. Butterfly experiments

Figure 6 sketches the coupled structure of the model, focusing on the artificial perturbation introduced at several stages of the simulation. The values ε_o^i and ε_a^i symbolize the perturbation and are equal to S (for SST) or F (for flux), respectively, when the perturbation is applied. The exponent i denotes the time step (in days) corresponding to the coupler time step, a and o stands for atmosphere and ocean, respectively.

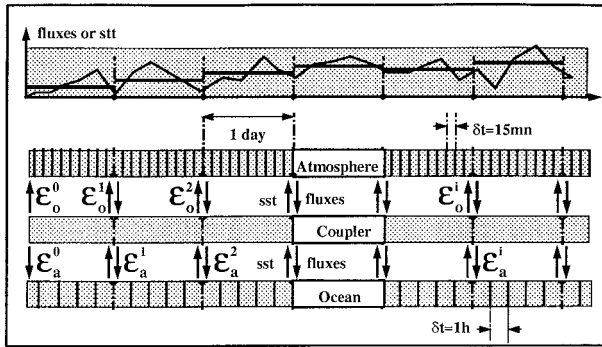


FIG. 6. Schematic representation of the artificial degradations. Perturbations ($\epsilon^i_{a/o}$) are applied after the interpolation on the forcings of each models (F averaged over 1 day for the ocean, SST for the atmosphere).

Experiments are named CT n with n being the number of the experiment (see Table 2). The various perturbations consist therefore in degrading the IFFs and/or the CFFs with the distinction between the oceanic and atmospheric forcings. The sensitivity to IFF perturbations has been first tested: only $\epsilon^0_{a/o}$ is active and equal to S/F. We then study the sensitivity to the CFFs perturbations which are performed in OASIS after each interpolation. Experiments can be easily identified by the four characters of Table 2, indicating the nature of the IFFs (first two characters) and the CFFs (last two characters) perturbations.

As reported before, simulations were performed on the CRAY C98 of Météo-France, except for CT7, which was carried out on the CRAY II with no degradation. Comparisons are made between a perturbed experiment and control experiment CT3 (full precision retained). The change in the type of supercomputers (compiler, scientific libraries) is tested by comparing CT3 and CT7,

which are identical twin with respect to the precision in exchanges and initial forcings.

The runs CT3, CT5, CT6, CT7 were extended to 1 yr in order to investigate the perturbation growth over a longer period. Run CT4 is also a control experiment (full precision) using OASIS-Interp/CLIM instead of using OASIS-Interp/CALCIUM. Giving strictly identical results compared to CT3, we have stopped this experiment after six months. The main purpose of CT4 was therefore to cross validate the CALCIUM and CLIM tools.

c. Control from selected grid points

The perturbation growth in coupling fields are monitored on 12 oceanic-atmospheric grid points. These points have been selected according to climatic areas around the planet. They are located on two longitudes centered on the Pacific (180°) and Atlantic Oceans (30°W) and three latitudes related to middle (60°), tropical (30°), and equatorial (1°) regions. Six points are therefore chosen in the Northern Hemisphere and six points, for symmetry, in the southern one. These grid points are represented in Fig. 7, superimposed on the CG3 experiment SST variance.

For each grid point and each time step, the SST value is extracted from the oceanic model whereas the solar flux, the total heat flux, the zonal and meridional wind stress, and the water flux come from the atmospheric model.

d. Temporal analyses

The general problem of sensitivity to the precision in exchanged data is discussed in terms of error growth. By definition, the instantaneous error E is given by the expression

TABLE 2. Overview of the perturbation experiments. All experiments have been run during 10 days, except CT3, CT5, CT6, and CT7, which have been extended to 1 yr. Perturbations are indicated by the letters S (for SST) or F (for flux) when applied at the first coupling time step (one day) or at the following ones. The sequence of these four characters is used to characterize the experiment, such as CT5(SF--) or CT8(SFSF). A classification of the experiments is performed according to the nature of the IFFs perturbations, that is, the first two characters of the coding. Three classes of experiments can thus be considered: SF family (CT5, CT6, CT8) where all the IFFs are perturbed (fluxes and SST); the -F family (CT10, CT12) for which only initial atmospheric fluxes are biased; and the S- family (CT9, CT11) for which only initial SST is perturbed.

Experiments				IFF Initial exchanges $i = 0$		CFF Next exchanges $i > 0$	
Control	Perturbed	Coupler	CRAY	$\epsilon^0_{a/o}$ SST	$\epsilon^0_{a/o}$ Fluxes	$\epsilon^i_{a/o}$ SST	$\epsilon^i_{a/o}$ Fluxes
CT3		CALCIUM	C98	—	—	—	—
CT4		CLIM	C98	—	—	—	—
	CT5	CALCIUM	C98	S	F	—	—
	CT6	CALCIUM	C98	S	F	—	F
CT7		CALCIUM	C II	—	—	—	—
	CT8	CALCIUM	C98	S	F	S	F
	CT9	CALCIUM	C98	S	—	—	—
	CT10	CALCIUM	C98	—	F	—	—
	CT11	CALCIUM	C98	S	—	S	—
	CT12	CALCIUM	C98	—	F	S	F

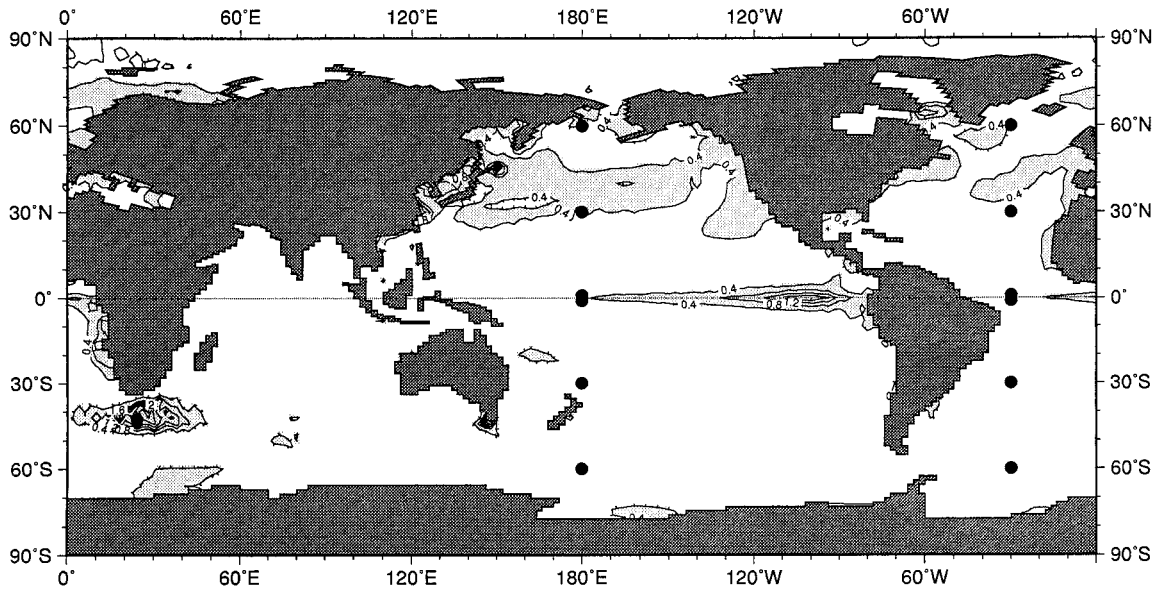


FIG. 7. Selected grid points centered on the Pacific (180°) and Atlantic Oceans (30°W). They are superimposed on the CG3 experiment variance of the monthly averaged SST computed from the last 15 simulated years.

$$E(t, x_i, y_i) = |E_{\text{pert}}(t, x_i, y_i) - F_{\text{cont}}(t, x_i, y_i)|, \quad (1)$$

where x and y , respectively, represent the latitude and the longitude, and t the time step. The value F stands for the different atmospheric fluxes and for the oceanic SST; F_{pert} refers to the perturbed experiments, and F_{cont} to the control experiment CT3. The error is therefore calculated at each time step, for each grid point represented by i , and each coupling field.

