DOE/ER/61014--T2

THE ROLE OF CLOUDS AND OCEANS IN GLOBAL GREENHOUSE WARMING

Final Report

US Department of Energy Grant DE-FGO2-90ER61014

by

Martin I. Hoffert, Principal Investigator
Department of Physics, New York University,
New York, NY 10003

OCTOBER 1996

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

THE ROLE OF CLOUDS AND OCEANS IN GLOBAL GREENHOUSE WARMING

Final Report

by

Martin I. Hoffert, Principal Investigator
Earth Systems Group, Department of Physics, New York University, New York, NY 10003

OBJECTIVES AND HISTORY

Our research at NYU aims at assessing connections between anthropogenic greenhouse gas emissions and global climatic change. It has been supported since the early 1990s in part by the DoE "Quantitative Links" Program (QLP). A three-year effort was originally proposed to the QLP (later extended by two years) to investigate effects of global cloudiness on global climate and its implications for cloud feedback; and to continue the development and application of NYU climate/ocean models, with emphasis on coupled effects of greenhouse warming and feedbacks by clouds and oceans. It is well-known that cloud and ocean processes are major sources of uncertainty in our ability to predict climatic change from humankind's greenhouse gas and aerosol emissions. And it has always been our objective to develop timely and useful analytical tools for addressing real world policy issues stemming from anthropogenic climate change.

There has, of course, been increasing concern about the specter of global warming stemming from the buildup of atmospheric CO_2 from fossil fuel burning, and from other greenhouse gases emitted by human activities. Some climatic change has most likely already occurred. In their latest assessment report, the Intergovernmental Panel on Climate Change (IPCC) concluded that "the balance of evidence suggests a discernible human influence on climate (Houghton et al., 1996, p. 4)." The human imprint thus far is in the pattern of temperature change associated with the buildup of sulfate aerosols (mainly in the Northern Hemisphere) and greenhouse gases (worldwide) emitted by fossil fuel burning. If fossil fuel use expands as predicted for "business as usual" the next century could experience historically unprecedented global warming albeit with significant regional variations.

The IPCC was established by the World Meteorological Organization and the United Nations Environment Program in 1988 to provide information to the UN on global climate change -- a process in which many US scientists participate, including ourselves. The IPCC First Assessment Report (Houghton et al., 1990) was the basis of the UN Framework Convention on Climate Change. In 1992 the "Rio Treaty" was agreed to by members of the

United Nations at the Earth Summit in Rio de Janeiro. Article 2 of the Framework Convention calls for stabilization of atmospheric greenhouse gases (UNEP/WMO, 1992):

> The ultimate objective of this Convention...is...stabilization of greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system. Such a level should be achieved within a time-frame sufficient to allow ecosystems to adapt naturally to climate change, to insure that food production is not threatened and to enable economic development to proceed in a sustainable manner.

Note that this Convention does not define what "dangerous interference" means, what rates of climate change permit ecosystems to survive, how changes in precipitation and soil moisture affect agriculture, or how economic development is supposed to proceed sustainably when developed nations consume 10-100 times the per capita energy of poor countries. However, any rational implementation strategy requires that one understand the links between greenhouse gas and aerosol emissions on the one hand and regional climate change and economic impacts on the other.

The paradigm for computing anthropogenic climate change is the coupled atmosphere/ocean general circulation model (A/O GCM) (Washington and Parkinson, 1986). For reasons discussed below, the development of credible three-dimensional A/O GCMs has proven elusive. The reality is that atmospheric CO_2 buildup, global mean temperature changes, and sea level rise in the next century were computed for various global emissions scenarios by IPCC scientists using simplified ocean/climate and ocean/carbon cycle models (Houghton et al., 1990; 1996) many of which were developed by our group at NYU. This work led to the Rio Treaty, and is now being employed by negotiators to analyze treaty implementations that could stabilize atmospheric CO_2. It is gratifying to report that our research has had an impact on real world problems.

That schematic climate system models -- as opposed to A/O GCMs -- are being used for policy analysis has not escaped the attention of Rio Treaty negotiators, who have asked the IPCC to provide documentation of these models comprehensible to the educated layman. A series of technical reports (TP's) on this subject will be issued by the IPCC shortly. The documentation provided below summarizing results on this grant is more technical. What is described are the main findings in series of interrelated tasks at NYU exploring "quantitative links" between anthropogenic greenhouse and aerosol emissions and climatic change. Reprints of publications supported partly or wholly by this grant are also appended to this Final Report.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

ACCOMPLISHMENTS AND MAJOR FINDINGS

A requirement for modelling the atmospheric buildup of fossil fuel CO_2 and anthropogenic global climate change is accounting for the role of oceans in storing carbon and heat. Since we proposed a one-dimensional upwelling-diffusion ocean (UD) coupled to a Polar Sea for transient climate change studies (Hoffert et al., 1980) many other researchers have applied, and in some cases improved on, our formulation; Tom Wigley of the National Center for Atmospheric Research (NCAR), Robert Watts of Tulane University, Mike Schlesinger of the University of Illinois at Urbana-Champaign (U of I/UC), Danny Harvey of the University of Toronto, and others.

The ocean carbon cycle UD ocean published a year later (Hoffert et al., 1981) has likewise been articulated by Atul Jain of U of I/UC, Haroon Kheshgi of Exxon Corporate Research Laboratories, and by others. UD models are now employed in a variety of modes including direct radiative forcing by changes in solar luminosity, aerosols and greenhouse gases, and for studies of climate variability. Normally, the carbon cycle version is forced by CO_2 (or carbon) emissions and, when coupled to a terrestrial biosphere, computes the evolution of atmospheric CO_2. These models have been employed to study the so-called "missing sink" needed to balance the present-day carbon cycle -- perhaps CO_2 fertilization of boreal forests. More recently, we have run the carbon cycle in "inverse mode" with 21st century concentration paths from IPCC CO_2 stabilization scenarios as inputs, and the carbon emission paths that produce those concentration paths as outputs (Jain et al., 1995).

UD Ocean/Climate Models

A schematic diagram of the UD ocean is shown in Figure 1. As mentioned above, there are several versions of the model, including some in which the oceans couple to a radiatively very sophisticated atmospheric model. But for illustrative purposes assume that the climate change is driven by some specified linear combination of solar, albedo and greenhouse gas radiative forcing, $\Delta Q\,(t) = \Delta Q_{sun} + \Delta Q_{albedo} + \Delta Q_{green}$. As usual, radiative forcing is the horizontally-averaged change in the radiative balance at the tropopause after the stratosphere has adjusted but before the troposphere has (see Houghton et al., 1990, for details).

For example, radiative forcing from changes in CO_2 concentration is parameterized by the IPCC as $\Delta Q_{green}\,(CO_2) = 6.3\ln[(CO_2/(CO_2)_o]$, where subscript zero denotes the preindustrial state (Houghton et al., 1990); so the forcing for doubling atmospheric carbon dioxide is $\Delta Q_{2x} = 6.3\ln 2 = 4.4$ W m^{-2}, a useful reference. *Climate sensitivity* is proportional to the ratio of the equilibrium global mean surface temperature change, ΔT, to the radiative forcing. Typically, it is measured by the *CO_2 doubling sensitivity*, ΔT_{2x} =

$\Delta Q_{2x}(\Delta T/\Delta Q)$. A related quantity is the *radiative damping coefficient*, $\lambda = (\Delta Q)/\Delta T) = (\Delta Q_{2x}/\Delta T_{2x})$, which is a net radiative cooling flux for the planet as a whole per unit surface temperature change.

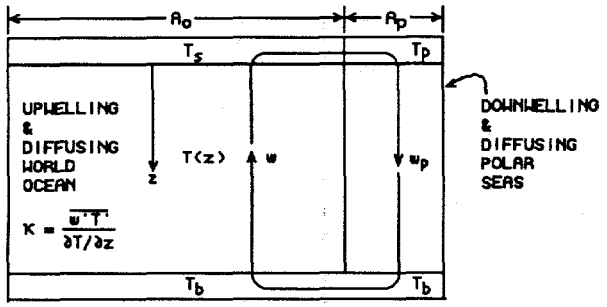


Figure 1. Schematic diagram of upwelling-diffusing world ocean coupled to high-latitude polar sea. Cold water at temperature Tp enters the base of the world ocean water column and upwells to the surface at velocity w ~ 4 m/y. At the same time warm water from the surface mixes downward by some vertical eddy diffusivity κ ~ 2000 m^2/yr. The balance between upwelling and diffusive fluxes results in an exponential steady-state temperature profile in deep ocean of the form $T(z) = T_p + (T_s - T_p)\exp(-z/z^*)$, where $z^* = \kappa/w$ ~ 500 m is a thermocline depth scale. Above the "deep ocean" is a well mixed layer of depth h in contact with the atmosphere.

The evolving global mean surface temperature, $T_s(t) = T_{so} + \Delta T$, is given in the UD ocean/climate formulation by the differential equation of a thermal uniform mixed layer near the surface (Hoffert and Flannery, 1984),

$$\frac{dT_s}{dt} = \left\{\frac{\Delta Q - \lambda(T_s - T_{so})}{\rho c_p h}\right\} + \left\{\kappa\frac{\partial T}{\partial z} + w(T - T_p)\right\}_{z=h} . \qquad [1]$$

where ρ (~ 1030 kg m^{-3}) and c_p (~ 4000 J kg^{-1} K^{-1}) are the density and specific heat of seawater, w (~ 4 m y^{-1}) is upwelling, κ (~2000 m^2 y^{-1}) vertical eddy diffusivity, and h (~ 70 m) an effective mixed layer depth, area-weighted for land and ocean fractions. The first bracketed term is the warming rate of the mixed layer from radiative forcing less cooling to space; the second bracketed term the warming rate from heat exchanges between the mixed layer and the upwelling-diffusion ocean of depth h_b (~ 4000 m). In the steady state, the bracketed terms go to zero individually.