For a given grid point, the fluctuation of E defined by Eq. (1) is considerable, due to the complexity in the dynamics of the coupled system (Nicolis 1992). To obtain a more representative growth rate, errors are thus averaged over the 12 selected points or with the distinction between the Northern and Southern Hemisphere, namely averaged over the six points belonging to the corresponding hemisphere. In addition to this arithmetic average $E(t)$ defined by

$$E(t) = \frac{1}{12} \sum_{i=1}^{12} |F_{\text{pert}}(t, x_i, y_i) - F_{\text{cont}}(t, x_i, y_i)|, \quad (2)$$

it is relevant to introduce the logarithm of the error as a new variable. The error growth is assumed to be modeled by the first-order differential equation (Lorenz 1969):

$$\frac{dE}{dt} = \alpha E \left(1 - \frac{E}{E_{\infty}} \right), \quad (3)$$

where α is the growth rate and E_{∞} the saturation value of the error.

Considering that $E/E_{\infty} \ll 1$ at the beginning of the simulation, Eq. (3) can be rewritten as

$$\begin{aligned} \frac{1}{E} \frac{dE}{dt} &= \alpha = \frac{1}{\tau} \Leftrightarrow E(t) = E_o \exp^{\alpha t} \Leftrightarrow E(t) \\ &= E_o \exp \frac{t}{\tau} = E_o 10^{t/t_*}. \end{aligned} \quad (4)$$

This equation is the one used to characterize the initial growth state of the error. The *10-folding time* t_* is directly given by the slope of the $\log_{10} E = E_o + t/t_*$ curve and the *doubling time* τ is related to t_* by the relation $\tau = t_* \log_{10} 2 \approx 0.3t_*$.

The error growth is investigated over two timescales: 10 days and 1 yr. The error grows dramatically over the first 10 days before reaching a plateau. The fluctuation around this plateau is investigated over 1 yr. It is worth separating the two timescales, each involving transient and steady mechanisms, respectively.

5. Error growth in the 10-day simulations

Figures 8a and 8b show the 10-day evolution of the logarithm of the error of the oceanic SST (OSST; i.e., extracted from the oceanic model) for all the types of perturbation. It is clear from this figure that any initial error reaches a nearly common value after 10 days. Moreover, the intensity of the initial error is not crucial (not shown) and does not change the trend of the growth. So, whatever the type of the perturbation, the behavior of all these experiments are both qualitatively and quantitatively similar.

a. Two classes of experiments

Two classes of experiments can be distinguished. The first class is characterized by a big jump in the OSST

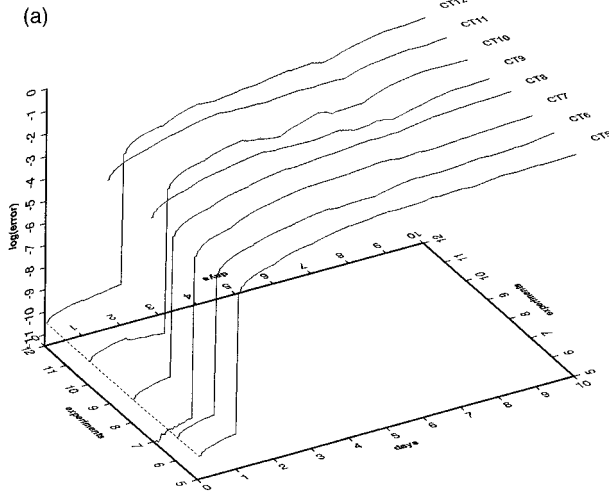
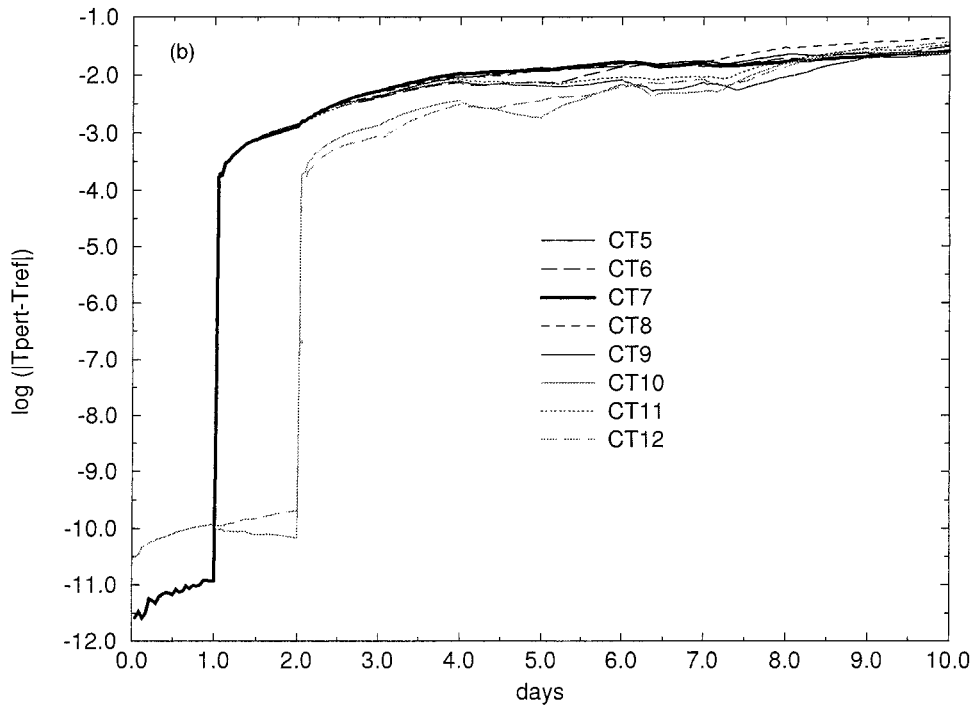


FIG. 8. SST error growth over 10 days, depending on the perturbation type. Panel (a) shows, in three dimensions, the logarithm of the error as a function of the integration time (in days) depending on the experiments. Panel (b) shows the same data in the plane ($\log E$, days).



error just after the first coupling step (i.e., first day). This class corresponds to the (SF) and (S-) families of experiments for which the initial SST given to the atmosphere is initially perturbed. The error growth in the atmospheric component during the first day is responsible for this big OSST jump and the intrinsic error growth of the oceanic model seems to be negligible compared to it.

In the second class of experiments, the big OSST error jump only occurs after the second coupling step (i.e., second day). This class corresponds to the (-F) family of experiments for which the initial SST is unperturbed. Thus, the first day of simulation of the atmo-

sphere is error free, and the atmospheric flux passed to the ocean after one day is unperturbed.

The features of these two classes of experiments show that the atmospheric model is very sensitive to SST anomalies and that the global error is mainly driven by the perturbations performed on the atmospheric SST (ASST; surface temperature seen by the atmospheric component). This means that the AGCM is responsible for the error growth in the coupled system for short time range simulation.

The saturation and the growth rate seem to be independent of additional perturbations. Comparisons between experiments where fields are degraded at each

coupling show that the first initial ASST perturbation has a significant effect whereas a second and further degradation at each coupling is without relevant influence.

We now present a detailed analysis of these two classes of experiments through the study of the CT5 (SF-) and CT10 (-F-) experiments.

1) THE (SF) AND (S-) EXPERIMENT FAMILIES

The CT5(SF-) is representative of the (SF) and (S-) families of experiments. Three different stages can be distinguished for the evolution of the error: the initial stage related to the error induced by perturbed fluxes given to the ocean, the intermediate stage after the first coupling, and the final stage corresponding to the saturation of the error.

At each stage, the error dynamics is fitted by Eq. (4). Figure 9 details the three periods for the OSST. The first stage extends up to 1 day: the doubling time is $\tau = 14$ h (i.e., $t_* = 1.9$ days). Fourteen time steps (1 h each) are thus necessary for the oceanic model to double the error induced by the perturbed atmospheric fluxes. A big jump of seven decades in the OSST error occurs just after the first coupling and within one time step. A value of $\tau = 1$ h is thus only an upper bound.

An intermediate period between days 1 and 3 is characterized by a doubling time similar to the $\tau = 14$ h of the first day (here $t_* = 2.2$ days). This intermediate period is followed by a slower error growth up to day 20, with $\tau = 5$ days (i.e., $t_* = 16.4$ days), where a plateau is reached.

Figure 10 describes the error growth of the horizontal wind stress and confirms the high sensitivity of the atmospheric model. Figure 10a shows that the error growth is rapid and tends quickly to a first saturation threshold with $\tau = 20$ min (i.e., $t_* = 0.04$ days). The long-term saturation is seen on Fig. 10b after 20 days. During this final stage, it is interesting to observe that the atmospheric doubling time $\tau = 5$ days is exactly the same as the oceanic one. This agreement is reasonable considering the coupling between the two models.

The CT7(CRAY II) experiment shares common features with the (SF) experiment family. The errors related to the use of the CRAY II compiler and the mathematic libraries give rise to similar behavior compared to other perturbations (see Fig. 10b).

2) THE (-F) EXPERIMENT FAMILY

We now turn to the CT10(-F-) experiment, which is representative of the (-F) family. Figure 11 highlights three specific time regions of the OSST errors over the ten simulated days. The first region, during day 1, is strictly similar to the one of the CT5 (SF-) experiment. The error growth is related to the perturbed initial fluxes sent to the ocean. However, the atmosphere is not perturbed during the first day and the atmospheric fluxes

sent to the ocean after the first day of integration are "error free." Thus, the OSST error of the CT10 experiment decreases during the second day (see Fig. 8b or Fig. 11) and results in a negative error growth with $\tau = 2$ days (i.e., $t_* = 6$ days). The growth of OSST error in the CT12 (-F-F) experiment (see Fig. 8b) is only due to the small perturbation imposed on these "error free" fluxes during the first coupling.

The last regions are defined from day 2 to day 4, and from day 4 to day 10. These regions are featured by steeper slopes compared to CT5(SF-). This is particularly emphasized for the second one, where $\tau = 1.7$ days (i.e., $t_* = 5$ days) compared to $\tau = 5$ days (i.e., $t_* = 16.4$ days) for CT5(SF-). However, it is obvious from looking at the curve that the method of determining all these growth rates is not very precise. It is therefore more relevant to deal with qualitative conclusions. The main one that can be drawn is that the general tendency is maintained, whatever the simulation.

b. Interpretation of the error growth

We now give an interpretation of the above observations based on studies of the error growth in GCMs (e.g., Stroe and Royer 1993). Each component of the coupled system is characterized by its own attractor defined by the boundary conditions (fluxes for the OGCM, SST for the AGCM). Perturbations introduced into the forcing induce therefore an attractor change corresponding to a perturbed statistical equilibrium state. An initial condition of the OGCM or AGCM will first converge to the new attractor. This convergence will first induce a growth in the error. If the perturbation of the attractor is small, the growth rate will be the linear convergence rate to the attractor. If the perturbation is larger, the error growth rate can be fast at first, due to a nonlinear convergence to the attractor, followed by the linear convergence rate. Such convergence to a perturbed OGCM attractor can explain the OSST doubling time value $\tau = 14$ h for small artificial perturbations of the fluxes (linear convergence) and the $\tau = 1$ h value for larger ones (nonlinear convergence) such as the ones resulting from an atmospheric error amplification. Similarly, the convergence to a perturbed AGCM attractor can explain the wind stress doubling time value $\tau = 20$ min (nonlinear convergence) for small perturbation of the ASST.

In addition to this change of attractor, the perturbed evolution will diverge exponentially from the reference one because of the intrinsic sensitivity to initial conditions on the attractor. This divergence explains the value $\tau = 5$ days observed in both the OSST and wind stress error. It seems to be only due to the divergence of trajectories on the atmospheric attractor, since this value is typical for uncoupled AGCM experiments. The divergence of the oceanic attractor, as could be measured on uncoupled OGCM experiments, is slow if any.

In conclusion, these experiments show that the general behavior of the error growth over 10 days is driven

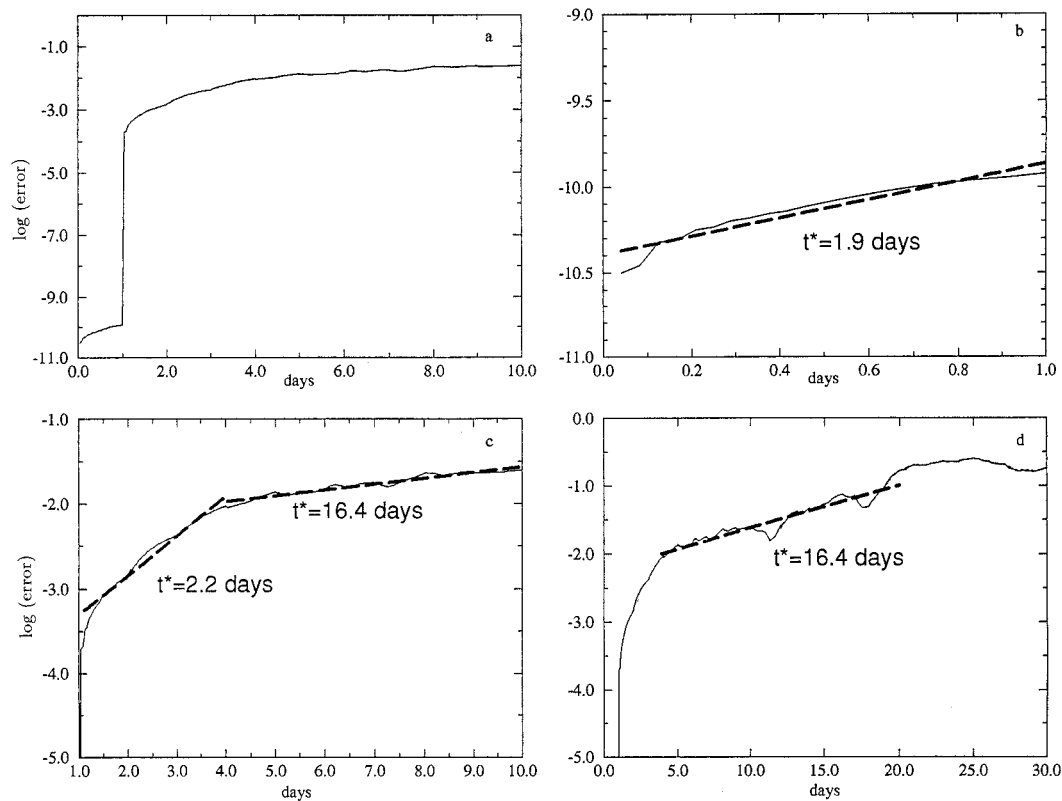


FIG. 9. SST error growth for the CT5(SF--) experiment. The logarithm of the error is represented as a function of time. The time t_* is given in days for the different stages: (a) error growth over 10 days; (b) error growth between the beginning of the experiment and the first coupling; (c) error growth between the first coupling and 10 days; and (d) error growth over 30 days.

to a large extent by the atmospheric component. This is a direct consequence of the highly nonlinear and chaotic character of the atmospheric dynamics, which is reinforced by the very large inertia of the ocean, compared to the atmosphere (Thompson 1957). Rather than an intensive study of predictability in coupled mode the present work identifies qualitative characteristics of the perturbation transient growth.

6. Error growth in the 1-yr simulation

a. Temporal analyses on the 12 selected points

The CGCMs vocation of long-term studies is now investigated. The error growth is studied by monitoring its time evolution over one year. We subdivide the 12 selected points into the two hemispheres for which averages are carried out from six points. Figure 12a describes the OSST behavior for the CT5(SF--) experiment. The long-dashed curve represents the time evolution of the error growth in the Southern Hemisphere, whereas the dotted line corresponds to the northern one. The solid curve represents the usual average over the 12 selected points. Attention must be given to the fact that the number of points is not high enough to assess a general behavior for the whole hemisphere. Never-

theless, this study will give a trend of the response for the different hemispheres.