To get the mixed-layer/thermocline exchange, we need to solve an upwelling-diffusion partial differential equation for potential temperature T(z,t) in an upwelling-diffusing ocean,

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial z}\left(\kappa\frac{\partial T}{\partial z}\right) + w\frac{\partial T}{\partial z} , \qquad [2]$$

subject to continuity boundary conditions at the mixed layer/thermocline interface, and conservation of heat flow between polar sea downwelling and world ocean upwelling at the sea floor:

$$@z = h: T = T_s(t); \qquad @\, z = h_b: \quad \kappa\left(\frac{\partial T}{\partial z}\right) + wT = wT_p(t). \qquad [2a]$$

To close this system we introduced the Polar Warming Parameter early on, $\pi = (T_p - T_{po})/(T_s - T_{so})$ -- a parameter subsequently employed by the IPCC and others for temperature and sea level rise scenario analysis (Wigley and Raper, 1992; Titus and Narayanan, 1996). For constant κ and w, the equations are linear, and soluble analytically in principle. In practice, for general radiative forcing histories $\Delta Q(t)$, it is more efficient to employ a numerical finite-difference approach. In our work at NYU we typically represent the four kilometer deep world ocean by forty hundred-meter-thick slabs, and use implicit Crank-Nicholson integration to find their evolving temperatures.

Many features of A/O GCM numerical model results were first predicted by our NYU UD ocean/climate models. And since the oceans are resolved in only one vertical (vertical) dimension (and the atmosphere in zero dimensions), calculations can easily be done on a personal computer -- as opposed to coupled three-dimensional A/O GCMs requiring huge chunks of supercomputer time.

Moreover, it has proven possible to simulate the global mean temperature evolution of A/O GCMs with UD/climate models for appropriate choices of ΔT_{2x}. The most recent IPCC report includes a comparison of global warming rates predicted by 10 A/O GCMs with CO_2 doubling sensitivities between 2.1 and 4.6 °C forced by a 1% per year increase in CO_2 (Kattenberg et al., 1996; Fig. 6.4, p. 300). This forcing roughly mimics the historical buildup of atmospheric greenhouse gases. Of course, these A/O GCMs, in addition to their 3D resolution, contain an enormous amount of explicit physics. Still, all of the GCM results straddle the ~ 0.3 °C/decade warming predicted by an UD ocean/climate model with a sensitivity of $\Delta T_{2x} = 2.5$ °C; and those GCMs with the highest equilibrium sensitivity show the fastest warming rates; those with the lowest sensitivity, the lowest rates; a result predicted by our modelling studies well before this intercomparison (see below).

UD Ocean/ Carbon Cycle Models

Predicting atmospheric CO_2 changes from carbon emissions requires a carbon cycle model that realistically represents penetration and mixing of carbon in the oceans. It is well-known that of the fossil fuel carbon emitted thus far, roughly half has remained in the atmosphere, with the balance going into the oceans. It is also understood that the rate of carbon uptake by the sea is not so much controlled by gas exchange at the surface as by mixing in oceanic interiors.

A consequence of aqueous chemistry that when carbon dioxide enters seawater only a small percentage remains as dissolved CO_2 gas. The rest of the carbon is "buffered" to bicarbonate (HCO_3^-) and carbonate ions (CO_3^-).

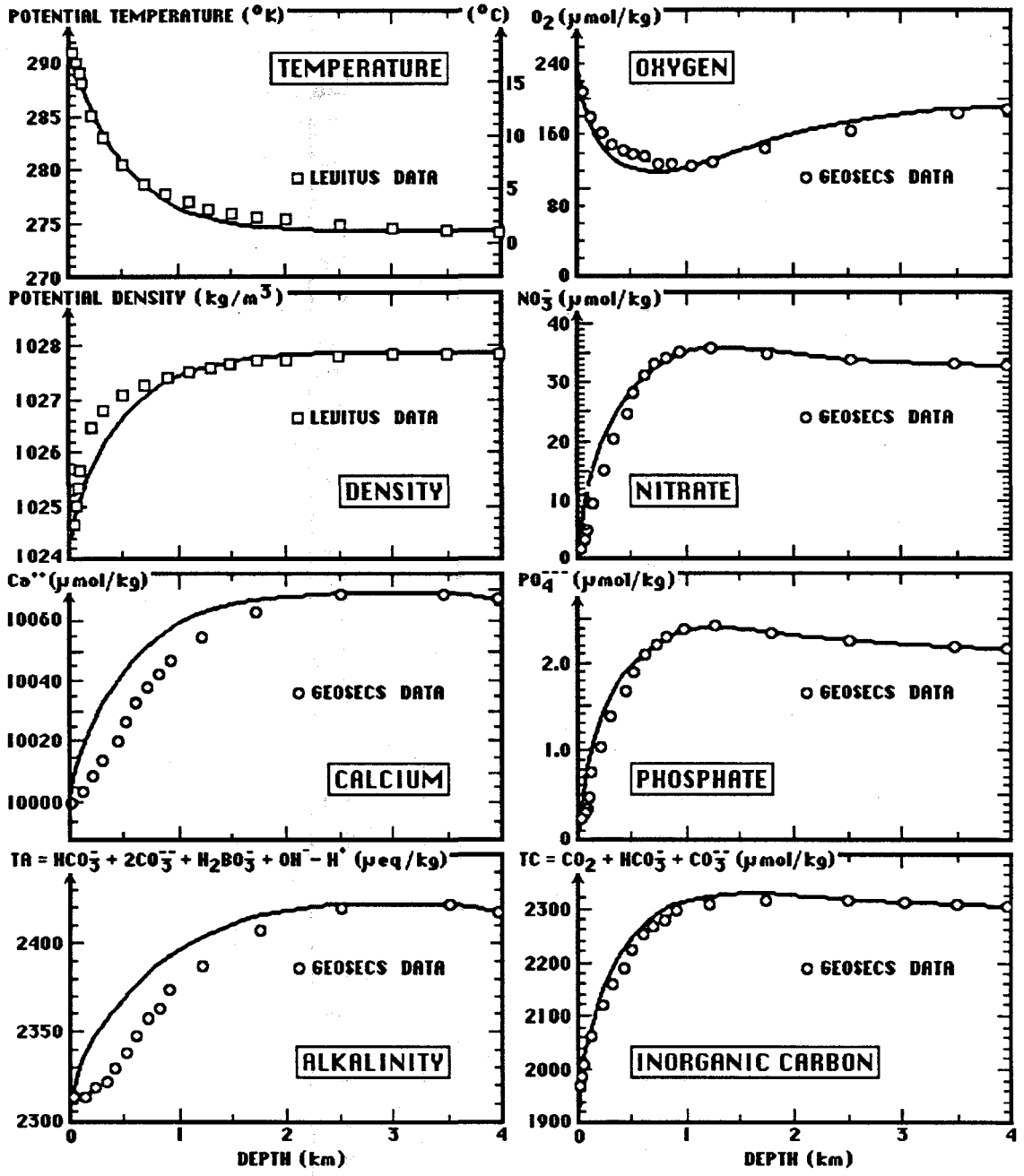


Figure 2. Observations of oceanic temperature, nutrient and carbon cycle profiles horizontally-averaged over the world's oceans compared with steady-state solutions of NYU one-dimensional upwelling-diffusion ocean model (T. Volk, Ph.D. thesis, 1984).

The present value of the surface ocean "buffer factor" is

$$\zeta = \frac{\Delta PCO_2/(PCO_2)_o}{\Delta TC/(TC)_o} \sim 10. \qquad [3]$$

That is, the effect of Henry's law and aqueous chemistry is that a change in atmosphere carbon dioxide concentration (or partial pressure, PCO_2) from fossil fuel burning increases total inorganic carbon concentration ($TC = CO_2 + HCO_3^- + CO_3^{-2}$) of the surface ocean by only ~10% of the relative increase in PCO_2. Consequently, changes in total carbon concentration from fossil fuel carbon invading the sea are small and hard to measure directly. The ~ 30% increase in PCO_2 from anthropogenic emissions since preindustrial produced an only ~ 3% increase in TC. This is much less than natural local variability, even if we had baseline data from which to measure changes relative to preindustrial -- which we don't.

Since direct observations of the oceanic uptake of fossil fuel are so difficult transient tracers like bomb radiocarbon are often used to calibrate carbon cycle models. Another test of ocean carbon cycle models pioneered at NYU is recovery of observed steady state tracer profiles. It has not been easy to get all the steady and transient tracers right in the same carbon cycle model. Figure 2 shows our UD/carbon cycle model recovers key features of carbon cycle related steady tracer concentration profiles in the World Ocean rather well (Volk, 1984) -- better in fact than present-day ocean carbon cycle models based on circulation fields from 3D ocean models.

Notice in Figure 2 that the total inorganic carbon (TC) profile exhibits a minimum near the surface where photosynthetic algae convert dissolved carbon dioxide to organic carbon (CH_2O): $h\nu + CO_2 + H_2O \rightarrow CH_2O + O_2$ -- a process also limited by diffusion to the surface of nitrate (NO_3^-) and phosphate (PO_4^{-3}) nutrients. When plankton die or excrete, organic carbon detritus sinks through the thermocline, oxidizing on the way down back to CO_2 , and releasing nutrients to the water column for recycling in the process. This generates peaks in TC, phosphate and nitrate, and a minimum in dissolved oxygen, at ~ 1 kilometer depth; features which recovered by our UD carbon cycle with appropriate sources and sinks (Figure 2).

Along with the organic carbon rain are sinking calcium carbonate shells of planktonic foraminifera. Although the upper ocean is supersaturated with respect to both calcite and aragonite (the mineral forms of $CaCO_3$) these only precipitate as hard parts of certain marine organisms. If the carbonate pellets fall to the sea floor below their so-called saturation horizon they re-dissolve in the water column. But particulate carbonate falling on sufficiently shallow parts of the sea floor accumulates as sediment. Although $CaCO_3$ is chemically "inorganic carbon" the precipitation of carbonates in water, into what ends up as sedimentary rock formations, has always involved biology. The oldest sedimentary rock, a 3.8 By BP water-deposited chert/carbonate of the Isua Formation of West Greenland, contains microbial fossils (Schopf, 1983). There

is some evidence now that meteorite ALH84001 of the Shergotty-Makhla-Chassigny (SNG) class -- impact-blasted from the surface of Mars, wandering through space 16 My, and eventually crashing into the Antarctic ice cap 13 kyr BP, where it was recently discovered -- also contains microbial fossils (McKay et al., 1996). If so, the simultaneous presence of carbonate globules deposited ~ 3.6 By BP on Mars during a prior wet climatic epoch suggests the Earth's carbon cycle is not the only one with biologically mediated carbonate precipitation.