After a characteristic time of one month, the error reaches a saturated value and seems to oscillate around a plateau. But the most striking effect is the hemisphere-dependence of this oscillation. The sensitivity is greater in the summer hemisphere than in the winter one. Maximum errors are observed from July to September in the Northern Hemisphere, whereas higher values are noticed in the southern one, from January to March and October to December. The sensitivity therefore follows the seasonal cycle, but it is interesting to note the higher variation in the Northern Hemisphere compared to the southern one. This difference is related to the stronger seasonal cycle in the Northern Hemisphere due to the presence of land and strong oceanic currents in the North Atlantic and North Pacific Oceans. Moreover the Pacific point located at 60°N is frozen during the winter. In that case, we should take into account that fluxes of frozen points are prescribed in the experiments. So, due to this climatological "prescription," any fluctuation is inhibited and the average error may be therefore artificially weakened.

Experiment CT7(CRAY II) exhibits the same shape as CT5(SF--) (Fig. 12b). The seasonal cycle is less pro-

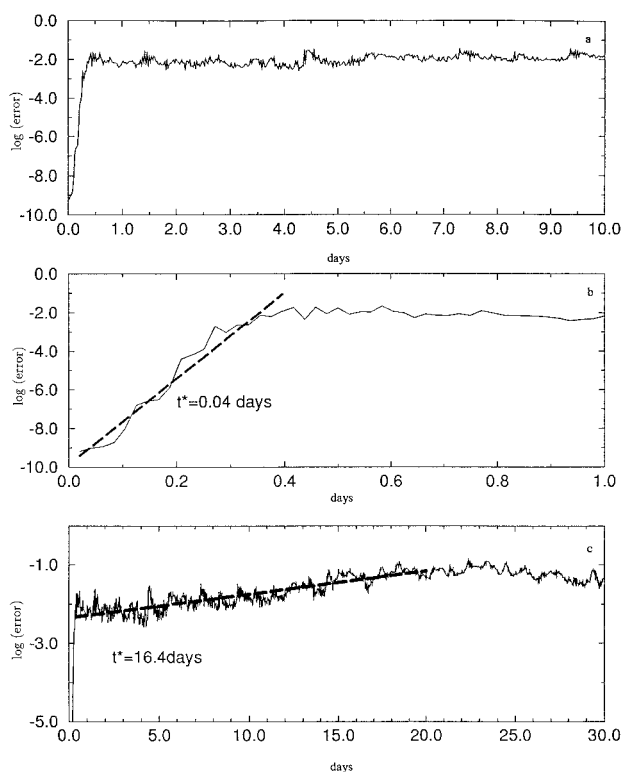


FIG. 10. Wind stress error growth for the CT5(SF-) experiment. The logarithm of the error is represented as a function of time. The time t_* is given in days for the different stages: (a) error growth over 10 days; (b) error growth between the beginning of the experiment and the first coupling; and (c) error growth over 30 days.

nounced but remains significant enough to be highlighted. This confirms the error independence on the type of the initial perturbation, since no difference can be found between artificial perturbations and intrinsic perturbations due to the change of remote machines.

Similar studies have been carried out with atmospheric fields. As shown in Fig. 13a, the solar flux exhibits a weak seasonal cycle whereas the zonal wind stress is more constant over the 12 months with a higher sensitivity in the southern hemisphere (see Fig. 13b). The seasonal cycle of the wind is known to be weak and can explain the weak seasonal oscillation of the error. On the other hand, the permanent higher sensitivity in the southern hemisphere is certainly explained by the wind intensity, which is relatively stronger over the year than in the northern one.

Goswami and Schukla (1991) assert that any initial error small or large would finally saturate at the value corresponding to the model's natural variability. All these results tend to confirm this hypothesis.

b. Spatiotemporal analyses over the planet

1) ZONAL AVERAGES

Figure 14a represents the zonal average of the OSST error for the CT5(SF-) experiment. It shows that the

one-month difference remains relatively small (below 0.6° at any latitude). However, the error tends to become quite large in the Northern Hemisphere at the beginning of May. After May, the error diverges rapidly and significantly between the 40° to 70°N , before reaching a plateau in July. The largest bias (2.2°) remains steady until the end of August and is centered between 50° and 60°N . Beyond August, the difference decreases rapidly to value of about 0.4° with a relatively narrow and low maximum in the Southern Hemisphere with 0.6° between 25° and 50°S . The largest difference occurs therefore in the Northern Hemisphere during summer and is located in the midlatitudes. No similar variations are seen in the Southern Hemisphere. It is partly due to a stronger seasonal cycle in the Northern Hemisphere and partly due to the presence of physical phenomena (western boundary currents) induced by the presence of land. The signature of higher sensitivity during the summer in the Northern Hemisphere is consistent with the previous section where 1-yr studies gave similar results.

Figure 14b examines the CT7(CRAY II) OSST error and draws similar conclusions. The general shape is retained with a maximum error in the Northern Hemisphere at midlatitudes and during the summer.

Similarly, Fig. 15 presents the OSST zonal average in terms of variance for the CG3 experiment (see section 2a), which refers to the T42 experiment. The variance σ is calculated over the last 15 years and is defined as

$$\sigma_m(x, y) = \frac{1}{N-1} \sum_{i=1}^N [\text{SST}_m(i, x, y) - \overline{\text{SST}_m(x, y)}]^2, \quad (5)$$

where i denotes the year ($N = 15$), x and y the latitude and the longitude, respectively, m the given month ($m = 1-12$), SST_m the monthly average SST of month m during a particular year, and $\overline{\text{SST}_m}$ the monthly climatological SST averaged over the selected 15 years. The variance is used to represent the interannual variability of the coupled model, that is, its natural variability. This “model variability” includes “random” interannual variability, such as the one due to atmospheric variability at midlatitudes, as well as “coherent” interannual variability, such as the one due to interannual coupled variability of the equatorial Pacific.

We observed that the maximum variance occurs at the same latitude, the same hemisphere, and the same period as the maximum error. These comparisons lead us to think that the most sensitive areas are related to regions of important model variability. Tropical areas are characterized by a very stable equilibrium state from where the model cannot be taken away. This might explain the very weak difference between 10° and 30° in both hemispheres. In contrast, several equilibrium states are possible in the midlatitudes and the model converges indifferently to one of them. Small perturbations would, therefore, have initiated a transition between two states.

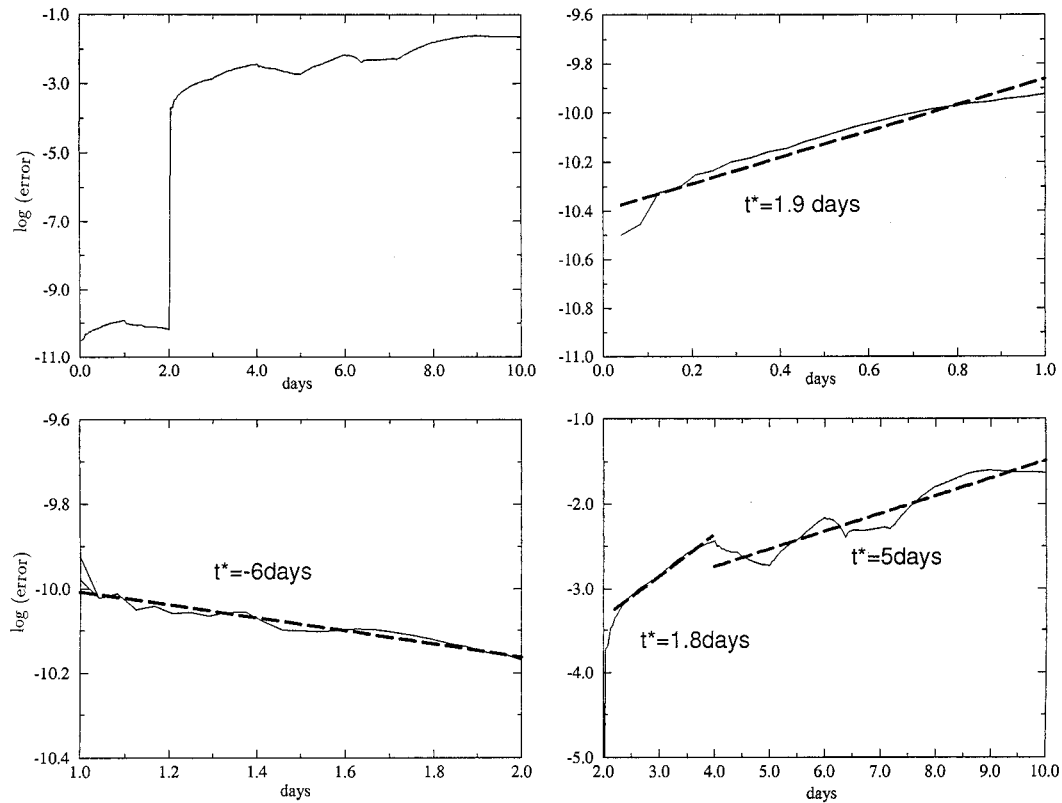


FIG. 11. SST error growth for the CT10 (-F-) experiment. The logarithm of the error is represented as a function of time. The time t_* is given in days for the different stages: (a) error growth over 10 days; (b) error growth between the beginning of the experiment and the first coupling; (c) error growth between the first coupling and the second coupling; and (d) error growth between 2 and 10 days.

It is also striking that the important model variability of CG3 around the equator is not present in the error analysis. This model variability in this area is assigned to El Niño events. Some studies (Cane et al. 1986; Cane and Zebiak 1988; Barnett et al. 1988) have indicated that there might be some seasonality in the growth of

errors in the coupled system. In order to obtain a more detailed examination of the error growth, it would be interesting to start a new simulation in July. It would allow us to check if the error in the Northern Hemisphere reaches similar values in summer and if the weak error in the southern one might be explained by the

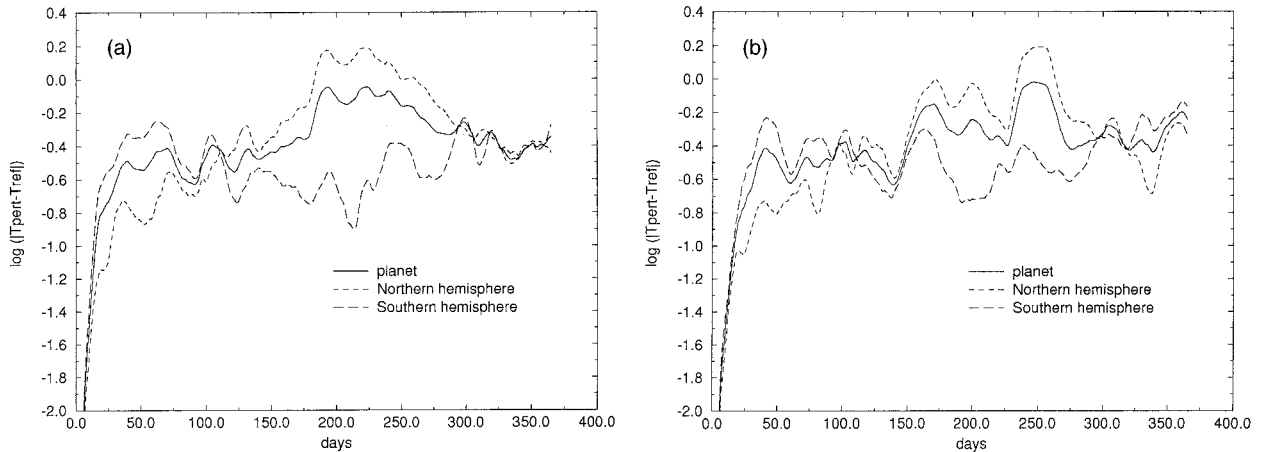


FIG. 12. SST error growth over 1 yr depending on the hemisphere. Averages are performed over 6 points for each hemisphere and 12 points for the planet curve: (a) CT5(SF-) experiment, (b) CT7(CRAY II) experiment.

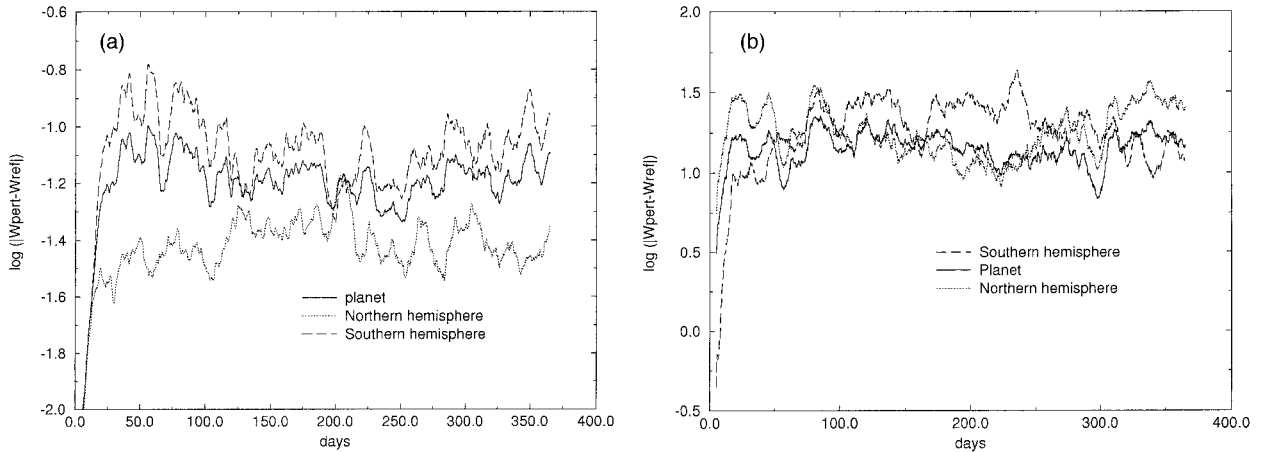


FIG. 13. Error growth over 1 yr depending on the hemisphere, for the CT5(SF--) experiment: (a) wind stress, (b) solar flux.

initial phase of the error growth, which is too short to be efficient.

Moreover, attention must be given to the fact that the number of oceanic points in the Northern Hemisphere at high latitudes is smaller than in the Southern Hemisphere.

This may indirectly enhance the imbalance noted in the zonal average of the variance between the two hemispheres.

2) GEOGRAPHICAL STUDIES

In the previous section, we attempted to explain the monthly dependence of the OSST error growth: this has been done by comparing zonal averages of CG3 experiment with CATHODE CTn experiments (see section 4). Comparisons between longitude–latitude charts of OSST error and CG3 charts of OSST variance are now presented.

January is the first simulated month. As expected and shown in Fig. 16a, the error remains small over the oceans. Some patchy features are visible on the west side of the Pacific Ocean and in the center of the Atlantic, for the mid-southern latitudes. In contrast, significant cores of error are present in July in the Northern Hemisphere (Fig. 16b). They are mainly located close to the warm oceanic currents such as the Kuroshio and the Gulf Stream. It is worth noting that a positive anomaly is always connected to a negative core, which grows in parallel and very close together. The major aspects of the midlatitude variability, such as the variability in the Gulf Stream areas, are mainly governed by the instability of the mean ocean current. So, the coupling between positive and negative anomalies suggests that the currents may have changed their location compared to that simulated in nonperturbed experiments: they do not extend toward the west and remain confined in the western part of the ocean. A cooling effect is therefore induced in the eastern part of the basin.