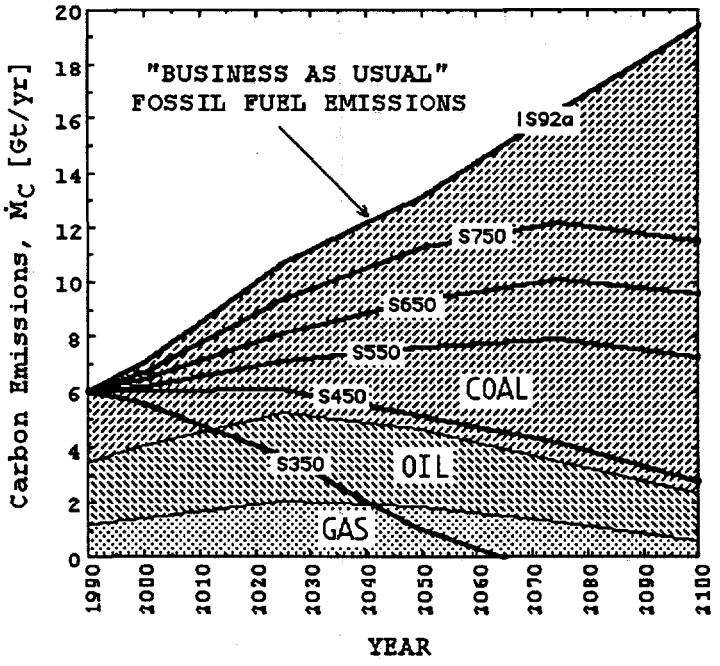


Figure 3. Fossil fuel carbon emissions in the 21st century. Total emissions for the IPCC "Business as Usual" scenario (IS92a) are the uppermost solid curve; with cumulative emissions from gas, oil and coal indicated as differently shaded areas. The other solid curves are carbon emissions constrained to approach atmospheric CO_2 concentrations of 750, 650, 550, 450 and 350 ppm according to the IPCC stabilization paths (Houghton et al., 1996, Fig. 7a). The calculations were done with an upwelling-diffusion ocean/carbon cycle model (Hoffert et al., 1981, Volk, 1984, Kheshgi et al., 1991) run in an inverse mode (Jain et al., 1995). Applying appropriate carbon emission factors for gas, oil and coal to these results we found that more carbon-free energy supply would be needed fifty years hence to stabilize at 550 ppm (~ a CO_2 doubling) than currently consumed by our global fossil-fuel based civilization (Hoffert et al. , in preparation).

In any case, our UD ocean/carbon cycle model provides important insights into the working of carbon "pumps" driven by circulation, carbonate chemistry and biology (Volk and Hoffert, 1985). Among other things the model may help explain changes in atmospheric CO_2 recorded in polar ice cores during glacial-interglacial transitions, and has been employed to predict future atmospheric carbon dioxide buildup from fossil fuel burning. The version of the UD/carbon cycle model that we use now was developed collaboratively with Atul Jain and Don Wuebbels at the University of Illinois at Champaign/Urbana and Haroon Kheshgi at Exxon Corporate Research Laboratory. It has been tested against historical bomb radiocarbon and

atmospheric CO_2 data, and is documented in Jain et al. (1995) appended to this report (including all equations). We are using the carbon cycle model now to assess implications of IPCC CO_2 stabilization scenarios on the demand for renewable energy (and other carbon-free energy sources) in the twenty-first century. A preliminary result is shown in Figure 3.

It turns out to be important for transient climate as well as transient carbon cycle calculations to have a base steady state that agrees with observations. Many problems of 3D GCMs arise because they don't recover the "present climate" or the "present oceanic state." This is believed to result mainly from incorrect explicit cloud processes (see below). Because computed climatic states of A/O GCMs tend to drift away from observations for simulations of the present climate, ad hoc "flux adjustments" of heat and fresh water fluxes are often employed to drive the base state toward reality.

It is often rationalized that climatic perturbations are relatively unaffected by flux-adjusting the base state. However, this is an unverified assumption. It would not, for example, be the case if the climate system is strongly nonlinear. It is precisely the possibility of capturing nonlinear "surprises" -- like the cutoff of North Atlantic Deep Water -- that is cited as an advantage of GCMs. But our point is simply that UD models get the base steady state very nearly right by design. And so flux-adjustments are not needed in transient calculations. We are not arguing that UD models are superior to GCMs; simply observing an important property that they have.

Climate Sensitivity

A early finding of UD ocean/climate models was that ΔT_{2x} is the single most important parameter determining the global mean transient climate response for a given radiative forcing (Hoffert and Flannery, 1984). This is evident in Figure 4, which appeared in the original IPCC report (Houghton et al., 1990). Note that the "best estimate" climate sensitivity range ($1.5 < \Delta T_{2x} < 4.5$ ^{o}C) and the conclusion that climate sensitivity is the most important parameter in transient climate response survives unchanged from the IPCC First Assessment (Houghton et al., 1990) to their recent Second Assessment five years later (Houghton et al., 1996). What has been added in IPCC 1995 is consideration of the role of anthropogenic aerosols and solar variability to radiative forcing, which improves the agreement between UD ocean/climate model predictions and historical temperatures variations over the past hundred years (Houghton et al, 1996, Fig. 8.4, p. 424).

Still, the persistent factor of three uncertainty can mean the difference between the need for a vigorous CO_2 stabilization strategy and a *laissez faire* approach to "prevent dangerous interference with the climate system." The main reason for the uncertainty is our inability to determine which of the many possible cloud parameterizations employed in GCMs is correct. All atmospheric GCMs are "tuned" to more or less recover present climate when

sea surface temperatures (SSTs) or oceanic poleward heat flows are specified. However, they can produce significantly different global warming when perturbed by same radiative forcing. To better understand the cloud feedback problem we developed the conceptual model summarized below. It has proven useful as a tool for analysis of both GCM and empirical results.

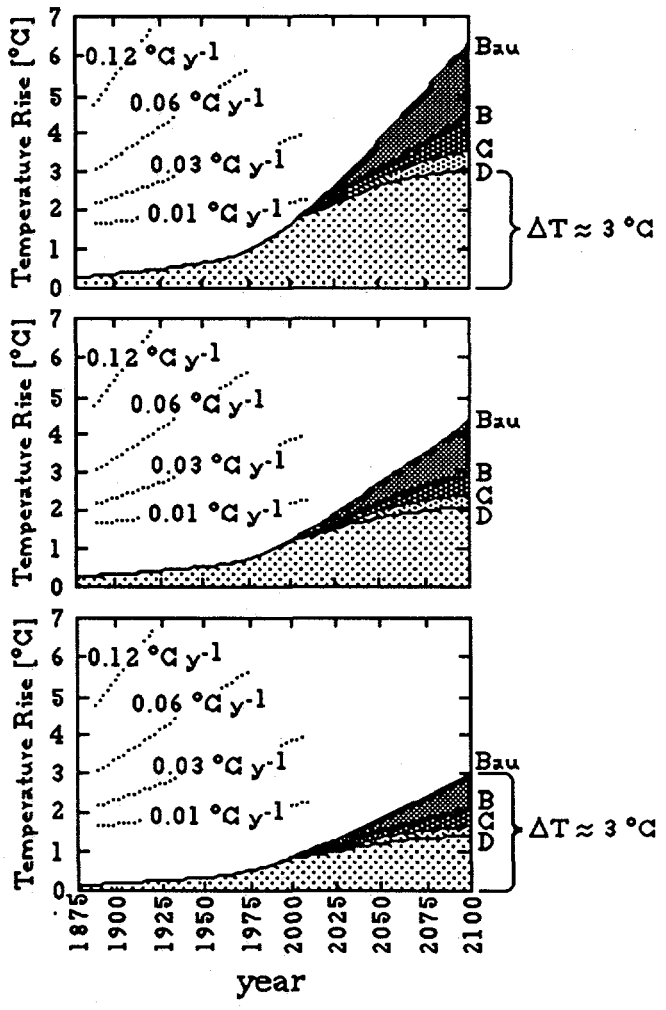


Figure 4. Twelve predictions of global mean surface temperature response projected to the year 2100 computed with the NYU ocean/climate model. Each panel shows the response to the same four IPCC 1990 radiative forcing scenarios: Bau (Business as Usual), B, C and D.

The IPCC climate sensitivity range is $1.5 < \Delta T_{2X} < 4.5$ °C. Panels differ in the climate sensitivity assumed in the calculations. The top panel is for $\Delta T_{2X} = 4.5$ °C, the middle for $\Delta T_{2X} = 2.5$ °C and the bottom for $\Delta T_{2X} = 1.5$ °C. By the year 2100 global warming is ~ 3 °C for the most constrained (and expensive) scenario if climate sensitivity is near the high end of the range; and also ~ 3 °C for Bau if sensitivity is near the low end.

The policy implications of global warming thus depend critically on climate sensitivity.

It follows from the global energy balance that the effective temperature at which the Earth radiates to space is $T_{eff} = [S_o(1 - \alpha_p)/(4\sigma)]^{1/4} \approx 255$ K (-18 °C) where S_o (≈ 1370 W m^{-2}) is the solar constant, α_p (≈ 0.30) the planetary albedo and σ (≈ 5.67 x 10^{-8} W m^{-2} K^{-4}) the Stefan-Boltzmann constant. That T_{eff} is 33 °C above the global mean surface temperature, $T_s \approx 288$ K (15 °C), is a consequence of atmospheric greenhouse gases. The entire troposphere -- roughly, the first 11 kilometers of atmosphere where temperature drops off with altitude at a lapse rate $\gamma \approx 6$ °C/km -- is "convectively coupled" to the surface: That is, the lapse rate remains approximately constant as the Earth warms or cools. Right now, the atmosphere radiates to space from an effective

altitude $z_{eff} \approx (T_s - T_{eff})/\gamma \approx (33\ ^oC/6\ ^oC\ km^{-1}) \approx 5.5$ km. Outside a narrow wavelength "window," the lower troposphere is made opaque in the infrared by greenhouse gases. An orbiting observer with infrared sensors would detect an effective temperature from the mid-troposphere, on average.

As the surface warms by ΔT, the entire convectively coupled troposphere warms by about the same amount, including at the altitude radiating to space where the temperature becomes $T_{eff} + \Delta T$. This "feedback" increases cooling to space until it balances increased heating from radiative forcing. The blackbody cooling rate per unit surface warming is $\lambda_b = [\sigma(T_{eff} + \Delta T)^4 - \sigma T_{eff}^4]/\Delta T \approx 4\sigma T_{eff}^3 \approx 3.8\ W\ m^{-2}\ K^{-1}$. It is this basic blackbody cooling feedback that prevents surface temperature from running away when the Earth is heated. However, blackbody cooling can be affected either positively or negatively by other climate feedbacks; for example, from water vapor, snow, ice and clouds.

At equilibrium, the blackbody cooling flux,

$$\Delta Q_{cool} = \lambda_b \Delta T,$$

balances the sum of direct radiative forcing plus and heating induced by clear sky and cloud feedbacks,

$$\Delta Q_{heat} = \Delta Q + \left[\left\{\frac{\Delta Q}{\Delta T}\right\}_{clear} + \left\{\frac{\Delta Q}{\Delta T}\right\}_{cloud} \right] \Delta T.$$

Clear sky feedbacks, $\{\Delta Q/\Delta T\}_{clear}$, include positive water vapor and negative snow and ice albedo feedback. Cloud feedbacks, $\{\Delta Q/\Delta T\}_{cloud}$, include the positive cloud greenhouse (mainly longwave) and negative cloud albedo (mainly shortwave) feedback. Setting $\Delta Q_{heat} = \Delta Q_{cool}$ yields

$$\frac{\Delta Q}{\Delta T} \equiv \lambda = \lambda_{clear} - \left\{\frac{\Delta Q}{\Delta T}\right\}_{cloud}. \qquad [4]$$

where $\lambda_{clear} \equiv \lambda_b - \{\Delta Q/\Delta T\}_{clear}$ is the radiative damping for clear sky regions. The solid line in Figure 5 is equation [4] with $\lambda_{clear} \approx 2.2\ W\ m^{-2}\ K^{-1}$.

We estimated λ_{clear} from satellite data as follows. Raval and Ramanathan (1989) express clear-sky infrared flux to space versus surface temperature as $I(T_s) = \sigma T_s^4 [1 - g(T_s)]$, where $g(T_s) = g_0 + g_1 T_s$ is a (mainly water vapor) greenhouse function. Their regression coefficients based on Earth Radiation Budget Experiment (ERBE) satellite data and sea surface temperature (SST) data sets are $g_0 = -0.658$ and $g_1 = 3.42 \times 10^{-3}\ K^{-1}$. Neglecting the weak temperature dependence of solar absorption gives the clear sky radiative damping coefficient $\lambda_{clear} = \partial I/\partial T_s = 4\sigma T_s^3 [1 - g_0 - (5/4) g_1 T_s]$. From

this expression, we computed the clear sky radiative damping and the CO_2 doubling climate sensitivity from the equator to the pole for the present zonal mean surface temperature distribution. This data yielded λ_{clear} in the range 2.1 to 2.3 W m^{-2} K^{-1} and $(\Delta T_{2x})_{clear}$ in the range 1.9 to 2.1 K.

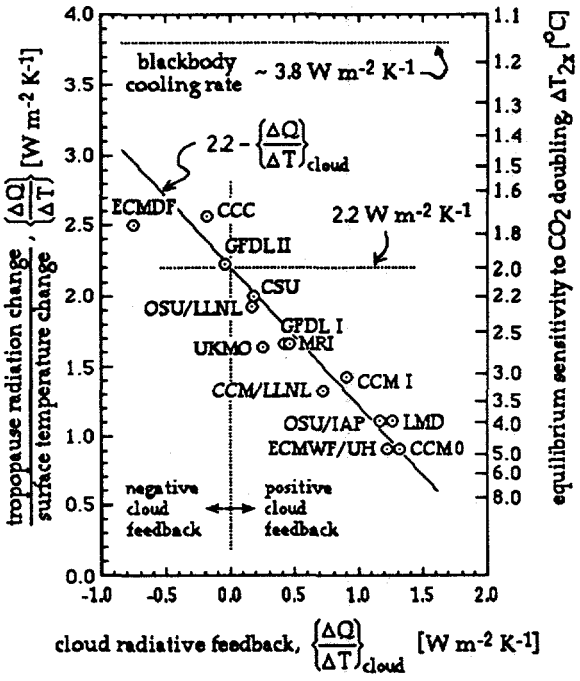


Figure 5. Radiative damping ($\Delta Q/\Delta T$, left vertical scale) and CO_2 doubling sensitivity (ΔT_{2x}, right vertical scale) versus the cloud radiative feedback parameter of 14 A GCMs. The GCM points ⊙ are derived from data in Cess et al. (1989, Table 2) for SSTs varied at all ocean latitudes and longitudes by ± 2 °C. This comparison is instructive even though temperature changes are unlikely to be uniformly distributed over latitude. That the GCM points line up so well with the correlation line indicates that most of the variance is due to differences in the way the different models parameterize cloud process.

What Figure 5 shows is that most of the differences in the equilibrium climate sensitivity of GCMs,

$$\Delta T_{2x} = \frac{\Delta Q_{2x}}{\lambda_{clear} - \left\{\frac{\Delta Q}{\Delta T}\right\}_{cloud}}, \qquad [5]$$

arise from differences in their net cloud feedback, $\{\Delta Q/\Delta T\}_{cloud}$. At this point, even the sign of cloud feedback is unknown. And cloud feedback uncertainties are unlikely to be narrowed unless observation-based techniques are discovered that credibly distinguish which parameterizations are more realistic.

Unfortunately, recent ingenious attempts to employ ERBE satellite radiation and SSTs observations to deduce cloud feedbacks in tropical zones by comparing El Niño with non El Niño years (Ramanathan and Collins, 1991) have been unsuccessful (Fu et al., 1992). The Ramanathan-Collins tropical thermostat regulated by highly-reflective cumulus towers decreasing solar absorption as the surface warms is an intriguing idea. But their method of proving it using local satellite-derived radiative correlations has not stood

up to scrutiny. We have also been down that path. The problem is separating local correlations for $\{\Delta Q/\Delta T\}_{cloud}$ from cloud-forming influences in the atmosphere and oceans remote from the zone in question.

The relationship between cloud and ocean processes in climate change is quite important, and just beginning to be appreciated. Gleckler et al. (1995) compared the implied ocean heat transport of fifteen atmospheric GCMs constrained by realistically varying SSTs as part of the Atmospheric Model Intercomparison Project (AMIP). The implied poleward heat flow by the oceans needed to satisfy global energy balance was critically sensitive to the radiative effect of clouds, to the extent that even the sign of the Southern Hemisphere ocean heat transport was affected by errors in simulated cloud-radiative interactions. Differences in cloud parameterizations affect the surface oceanic fluxes of solar and infrared radiation, and hence net heat fluxes into and out of the sea which control the poleward heat flow by the oceans when SSTs are specified. Cloud-ocean interactions work both ways. It is known from observational studies of the tropical warm pool during El Niños that cloud formation depends on oceanic dynamics.

As with global climate sensitivity, our lack of understanding about how to parameterize clouds is the culprit. Lurking behind this finding is the possibility, which cannot logically be excluded, that it may not be possible to accurately parameterize the climatic effect of thousands of clouds with complex microphysics inside GCM finite-difference grids hundreds of kilometers on a side.

In the long run, it is vital to continuously monitor the Earth System. Inexpensive satellite measurements of radiative forcings and feedbacks of the type proposed by Hansen et al. (1995) at the Goddard Institute for Space Studies (GISS) are particularly attractive. We are, in a related effort, working with NASA/GISS scientists on such an experiment, perhaps to be conducted under the NASA Earth System Science Pathfinder Program. Even if such monitoring is successful -- a major technological challenge in itself -- it is unlikely to resolve the cloud feedback issue decisively before the climate itself changes.

We therefore asked ourselves the following question: Is there a productive research path to narrowing global climate sensitivity uncertainties in the meanwhile?

Paleocalibration

The answer we arrived at was *paleocalibration* -- the derivation of global climate sensitivity from reconstructions of the radiative forcing and temperature responses of ancient climates (Hoffert and Covey, 1992; Covey et al., 1996). Our paleocalibration work at NYU was facilitated by collaborations with Curt Covey at Lawrence Livermore National Laboratory LLNL), Lisa Sloan at the University of California at Santa Cruz, Haroon Kheshgi at Exxon

Corporate Research Laboratories, and particularly Andrei Lapenis, who joined us at NYU after working closely for many years with Michael Budyko at the State Hydrological Institute in Leningrad (now, St. Petersburg) in the Former Soviet Union (FSU).

Unlike climate research in the West where future climate states are forecast primarily using numerical computer models, the approach developed by Budyko in the FSU was to develop global data sets from paleoreconstructions, and to project future changes assuming that the pattern of changes normalized to hemispheric mean temperature change remains constant -- a technique dubbed the *paleoanalog* method.

The Russian paleoanalog path may have been partly a consequence of insufficient computational resources for nonmilitary research in the FSU. But a paleodata-based approach does have the conceptual advantage of being observationally based. In practice, there are questions of the coverage and accuracy of reconstructions. How good is the Russian data? And objections can be made that patterns of temperature and precipitation change depend on the spatial patterns of the forcing (not only on the global mean); for example, in the highly nonuniform sulfate aerosol forcing from humankind's sulfur emissions.

On the other hand, paleocalibration does not make explicit assumptions about the constancy of spatial patterns, although some pattern assumptions may be implicit in reconstructing the global temperature change from data at isolated locations. And we have tried to use all available data from researchers worldwide. Dr. Lapenis did make available to our group at NYU hitherto inaccessible Russian paleodata for many geologic epochs which greatly facilitated our studies. Much of this data was eventually published in English under this grant (Kheshgi and Lapenis, 1996). In addition to temperatures, it is necessary to reconstruct radiative forcing from changes in solar output, albedo and greenhouse gases to apply the method. In our research we attempted to address issues of uncertainty in the paleodata and to provide error estimates.

Our major finding is based on analysis of two very different climates -- the Last Glacial Maximum (LGM, a cold period 21.5 kyr BP) and the Mid-Cretaceous Maximum (MCM, a warm period 100 Myr BP). By reconstructing the forcing and response of these climates we determined a CO_2 doubling sensitivity of $\Delta T_{2x} = 2.3 \pm 0.9$ °C (Hoffert and Covey, 1992). Not only is this comfortably within the IPCC range ($1.5 < \Delta T_{2x} < 4.5$ °C), but our central value is very close to the clear sky value of $\Delta T_{2x} \sim 2.2$ °C. The clear sky sensitivity is also the cloud-neutral value at which cloud greenhouse and cloud albedo feedbacks cancel. Interestingly, the most recent IPCC report cites a recent intercomparison of GCMs in which net cloud feedbacks are much closer to neutral than previously (Houghton et al., 1996, Fig 4,2, p. 206). But some caution may be appropriate. Because prior discrepancies in GCM cloud

feedback are widely known, this finding may be more the result of psychological "convergence to the mean" than of real cloud physics. Still, it does support our results.