Comparisons with the CG3 charts of variance (Fig. 17a for January and Fig. 17b for July) confirm this hypothesis. The model variability is important in the zonal motion of the oceanic drifts. Small perturbations on exchanged fluxes would therefore have triggered a cold year in the eastern part of the ocean and forced

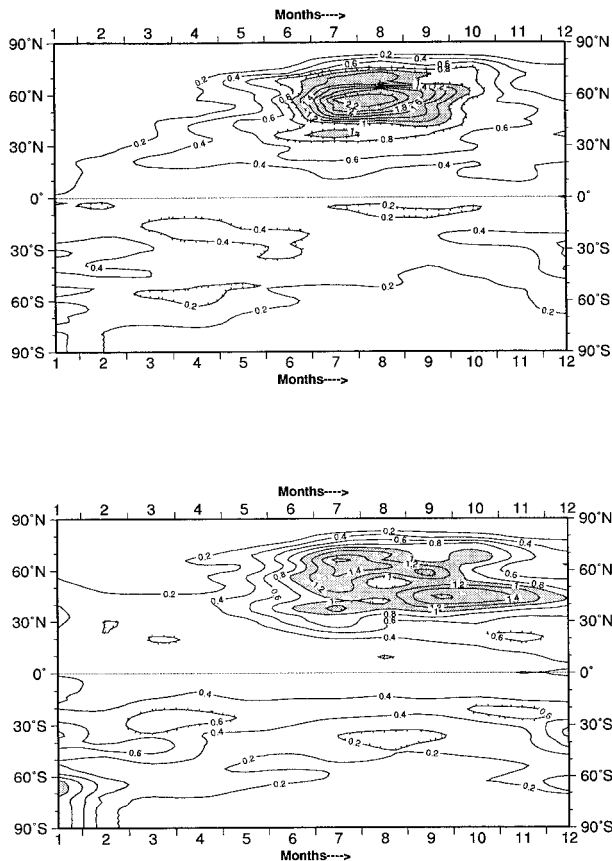


FIG. 14. Latitude–time diagram of the monthly and zonal average of the SST error, over 1 yr. Values greater than 1° are shaded and contour interval is 0.2°. (a) CT5(SF--) experiment, (b) CT7(CRAY II) experiment.

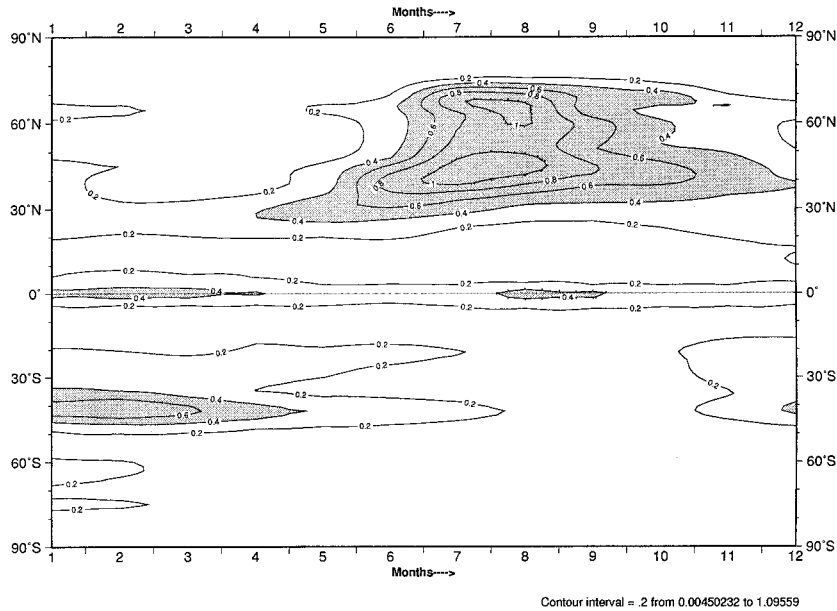


FIG. 15. Latitude-time diagram of the monthly and zonally average of the SST variance over the last 15 simulated years of CG3. Values greater than 0.4 are shaded and contour interval is 0.2.

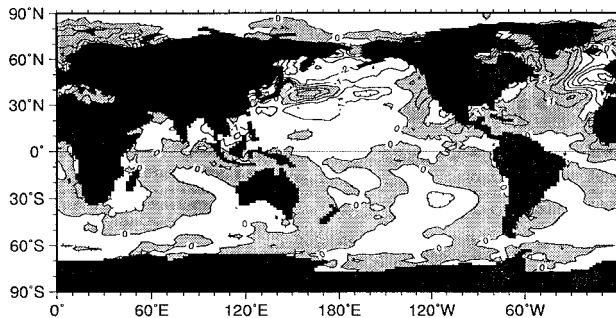
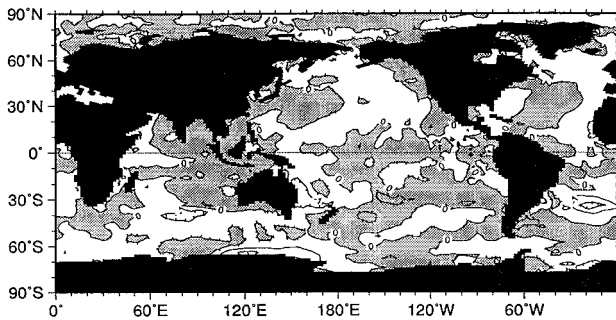


FIG. 16. Difference map between the SST of the CT5(SF--) experiment and the SST of the CT3(--) control experiment. Values greater than 0° are shaded contour interval is 1° (a) January and (b) July.

the Kuroshio and the Gulf Stream to adopt a new equilibrium state. Finally, perturbations seem to have a significant effect both on large- and small-scale because patchy error cores in the Southern Hemisphere in January are associated with oceanic eddies and are not related to large currents like in the northern one.

We have shown that after 1 yr, the error level is of the order or below the variability of the system. This study points out that it is not necessary to keep all the exchanges at full precision. The truncation from 64 to 32 bits is not critical for the long-term behavior of the model. Whatever the case, simulations still give results in agreement with the equilibrium state of the whole system in terms of model variability, and intensity of oceanic and atmospheric fields. The loss in precision gives rise to fluctuations (cold/warm phases) around these steady states but the statistical properties are not affected by the precision of exchanged data. Long-term studies do not require full precision because the fluctuations between states are averaged over the period. Above all, the imprecision on observations is dramatically higher than the error introduced by the truncation from 64 to 32 bits.

7. Conclusions

A first goal of the CATHODE project has been to realize a distributed CGCM. A climate model was decomposed into atmospheric and oceanic components that were located at different computing sites. The model consisted of a high-resolution AGCM (ARPEGE T42) coupled to a high-resolution OGCM (OPA) via the different configurations of the OASIS coupler. Different

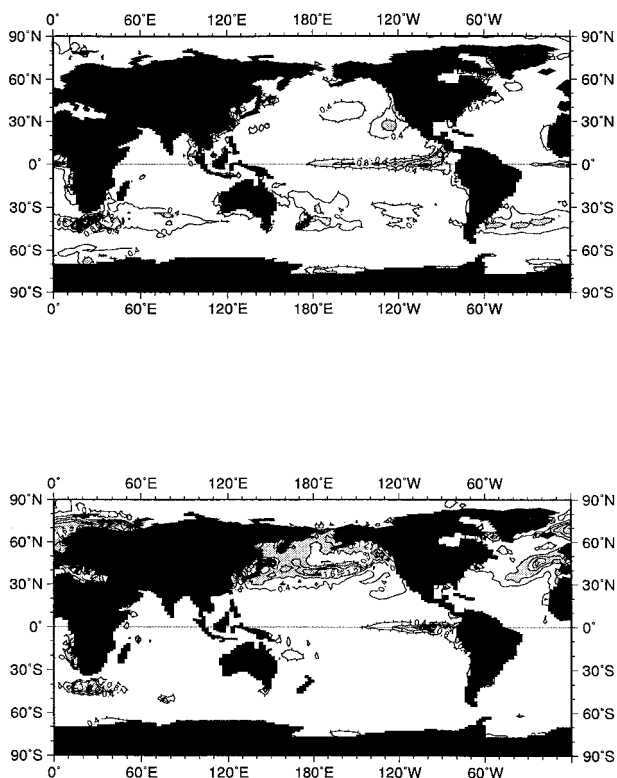


FIG. 17. Variance averaged over the last 15 simulated years of the CG3 experiment. Values greater than 0.8 are shaded and contour interval is 0.4: (a) January and (b) July.

CRAYs have been used: CRAY C98 and CRAY II for the atmosphere and the ocean, respectively. Communication between the two remote supercomputers was ensured by the French RENATER network.