An important milestone of our work was the Workshop on Paleocalibration of Climate Sensitivity held at the NOAA Metro Center, Silver Springs, Maryland, 15-17 August, 1994 (Broccoli, 1994). The Workshop was organized by the PI and Andrei Lapenis of NYU, Curt Covey of LLNL and Lisa Sloan of UC Santa Cruz, to assess the potential of paleoclimate reconstructions for reducing the uncertainty of climate sensitivity. Some forty scientists specializing in climate modelling and paleoclimatology participated. While recognizing the importance of paleocalibration as an approach to climate sensitivity, there were some reservation raised regarding the accuracy of the Russian paleodata -- reservations which we believe have been addressed since the meeting by publications in the peer-reviewed literature.

Among other things, our Paleocalibration Workshop was enlivened by the presence of Prof. Richard Lindzen of MIT -- outspoken critic of present-day GCMs who argues that global climate sensitivity is as low as $\Delta T_{2x} \approx 0.2$ °C because of negative water vapor feedback associated with upper troposphere drying by tropical cumulus cloud towers (Lindzen, 1995). In that case, radiative forcing would had to have been nonphysically huge to have driven paleotemperature changes observed in the geologic record; negatively huge in the LGM to cool the planet 3-5 °C, positively huge in the MCM to have warmed it by ~ 9 °C (Hoffert and Covey, 1992). Lindzen responds by discounting the radiative forcing paradigm and questioning paleotemperature reconstructions.

The Workshop led to ongoing analyses of the empirical calibration of climate sensitivity. A residual point which needs explanation is the apparent stability of tropical temperatures over geologic time, particularly during past warm periods. Tropical temperature stability has been studied as part of our DoE climate research at NYU and is the subject of Dan Collins Ph.D. thesis (Collins, 1996).

It is not necessary to accept the "paleoanalog" technique as proposed by Budyko to recognize the importance of paleodata for assessing future climate change. A specific recommendation of the Paleocalibration Workshop, which we reiterate here, is for a global effort to document and archive paleodata on temperature and past greenhouse gases to test climate models -- particularly data from prior warm epochs. More to the point would be a focused paleodata acquisition program targeted at climate sensitivity uncertainties, including new geological observations where necessary (cores, pollen, paleosols, etc.). Despite the obvious importance and potentially high payoff of such a program no funding agency has stepped forward to address the challenge as yet.

Ocean Mixing

We also studied the sensitivity of climate change predictions to parameterizations of vertical mixing. To test the impact of a density-dependent vertical eddy diffusivity law on global warming predictions we did calculations with a standard UD ocean/climate model driven by the IPCC historical anthropogenic greenhouse radiative forcing from 1860 to 1990 for a vertically uniform eddy diffusivity κ_0 and a $\kappa = a\,N^{-1}$ eddy diffusivity law similar in functional form to that predicted for ocean produced by internal wave-breaking (Gargett, 1989; Garrett, 1989; Kraus, 1990), where $N = [(g/\rho)(\partial\rho/\partial z)]^{1/2}$ is the Brunt-Väisälä frequency.

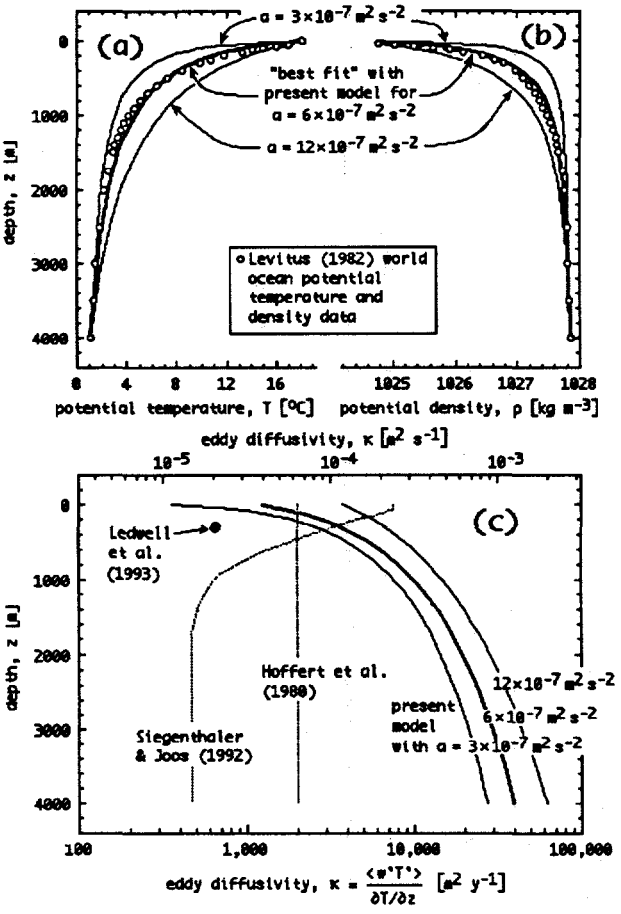


Figure 6. Vertical ocean profiles of (a) potential temperature, (b) potential density and (c) eddy diffusivity, κ. The upper panel shows steady state solutions of the present model matched to ocean surface and bottom boundary conditions for three values of the constant: $a = \kappa N$. The best fit value, $a = 6 \times 10^{-7}$ m^2 s^{-2}, is used in our transient model calculated. The constant eddy diffusivity of Hoffert et al. (1980), the radiocarbon-calibrated HILDA diffusivity distribution of Siegenthaler and Joos (1992) and the SF_6 tracer measurements of dipycnal diffusivity 310 m deep at a station 1200 km west of the Canary Islands by Ledwell et al. (1993) are also shown in (c).

The vertical turbulent mixing of heat (more precisely, the "diapycnal" mixing of heat" perpendicular to constant potential density surfaces slightly inclined to oceanic surfaces) is the return limb of the oceanic thermohaline circulation. The vertical eddy diffusivity typically used in UD ocean/climate models ($\approx 0.63 \times 10^{-4}$ m^2 s$^{-1} \approx 2000$ m^2/y) was derived by matching observed global mean ocean temperature profiles to the exponential steady state temperature profile assuming global mean upwelling typical of thermohaline

overturning ($\approx$ 0.13 μm $s^{-1} \approx$ 4 m y^{-1}). I will show in a later section of this report that the thermohaline overturn and upwelling are not independent of κ, but depend critically on the vertical diffusivity -- a fact with important implications for ocean GCMs.

Throughout the bulk of the ocean turbulence is distributed not uniformly but in patches that arise from internal wave instabilities, and in microstructure layers that result from the spreading of these patches. Internal wave- breaking is a highly intermittent phenomena and the fraction of ocean that is actually turbulent at any given time small. Monin and Ozmidov (1985, p. 34) estimate ~ 1%. Moum and Osborn (1986) report estimates of κ from velocity microstructure measurements of order 1×10^{-4} m^2 s^{-1} at ~ 1 km, increasing with depth at their site as expected from an N^{-1} dependence on density stratification. But Ledwell et al. (1993) found only $\kappa \sim 0.1 \times 10^{-4}$ m^2 s^{-1} three hundred meters beneath the surface using SF_6 dispersion as a dye tracer, similar to what Toole et al. (1994) found from microstructure data several kilometers deep where the internal wave field is presumably at background intensity. Toole et al. (1994) report much higher values near seamounts, speculating that basin-averaged mixing rates may be dominated by vigorous wave-breaking at boundaries.

There are other possibilities, like isopycnal diffusion producing net vertical mixing (Hoffert and Flannery, 1984). But the apparent discrepancy between measured κ values and what is needed to explain large scale thermohaline circulation is still not well understood. We suspect that ocean turbulence has been inadequately sampled; a potentially serious problem with intermittency at only 1%.

A unique feature of our ocean mixing study was evaluating the coefficient $a = \kappa N = 6 \times 10^{-7}$ m^2 s^{-2} in the stratification-dependent runs by matching the steady solution to observed ocean temperature profiles (Figure 6, top panel). Unexpectedly, we found that an a/N diffusivity can be calibrated to match Levitus (1982) data at least as well as a constant diffusivity. The lower panel of Figure 6 compares stratification-dependent κ-profiles with the constant κ_o case, and with the $\kappa(z)$ profile of Siegenthaler and Joos (1992) employed by their ocean carbon cycle. Note the major differences in values and depth-dependences.

Besides different World Ocean diffusivity laws, we studied the effect of different lower boundary conditions bracketing assumptions about Polar Sea internal mixing. Here we allowed for diffusive heat flux within the Polar Sea box proportional to the difference between is top and bottom temperatures, $T_p - T_b$ -- a departure from the purely downwelling Polar Sea of Hoffert et al. (1980). The net (downwelling plus mixing) heat flux at the bottom is then $F_b = \rho c A_p w_p [T_p + \gamma(T_p - T_b)] = \rho C A_o w [T_p + \gamma(T_p - T_b)]$, where ρc is the specific heat per unit volume and γ an empirical mixing coefficient. Equating the

downward heat flux at the sea floor from the Polar Sea to the net upward heat flux into the World Ocean, $F_b = \rho c A_o[\kappa(\partial T/\partial z) + wT]_{z=h_b}$ gives

$$@\ z = h_b: \quad \kappa\left(\frac{\partial T}{\partial z}\right)_{z=h_b} + wT_b = wT_p + \gamma w(T_p - T_b),$$

which we used as a bottom boundary condition instead of equation [2a]. The limit of a nondiffusive Polar Sea ($\gamma = 0$) recovers the lower boundary condition of a purely downwelling Polar Sea. The infinitely diffusive limit ($\gamma \rightarrow \infty$) maintains the bottom temperature equal to the Polar Sea surface temperature as the system evolves: $T(h_b) = T_p$. Numerically, we set $\gamma >> 1$ to compute the infinitely diffusive Polar Sea case. As previously, we parameterize warming of the surface Polar Sea by the *Polar Sea warming parameter*, $\pi = (T_p - T_{po})/(T_s - T_{so})$.

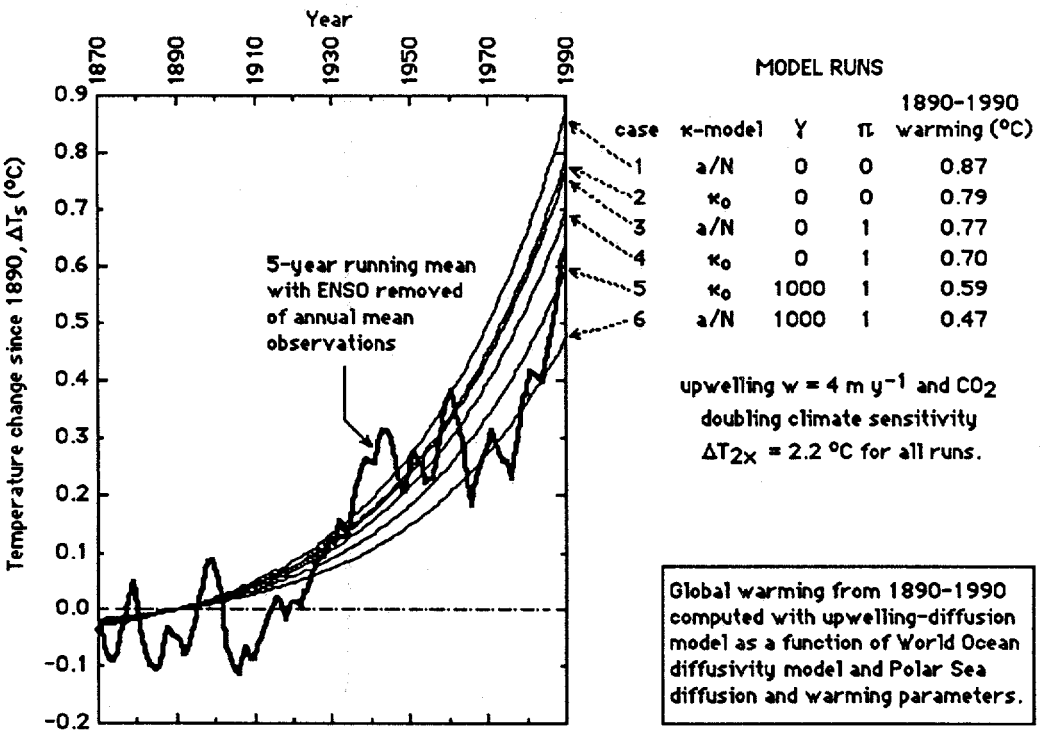


Figure 6. Observed global mean surface temperature variation relative to the year 1890 compared with global warming from 1870-1990 predicted from an upwelling-diffusion ocean/climate model for different World Ocean diffusion laws and Polar Sea mixing and surface warming rates.

Transient global warming responses of the past hundred years to the same radiative forcing scenario (IPCC 1990, no aerosols), climate sensitivity (ΔT_{2x} = 2.2 °C) and upwelling (w = 4 m y^{-1}), computed with an UD

ocean/climate model modified as described above are shown in Figure 6. The six model runs correspond to different assumptions about κ, γ and π. The thick solid line is a five-year running mean of observed global mean annual surface air temperature anomalies relative to the year 1890 corrected for the influence of the El Niño Southern Oscillation (ENSO).

The predicted global warming from 1890 to 1990 ranged from 0.47 to 0.87 ^{o}C depending on differences in model assumptions. The first two cases with $\pi = 0$ resulted in the greatest warming because with the Polar Sea surface temperature held constant ($T_p = T_{po}$) heat is constrained by the model to enter the deep sea though the main thermocline -- there is no "fast track" to the abyss bypassing thermocline bottleneck. When the Polar Sea route is suppressed heat transfer to the deep oceans is diffusion-limited in the stably stratified World Ocean thermocline. A $\kappa = a/N$ diffusivity law results in more warming than constant diffusivity because of the positive feedback from increases in the Brunt-Väisälä frequency $N = [(g/\rho)(\partial\rho/\partial z)]^{1/2}$ near the surface, as the surface warms, and temperature and density gradient $\partial T/\partial z$ and $\partial\rho/\partial z$ steepen.

Of the six cases, penetration of heat to the oceans was least for $\kappa = a/N$ with $\pi = \gamma = 0$ (case 1), corresponding to the *largest* global warming over the past hundred years -- 0.87 ^{o}C per century. Surprisingly, the effect of a stratification-dependent diffusivity law on the global warming rate can reverse if Polar Sea warming and diffusive mixing occur simultaneously. Cases 3 to 6 correspond to high-latitude warming equal to that of the global mean as assumed by the IPCC ($\pi = 1$). So long as the Polar Sea pumps heat to the sea floor by downwelling ($\gamma = 0$) (cases 3 and 4) the $\kappa = a/N$ diffusivity remains a positive feedback relative to the constant-κ law although both cases warm less than their $\pi = 0$ counterparts.

But oceanic heat uptake is greatest when the Polar Sea both warms and becomes diffusion-dominated ($\pi = 1$ and $\gamma \rightarrow \infty$), as in cases 5 and 6. The stratification-dependent World Ocean diffusivity law (case 6) gave the largest oceanic heat penetration and the *smallest* global warming of any case studied -- 0.47 ^{o}C per century, the most consistent with the observed ~0.5 ^{o}C per century. The paradoxical result that $\kappa = a/N$ produced both the greatest and smallest warming is a consequence of differences in the depth at which heat is injected to the World Ocean. When the high-latitude path to the deep is suppressed (case 1) eddy diffusivity and heat penetration is reduced. When high-latitude heat transfer to the deep sea is most active (case 6) the density gradients and the Brunt-Väisälä frequency N, become smaller at the base of the water column where $\kappa = a/N$ increases. Our calculations indicate temperature rises at the sea floor then propagate rapidly up the water column leading to more vigorous heat uptake via the Polar Sea route.

Our study shows that inadequately understood turbulent mixing occurring in the interior oceans and at high latitudes in deepwater-forming

zones can produce surprisingly different transient responses to the same radiative forcing, even when climate sensitivity, upwelling and starting profiles are the same. Evidently, much more work is needed to understand and characterize in models these subgrid ocean mixing processes.

Equator-to-Pole Temperature Distributions

If global mean temperature change from a specified radiative forcing ΔQ is the zeroth-order quantity needed in climate change assessments then the first-order climatic property is the equator-to-pole temperature distribution. We are, of course, far from able to *reliably* predict three-dimensional regional climate changes. What about zonally-averaged surface temperature change? The emphasis in the most recent IPCC assessment on transient climate simulations by coupled A/O GCMs tends to obscure that these models differ widely not only in equilibrium climate sensitivity but also in their zonal mean equator-to-pole patterns of temperature change (Kattenberg et al. 1996). The information is there but it does not jump off the page.

For this problem we again used schematic models, observational data and paleoclimate reconstructions to analyze and test GCM results. The simplest model for the meridional temperature distribution is a global energy balance model (EBM) resolved in latitude (North et al. 1981, Hoffert et al., 1983). It is convenient to work with the meridional coordinate $x = \sin\phi$ proportional to the surface area from the equator to latitude ϕ. A differential surface area element is then $dA = 2\pi R^2 dx$, where R ($\approx 6.37 \times 10^6$ m) is the Earth's radius.

In the steady state two differential equations describe poleward heat flow by the atmosphere/ocean system, F(x), and surface temperature, T(x):

$$\frac{dF}{dx} = 2\pi R^2[S(1-\alpha) - I + \Delta Q], \qquad [7]$$

$$\frac{dT}{dx} = -\frac{F}{2\pi R^2(1-x^2)D}. \qquad [8]$$

Here S(x) is incident solar flux, $\alpha(x)$ planetary albedo [$\alpha(x,T)$ if albedo-temperature feedback is included] and I(T) outgoing longwave flux to space parameterized in terms of surface temperature. We use a simple linear law, $I(T) = A + \lambda T$, where A is a calibrated constant and λ (≈ 2.2 W m^{-2} K^{-1}) the radiative damping coefficient.

Our base state ($\Delta Q = 0$) solutions are compared with satellite data and meteorological observations of "current climate" in Figure 7 assuming symmetry about the equator; that is, data at a given (positive or negative) latitude are averaged. Again, the observations are recovered reasonably well by a schematic NYU model -- one that is presumably not "too simple" to capture the essential physics. Equation [7] says the net radiant energy flux

absorbed (absorbed solar less outgoing longwave plus forcing) is balanced by the increase in poleward heat flow per unit surface area. Equation [8] represents poleward heat flow by Fourier heat conduction law in terms of an empirical "thermal diffusion" parameter, D. The factor $(1 - x^2)$ in the numerator comes from the sphericity of the Earth. North et al. (1981) assume D is constant.

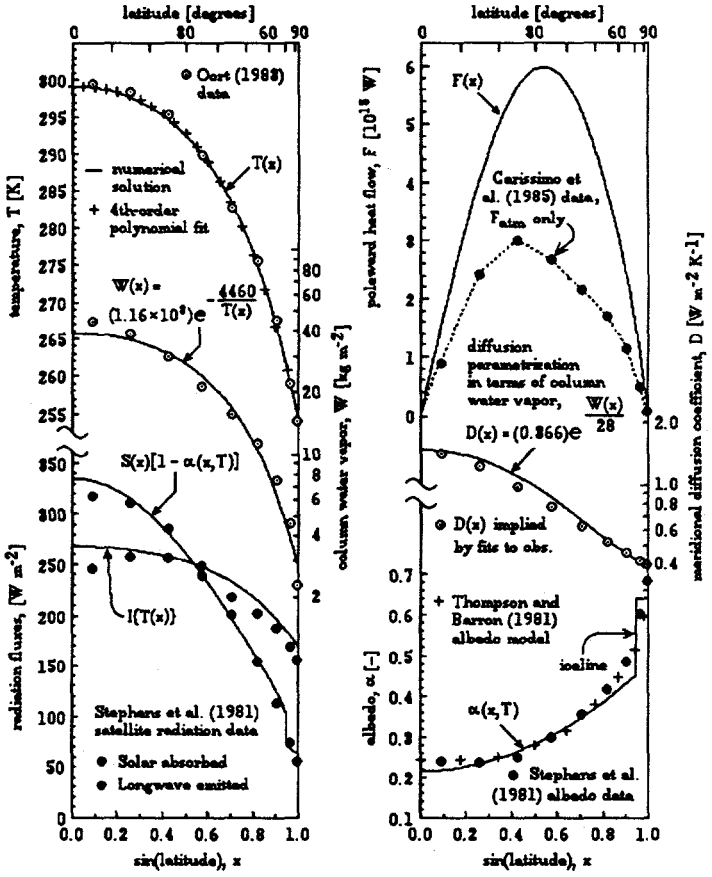


Figure 7. Latitudinal distribution of properties characterizing the zonal mean energy balance of the present climate (The NH and SH are averaged). The data are from satellite and meteorological observations, as cited. The theory (solid lines) is the NYU EBM for the current climate. The left panel shows surface temperature, column water vapor, absorbed solar radiation and outgoing longwave radiation. The right panel shows poleward heat flow, the thermal diffusion coefficient that recovers the observed T(x) and the planetary albedo.