Real-case climate simulations have been carried out in a distributed environment, with their complete operational structure, with acceptable security modification and restitution times. The CALCIUM module was validated by 1-yr simulations. Nevertheless, its general nature leads to low efficiency in the ocean–atmosphere coupling framework. A second version of the CALCIUM tool will benefit improvements (possibility of task-to-task communication), but the specification of duplicated messages still remains, leading to a nonoptimal efficiency in our case.

Then, a general purpose coupling library for interfacing models (CLIM) has been designed, based on dynamic link creation and deterministic behavior. Experiments using OASIS-Interp/CLIM show a good reliability and an improvement in efficiency. The distributed coupling could really become an interesting alternative to standard solutions with a very high-performance, high-speed network and well adapted code structure. Parallel models would be easily integrated in this framework, according to a low degree of parallelism in the matching climate versions.

Finally, the sensitivity to the precision in exchanged

data has been explored in coupled mode. In this study, we attempted to evaluate the growth of small perturbations introduced into the model by systematic imprecision of miscellaneous origins in such a distributed context.

We have shown that the perturbations of the flux induce rapid relaxations to perturbed AGCM or OGCM attractors with a doubling time $\tau = 20$ min for the atmosphere and $\tau = 14$ h for the upper ocean. In the atmosphere, the growth is so rapid that it reaches a first saturation stage before the first coupling. Therefore, additional small perturbations have no significant effect. Instead, the errors are dramatically amplified by the atmospheric model with a doubling time $\tau = 5$ days. The dynamics of the atmosphere seems to control the whole system at least on timescales up to a month.

As soon as the saturation is reached, the error seems to be highly sensitive to the seasonal cycle. Careful attention to spatiotemporal analyses between the two hemispheres leads to deeper insight about this type of sensitivity. The regions where the model variability is the largest are much more sensitive to the precision in exchanged data. This is the case at midlatitudes in the Northern Hemisphere where the oceanic currents exhibit a large variability due to their extension to the east. Here, we show that small initial perturbations have induced change in the magnitude of the SST, which can be compared to the model variability. Some important questions arise from this study, especially for medium/long time range forecast. But even if the simulation is very sensitive to small perturbations down to the change of the compiler, it is shown that the statistical properties of the simulation (i.e., the climate) are not changed.

Acknowledgments. The authors would like to thank Jacques Cahouet for his collaboration in the framework of the CATHODE Project that he initiated. We also thank the following people for useful support during the experiments: Nicole Girardot and Dominique Bielli (Météo-France) and Yves Dherbecourt (EDF) who made life across networks and security barriers easier. Finally, we wish to thank J. F. Royer and P. Delecluse for helpful discussions, J. Y. Caneill for his fruitful suggestions throughout the whole project, and M. Pontaud and E. Guilyardi for many tricks and assistance with the computations.

APPENDIX

The CALCIUM and the CLIM Software

a. The CALCIUM coupler

CALCIUM is a general coupling tool available for any type of code. It is based on a clear separation between two kinds of tasks: the development of a given code that must remain independent of the development of any other codes or any coupled applications, and the

coupling, which must remain completely separate from the codes that are considered black boxes by the coupler.

The design of CALCIUM is then based on the following characteristics (Beaucourt et al. 1994):

- *A basic component referred as “code connection points.”* By means of the CALCIUM facilities, each code can import or export values of variables on input or output ports named connection points that define the coupling interface of the code. It is worth noting that such ports allow anonymous communication: the code knows neither the origin nor the final destination of incoming/outcoming data. This is the key point of keeping a clear and real separation between codes. In order to facilitate the interface specification, symbolic names (temperature, pressure, etc.) and data types (real, integer, etc.) are given to CALCIUM connection points.
- *Coupling consists in linking “connection points” of different codes:* A CALCIUM application is therefore a set of several codes whose “connection points” have been joined. Thanks to the coupling language offered by CALCIUM library, defining connection links and distributing codes over available computers is a very simple task because the coupling is performed outside of the codes themselves. In CALCIUM, a link represents the definition of a single shared variable accessed by connected codes under different names. Only one code can write a shared variable. Furthermore, as each exported or imported variable is associated with a time value, CALCIUM is able to suspend a code asking for values which are not yet computed at the request date. It is able to detect infinite loops in the set of suspended codes that would lead to a deadlock.
- *A centralized architecture easy to implement:* Up to now, CALCIUM was seen as a library and a coupling language. Another important component is the coupling monitor. It is a set of processes in charge of managing all input and output data from the coupled codes. The communication between codes and monitor is based on message passing using PVM.

In the specific case of CATHODE, the topology of connected codes is represented on Fig. 4. ARPEGE and OASIS are implemented on the CRAY C98 of EDF, whereas CALCIUM and OPA are located on the CRAY II of Météo-France (Sevault et al. 1994).

b. The CLIM library

In order to integrate the pipe and distributed modes in a single OASIS source package, the Climate Modelling and Global Change team have carried out the design of a Coupling Library for Interfacing Models (CLIM) (Sevault and Terray 1995). Taking into account the outcomes of the CATHODE experience as well as some specifications of the Argonne National Labora-

tory's Fortran M (Foster 1992; Foster et al. 1992), the following points have been emphasized:

- *The message passing library:* The choice of a PVM-based library for a coupling tool appears to be the most flexible one. But the recommendations of the expert meeting held in Espoo (Kauranne 1994) about parallel programming models must be kept in mind, and further developments will use the message passing interface (MPI) as soon as some points are added or lightened in the MPI specifications (process spawning and location, information on the current physical configuration . . .).
- *Process status:* Programs should remain independent Unix processes without any hierarchical dependence. Therefore, no model is in charge of spawning the others. The launching of processes is up to the user.
- *Reliability:* The safety of data exchanges and an appropriate behavior to nonpredictable events in PVM (code aborts, host failure, network time out . . .) is of major importance for long-term simulations. Therefore, multiple checks are performed before any send or receive operation associated with an informative panels of error codes. Nonblocking PVM functions are preferred to blocking ones. The blocking receive is always used with a time-out control. A detailed trace file is generated for each model in the coupled application.
- *Flexibility:* The global configuration of the coupled application is hidden as much as possible from every process. A process is basically defined by its ports to the “outside” world. Ports are typed and they receive a symbolic name as well as a “port status” (*In* or *Out* or *InOut*). No assumption is made on the number of links (or Fortran M's *channels*) attached to a particular port. This should be fixed only at execution time. Several programs can connect to an identical port. We call this mechanism the dynamic link creation. A set of data can be imported or exported by a process with the same calling sequence, independently of its location, parallel decomposition or number of copies to be exported.
- *Integration in a parallel framework:* Parallel programs (using PVM) can safely use CLIM without any chance of interfering with intrinsic CLIM messages. Every message related to CLIM is identified by a unique (*process identifier, message tag*) pair, and there is no wildcard in receive functions. With the incoming of parallel versions of global circulation models, the coupling task becomes more problematic. At the time being, with multitasked GCM, the coupling is performed either in the sequential part of the temporal loop (ARPEGE-climat) or by the first multitasked task (OPA7). Assuming that the parallelism is expressed outside of the temporal loop, the coupled fields are no more available in their globality on one process. However, the OASIS coupler needs the complete fields to perform the interpolation

task. We consider some strategies with a particular focus on the atmospheric model since a parallel operational oriented version of IFS/ARPEGE is nearly available at the ECMWF (Barros et al. 1994). According to the “3D transposition strategy” each process is responsible in physical space for a set of contiguous grid points. In this case, CLIM offers an elegant solution through the multiple links–single port paradigm: all the parallel processes define an identical port, but with a different parallel data decomposition, while the OASIS coupler defines its matching port as usual. If we consider the so called “apple strategy” the data decomposition is described with the pair (*number of grid points, offset of the first point*). Further, CLIM also handles a “box strategy” where each process is responsible for a rectangular box as well as the more generic “orange strategy,” where each process is responsible for arbitrary pieces of the field.

- *Efficiency*: As CLIM is a PVM-based library, we expect to get roughly the same performance as PVM with the most suitable options for communication purposes. This is achieved by dynamically checking architecture and number of processes to allow TCP protocol and to avoid duplicate memory copies when packing as often as possible.