We found that matching the present equator-to-pole temperature distribution requires that D vary with latitude and used observations to derive a correlation for D based on the water vapor in the atmospheric column (Figure 7). Physically, this suggests poleward heat flow by latent heat (which depends on water vapor content) is an important mechanism. A correlation with water vapor is equivalent to a correlation of with temperature, D(T), since water vapor is a function of T. The results is that the equations are nonlinear even if albedo-temperature feedback is excluded. Accordingly, it is more efficient to solve them by numerical integration than to use the Legendre polynomial expansions of North et al. (1981).

Mathematically [7] and [8] for the hemispherically symmetric case are coupled ordinary differential equations (ODEs) to be solved between the equator (x = 0) and pole (x =1) with the slight complication that the temperature derivative is singular at the pole.

Integration is by a 4th-order Runge-Kutta scheme from x = 1 toward x = 0 using iterated "shooting" to satisfy split boundary conditions. The first step is guessing a polar temperature, $T(1) = T_p$, and finding the corresponding temperature derivative from the regularity condition,

$$\left\{\frac{dT}{dx}\right\}_{x=1} = \frac{S_p(1 - \alpha_p) - I_p}{2D_p},$$

where subscript p denotes conditions at x = 1. On reaching the equator the program tests for hemispheric symmetry: $\{dT/dx\}_{x=0} = 0$. If hemispheric symmetry is violated by more than a specified error the program uses Newton-Raphson iteration to chose the next T_p, and the integration is repeated. Convergence is normally obtained in four or five iterations.

Equation [7] can also be integrated for the total poleward flux from satellite observations of S(x), $\alpha(x)$ and I(x) without invoking [8]. Total poleward heat flow versus latitude derived from the satellite radiation balance typically peaks at $F_{max} \sim 6$ PW (1 PW = 10^{15} W) at 30 degrees latitude (Figure 7, right panel). Poleward heat flow by the oceans is difficult to measure directly and is normally evaluated as a residual by subtracting "observed" atmospheric poleward heat flow from F(x). Estimates by Carissimo et al. (1985) were that oceanic fluxes peak at ~ 3 PW although this has been revised downward to ~ 2 PW lately (Trenberth and Solomon, 1994). To make credible equator-to-pole temperature change predictions it is important to first determine how total F(x) distribution changes when the system is radiatively forced. Only then is appropriate to attribute these changes to mechanisms in the atmosphere and oceans.

To study these issues we computed the equator-to-pole temperature distributions of the Mid Cretaceous (100 My BP) when mean temperature was ~ 9 °C warmer with our EBM under various assumptions. Although the Cretaceous is an extreme case, there is enough carbon in the fossil reserve to push surface temperatures this high after a century or so of unrestrained coal burning (Hoffert, 1992). Paleoclimate reconstructions indicate that equator-to-pole temperature gradients 100 My BP were only half today's (Barron and Washington, 1985). And such more uniform meridional distributions of temperature are characteristic of other warm periods as well: the Holocene Climatic Optimum (6 ky BP), the Eemian interglacial (125 ky BP), and the Pliocene (3-4 My), and Eocene (45 My) warmings (Crowley and North, 1981; Covey et al., 1996).

This presents a fluid dynamical paradox: More poleward heat flow is needed to produce flatter equator-to-pole temperature distributions; whereas most poleward heat transport mechanisms *decrease* in intensity as the meridional temperature gradient gets smaller -- these include thermal wind and baroclinic eddies in the atmosphere, and the wind-driven and "thermo" part of the thermohaline circulations in the oceans (Hoffert et al., 1983). How does the climate system accomplish this trick? Mechanisms which could possibly increase poleward heat flow during global warming with diminished thermal gradients are latent heat flow in the atmosphere and an enhanced equator-to-pole salt gradient in the oceans.

In any event, a corollary of flatter equator-to-pole temperature distributions as climate warms is that tropical temperatures remain stable. Most warming occurs at high latitudes. The reasons for this are still being sorted out and are highly controversial (see below).

Implication of a "Universal" Normalized Temperature Change

There are fundamental problems with apportioning poleward heat flow in GCMs applied to both present and paleo- climates. For example, the most recent GCM simulations of the mid-Cretaceous climate by Barron et al. (1995) has a control run (present climate) with a peak ocean heat of ~ 0.4 PW -- five times less than the best estimate for the present-day ocean (Trenberth and Solomon, 1994). It is difficult to know if an increase in ocean heat flow is needed to explain warmer climates, as proposed by Rind and Chandler (1991) based on their GCM simulations, when the present ocean heat flow is so problematical. A logical first step is to determine how *total* poleward heat flow changes. Paleotemperature reconstructions from the Former Soviet Union supplied by Andrei Lapenis allowed us to make progress on this question.

The Russian data summarized in Figure 8 supports Michael Budyko's hypothesis that equilibrium equator-to-pole surface temperatures changes , $\Delta T(x) = T_s(x) - T_{so}(x)$, have a nearly "universal" form when normalized by their hemispheric mean temperature changes, ΔT_m. Budyko argues that the distributions of zonal mean temperature change for different climates exhibit less variability than the corresponding distributions for $2 \times CO_2$ computed by different GCMs (personal communication). This is an interesting point, but it is obviously contingent on the reliability of the paleodata. Many discussions of this point tend to devolve to questions about the Russian data. Among other things the present DoE grant has supported, in part, the publication and archiving of the Russian paleodata including a preliminary error analysis (Kheshgi and Lapenis, 1996). This at least permits the issue to be discussed in peer-reviewed publications. However, much more needs to be done.

Universality of the temperature change pattern is one thing, the specific pattern another. We fit the nine paleoclimate data sets for zonal mean

temperature change ΔT normalized to the hemispheric mean temperature change ΔT_m with a forth- degree polynomial in x = sin(latitude),

$$\frac{\Delta T(x)}{\Delta T_m} = \delta_0 + \delta_2 x^2 + \delta_4 x^4, \qquad [9]$$

where $\delta_0 = 0.25$, $\delta_2 = 1.26$ and $\delta_4 = 1.65$. The GCM climate sensitivity intercomparisons of Cess et al. (1989) also assume a "universal" pattern of temperature change; one in which SST changes by the same amount at all latitudes ($\pm$ 2 °C). This is the horizontal line $\Delta T/\Delta T_m = 1$ in Figure 8. This assumption was made for reasons of expediency, but must certainly be considered unrealistic in light of Figure 8.

What is most striking in the data is the relative insensitivity of tropical temperature. Most of the data is for prior warm periods where the stability of tropical temperatures is well-supported by data. But some questions have been raised regarding the tropics during the Last Glacial Maximum (LGM).

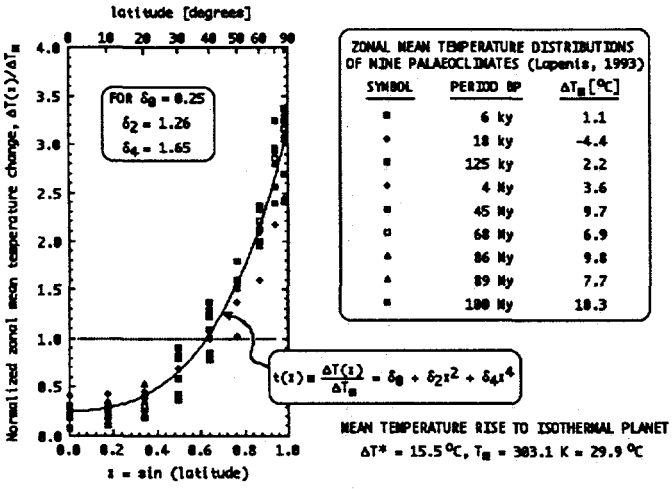


Figure 8. The equator-to-pole temperature *change* distribution of nine paleoclimates normalized to the hemispheric mean temperature change (Kheshgi and Lapenis, 1996). Most of these data are from prior warm periods, although the Russian interpretation of CLIMAP and other data for the LGM 18 ky BP are also shown. The solid curve is a forth degree hemispherically symmetric polynomial fit to this data set.

The paleoclimate findings of the CLIMAP program, based on species distributions of planktonic foraminifera in ocean cores, were that tropical cooling during the LGM was only 1±1 °C -- significantly less than the 3-5 °C cooling of the global mean. But new evidence summarized by Broecker (1996); based on strontium-calcium ratios in coral, the descent of snow lines on tropical mountains, and noble gas isotopes in Brazilian ground water; indicates cooling similar to the global mean for tropical land and shoreline environments. To complicate matters, in addition to the foraminifera speciation of CLIMAP, evidence from oxygen isotopes and alketones also support weak cooling of the tropical ocean surface. Why land and ocean LGM

data disagree is unclear, but does not invalidate equation [9] which predicts ~ 1/4 the temperature change at the equator as the global mean for *warmer* climates. Indeed, the evidence for stable tropics during global warming is holding up rather well.

The hot issue now (no pun intended) is why the tropical warm pool of the Pacific remains thermally stable in the present climatic regime. Ramanathan and Collins (1991) advanced the "Thermostat Hypothesis" that the atmosphere-ocean system limits SSTs to the observed maximum of 303 K (30 °C) because strong negative feedback is produced by shielding of surface insolation by highly reflective cirrus anvil clouds associated with tropical deep convection. Their theory is based on three premises: (1) the "super greenhouse effect" results in the outgoing longwave flux ceasing to function as a negative feedback on SSTs; (2) local analysis of changes in SST versus top-of-the-atmosphere, shortwave cloud forcing occurring in association with the 1985-87 El Niño Southern Oscillation "cycle" indicates cloudiness acts as a negative feedback of ~ 22 W m^{-2} K^{-1}; and (3) surface evaporation is unable to function as a limiting mechanism because evaporation adds moisture to the boundary layer and enhances the super greenhouse effect and net water vapor is imported to the tropics weakening evaporation there.

As Waliser (1996) has summarized contemporary critics of Ramanathan and Collins (1991) basically challenge the model on two grounds: Firstly that evaporation, more so than cloud forcing, is the fundamental mechanism limiting SST, and secondly that the large-scale circulation plays an important role, particularly the coupling between the large-scale circulation and the local environment -- points made over a decade ago by Hoffert et al. (1983) in connection with the tropical temperature stability of warm paleoclimates.