REFERENCES

- Arakawa, A., 1972: Design of the UCLA general circulation model. Numerical simulation of weather and climate. Tech. Rep. 7, Dept. of Meteorology, University of California, 116 pp.
- Barnett T. P., N. Graham, M. Cane, S. F. Zebiak, S. Dolan, J. O'Brien, and D. Legler, 1988: On the prediction of the El Niño of 1986/1987. *Sciences*, **241**, 192–196.
- Barros, S. R. M., D. Dent, I. Isaksen, and G. Robinson, 1994: The IFS model: Overview and parallel strategies. *Proc. Sixth ECMWF Workshop on the Use of Parallel Processors in Meteorology*, Reading, United Kingdom, ECMWF, 303–318.
- Beaucourt D., and C. Caremoli, 1994: Calium: A new tool for code coupling. *Proceedings Fall CUG 94*.
- Blanke, B., and P. Delecluse, 1993: Low-frequency variability of the tropical Atlantic ocean simulated by a general circulation model with mixed layer physics. *J. Phys. Oceanogr.*, **23**, 1363–1388.
- Bougeault, P., 1985: A simple parameterization of the large-scale effect of deep cumulus convection. *Mon. Wea. Rev.*, **113**, 2108–2121.
- Cane, M., and S. E. Zebiak, 1988: Dynamical forecasts of the 1986/87 ENSO with a coupled model. *Proc. 13th Annual Climate Diagnostics Workshop*, Washington, DC, NOAA, 278–282.
- , —, and S. C. Dolan, 1986: Experimental forecasts of El Niño. *Nature*, **321**, 827–832.
- Dandin, P., 1993: Variabilité basse fréquence simulée dans l’océan Pacifique Tropical. Thèse de doctorat de l’université Paris 6, 273 pp.
- Delecluse, P., G. Madec, M. Imbard, and C. Levy, 1993: OPA Version 7 Ocean General Circulation Model reference manual. LODYC Internal Rep. 93/05, 90 pp.
- Dent, D., L. Isaksen, S. Barros, F. Wollenweber, and G. Robinson, 1994: The message passing version of ECMWF’s weather forecast model. *High Performance Computing and Networking, April 1994, Lecture Notes in Computer Science*, Springer-Verlag, 299–318.
- Déqué, M., C. Drevet, A. Braun, and D. Cariolle, 1994: The ARPEGE/IFS atmosphere model: A contribution to the French community climate modelling. *Climate Dyn.*, **10**, 249–266.
- Foster, I., 1992: FORTRAN M as a language for building earth system models. *Proc. Fifth ECMWF Workshop on the Use of Parallel Processors in Meteorology*, Reading, United Kingdom, ECMWF, 144–155.
- , R. Olson, and S. Tueckle, 1992: Programming in Fortran M, version 2.0, August 30.
- Gates, W. L., 1992: AMIP: The Atmosphere Model Intercomparison Project. *Bull. Amer. Meteor. Soc.*, **73**, 1962–1970.
- Geleyn, J. F., 1987: Use of a modified Richardson number for parameterizing the effect of shallow convection. *J. Meteor. Soc. Japan, Special NWP Symp. Vol.*, 141–149.
- , and A. Hollingsworth, 1979: An economical analytic method for the computation of the interaction between scattering and line absorption of radiation. *Beitr. Phys. Atmos.*, **52**, 1–16.
- Goswami, B. N., and J. Shukla, 1991: Predictability of a coupled ocean–atmosphere model. *J. Climate*, **4**, 3–22.
- Guilyardi, E., and Coauthors, 1995: Simulation couplée océan-atmosphère de la variabilité du climat. *C. Roy. Acad. Sci. Paris*, **320**, 683–690.
- Guyon, M., M. Chartier, F. Roux, and P. Fraunie, 1994: A domain decomposition method applied to the baroclinic part of an ocean general circulation model on a MIMD machine. *Ocean Modelling* (unpublished manuscript), **103**, 7–20.
- Kauranne, T., 1994: Minutes of an expert meeting in Espoo. *European and Nordic Collaboration in Parallel Short Range Weather Forecasting*, Espoo, Finland.
- Lorenz, E. N., 1969: Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.*, **26**, 636–646.
- Louis, J. F., M. Tiedke, and J. F. Geleyn, 1982: A short history of the operational PBL-parameterization at ECMWF. *Proc. ECMWF Workshop Planetary Boundary Layer Parameterization*, Shinfield Park, Reading, United Kingdom, ECMWF, 59–80.
- Madec, G., and M. Imbard, 1995: A global ocean mesh to overcome the North Pole singularity. *Climate Dyn.*, **12**, 381–388.
- , M. Chartier, P. Delecluse, and M. Crépon, 1991: A three dimensional numerical study of deep water formation in the northwestern Mediterranean Sea. *J. Phys. Oceanogr.*, **21**, 1349–1371.
- Mahfouf, J. F., D. Cariolle, J. F. Royer, J. F. Geleyn, and B. Timbal, 1994: Response of the Météo-France climate model to changes in CO₂ and sea surface temperature. *Climate Dyn.*, **9**, 345–362.
- Mechoso, C. R., C. C. Ma, J. D. Farrara, J. A. Spahr, and R. W. Moore, 1993: Parallelization and distribution of a coupled atmosphere–ocean general circulation model. *Mon. Wea. Rev.*, **121**, 2062–2076.
- , and Coauthors, 1995: The seasonal cycle over the tropical Pacific in coupled ocean–atmosphere general circulation models. *Mon. Wea. Rev.*, **123**, 2825–2838.
- Meehl, G. A., 1995: Global coupled general circulation models. *Bull. Amer. Meteor. Soc.*, **76**, 951–957.
- Nicolis, C., 1992: Probabilistic aspects of error growth in atmospheric dynamics. *Quart. J. Roy. Meteor. Soc.*, **118**, 553–568.
- Pontaud, M., L. Terray, E. Guilyardi, E. Sevault, D. B. Stephenson, and O. Thual, 1995: Coupled ocean–atmosphere modelling: Computing and scientific aspects. *Proceedings of the Second UNAM-CRAY Supercomputing Conference*, Cambridge University Press, 16–23.
- Ritter, B., and J. F. Geleyn, 1992: A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. *Mon. Wea. Rev.*, **120**, 303–325.
- Royer, J. F., S. Planton, and M. Déqué, 1990: A sensitivity experiment for the removal of Arctic sea ice with the French spectral general circulation model. *Climate Dyn.*, **5**, 1–17.
- Sevault, E., and L. Terray, 1995: The Coupling Library Interfacing Coupling (CLIM): User’s guide and reference manuel. CERFACS Tech. Rep. TR/CMGC/95-47, 23 pp.
- , Noyret, P., L. Terray, and O. Thual, 1994: *Proc. VI ECMWF*

- Workshop on the use of Parallel Processors in Meteorology*, Reading, United Kingdom, ECMWF, 370–394.
- Stroe, R., and J. F. Royer, 1993: Comparison of different error growth formulas and predictability estimation in numerical extended-range forecasts. *Ann. Geophys.*, **V11**, 296–316.
- Stockdale, T., M. Latif, G. Burgers, and J. O. Wolff, 1994: Some sensitivities of coupled ocean–atmosphere GCM. *Tellus*, **46A**.
- Terray L., 1994: The OASIS Coupled User Guide Version 1.0, CER-FACS Tech. Rep. TR/CMGC/94-33, 123 pp.
- , O. Thual, S. Belamari, M. Déqué, P. Dandin, P. Delecluse, and C. Levy, 1995: Climatology and interannual variability simulated by the ARPEGE–OPA coupled model. *Climate Dyn.*, **11**, 487–505.
- Thompson, P. D., 1957: Uncertainty of initial state as a factor in the predictability of large scale atmospheric flow patterns. *Tellus*, **9**, 275–295.
- UNESCO, 1983: Algorithms for computation of fundamental property of sea water, UNESCO Tech. Paper Marine Science 44, 53 pp.