Notice that we do not have to understand the physics leading to the normalized temperature change distribution of equation [9] to analyze its implications for total poleward heat flow. Adding equation [9] to the present surface temperature $T_o(x)$ expressed as an even 4th degree polynomial gives the equator-to-pole temperature distribution for arbitrary radiative forcing: $T(x) = T_o(x) + \Delta T(x, \Delta T_m)$, where $\Delta T_m = \Delta Q/\lambda$. Total poleward heat flow distribution for the symmetric planet case follows from the integral of [7]:

$$F(x; \Delta T_m) = 2\pi R^2 \int_0^x \{S(x)[1 - \alpha(x)] - I[T(x; \Delta T_m)] + \lambda \Delta T_m\}\, dx \qquad [10]$$

The result is shown in Figure 9 for our nominal climate sensivity of ΔT_{2x} = 2.3 °C.

In this model the equator-to-pole temperature distribution gets flatter and flatter as the planet warms until it becomes uniform for a global warming of ~ 15.5 °C. We have no data for climates this hot; although it may well be

that the model breaks down before then. An important conclusion that follows from the global energy balance plus the universal curve for temperature change in Figure 8 is that total poleward heat flow increases by ~ 0.13 PW per °C of global mean warming if the albedo distribution remains constant. This too can be seen in Figure 9.

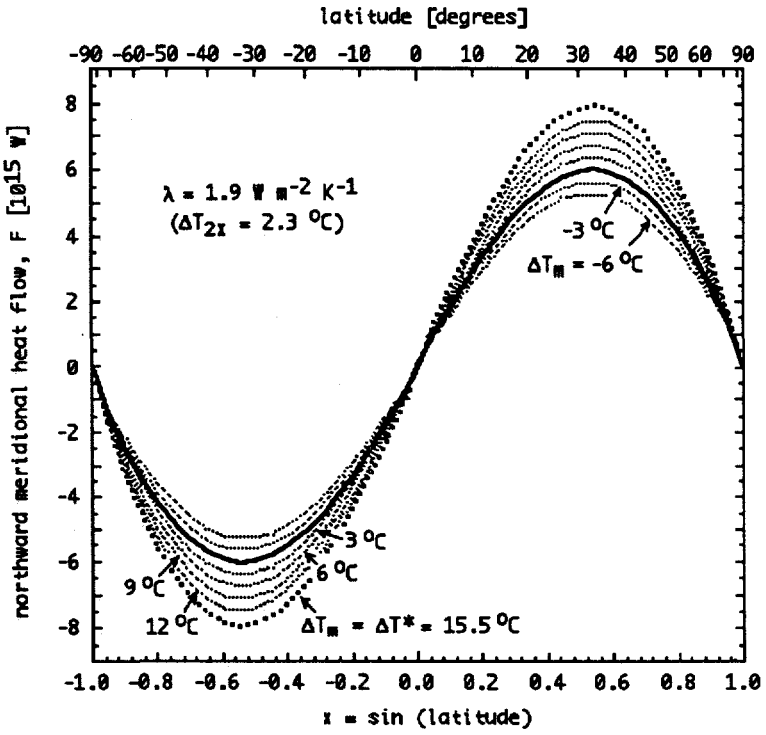


Figure 9. Variation of poleward heat flow as a function of latitude for various global mean temperature changes ΔT_m implied by the universal curve for normalized temperature and the global energy balance model.

Poleward Heat Flow by the Oceans

The preceding analysis indicated that additional heat must be exported poleward during global warming. Rind and Chandler (1991) and others, based on GCM results, have suggested that poleward heat flow increases are produced by ocean circulation changes to get the surface temperature distributions of warm paleoclimate. However, this is based on the assumption that the atmospheric transports are being correctly computed by GCMs and that changes in these transports are also correctly computed. We noted earlier the Cretaceous simulations of Barron et al. (1995) with a state-of-the-art GCM (GENESIS) which had present oceanic peak poleward fluxes of 0.4 PW.

As previously, we invoke a schematic model to analyze ocean heat flows. Thermohaline overturning driven by the equator-to-pole density gradient at the ocean surface is known to be the major mechanism for poleward heat flow (Bryan, 1987) It can be shown from order of magnitude

analysis the horizontal momentum balance of the thermocline in the geostrophic approximation, and from continuity, that the vertical upwelling in the thermocline is (Pedlosky, 1986) $W \sim [g\Delta\rho/(\rho f)](z^*/L)^2$, where g is gravitational acceleration (~ 10 m s^{-2}), f the Coriolis parameter (~ 1×10^{-4} s^{-1}), ρ the density (~ 1000 kg m^{-3}), $\Delta\rho$ the equator-to-pole density difference (~ 4 kg m^{-3}), L a horizontal length scale (~ 6×10^6 m) and z^* a thermocline depth scale. Substituting $z^* \sim \kappa/W$ for an upwelling-diffusing ocean, and solving for the upwelling scale gives $W \sim [(g/f)(\Delta\rho/\rho)]^{1/3}(\kappa/L)^{2/3}$. This is actually a scale for the peak upwelling velocity. In this thermohaline-driven picture, upwelling goes to zero at the top and bottom of the ocean. The mean value for uniform upwelling calculations can be written

$$w(\kappa) = \beta\left(\frac{g\Delta\rho}{f\rho}\right)^{1/3}\left(\frac{\kappa}{L}\right)^{2/3}, \qquad [11]$$

where β is a coefficient of order unity. It is encouraging that $\beta = 0.36$ and the other parameter values given above recover our calibrated upwelling, $w \sim 1.3 \times 10^{-7}$ m s^{-1} (4 m y^{-1}) for the eddy diffusivity that recovers the global temperature profile (Hoffert et al., 1980): $\kappa = 0.63 \times 10^{-4}$ m^2 s^{-1} (2000 m^2 y^{-1}) What is important is that upwelling and diffusion are not independent. If this scaling is correct, we expect the intensity of the thermohaline circulation and poleward heat flow to both scale with $\kappa^{2/3}$. Carrying the analysis a few steps further we derived the following approximate expression for oceanic poleward heat flow by the ocean as a function of latitude

$$F_{ocean}(x;\kappa) = F^*(x - x^3), \qquad [12]$$

where $F^* = (1/2)\rho c A_O \Delta T^* w(\kappa)$ is a heat flow scale; c being the specific heat per unit mass of seawater (~ 4000 J kg^{-1} K^{-1}), A_O the oceanic area (~ 3.3×10^{14} m^2) and ΔT^* the equator-to-pole ocean temperature gradient (~ 26 K). $w(\kappa)$ is from [11].

Figure 10 reveals major inconsistencies between observations and models of poleward ocean heat flows. There are no reliable direct observations of global scale oceanic poleward heat flow. (The difficulty is making highly resolved simultaneous measurements of temperature and velocity across an entire basin.) The left panel of Figure 10 illustrates a conventional residual estimate of ocean poleward heat flow. The implied peaks near 30 degrees latitude in both hemispheres have magnitudes of ~ 3 PW (Carissimo, 1985). The right hand panel illustrates that current ocean models produce only ~ 1 PW peak heat flow per hemisphere. The three-dimensional Semptner-Chervin (1992) eddy-resolving OGCM and our schematic model of equation [12] are both shown. Interestingly, our simple model produces poleward ocean fluxes of the same order as Semptner-

Chervin's horizontal eddy-resolving GCM while avoiding its pathological heat flow reversal in the Southern Hemisphere.

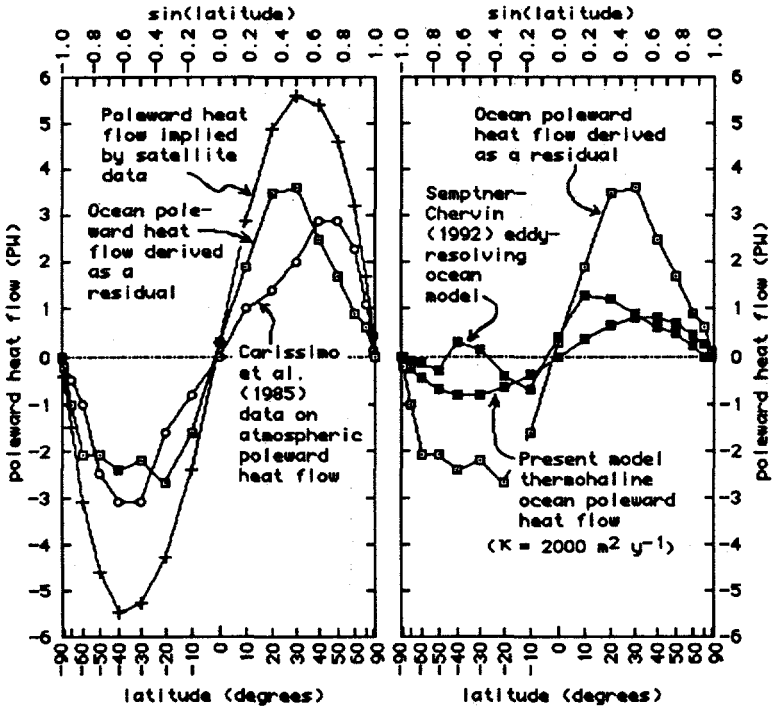


Figure 10. Global meridional heat flows in PW (1 Petawatt = 10^{15} W) versus latitude (positive toward the North Pole). The left panel shows total, atmospheric and ocean heat flow derived as a residual by Carissimo et al. The right panel compares the Carissimo residual with ocean poleward heat flows computed by the Semptner-Chervin horizontal eddy-resolving O GCM and our schematic model.

Thus, present day model-derived oceanic heat flows are only ~ 1/6 of the 6 PW peak heat flow to the poles implied by satellite data and ~ 1/3 of the 3 PW oceanic heat flow we expect as the residual of the total less the Carissimo (1985) atmospheric heat flow. The apparently weak poleward heat flows of modeled oceanic flow components compared with the residual-implied values has been recognized for some time. Current research, including our proposed effort, is more oriented toward reinterpreting present atmospheric heat flows as closer to 5/6 of the total than modelling higher oceanic heat flows. The poleward oceanic heat flow shown in the right panel of Figure 10 from our approximate analytical model developed at NYU scales with $\kappa^{2/3}$, where κ is the vertical eddy diffusivity. It is extremely difficult to raise the poleward heat flow for fixed surface temperature, salinity and stress distributions to 3 or even 2 PW without invoking unrealistically large subgrid vertical diffusivity (Bryan, 1985).

With the ocean heat flow such a small component of the total, it is hard to imagine increases in ocean heat flow explaining increased total heat flow with global warming. Indeed, numerical simulations of equilibrium climate change with coupled atmosphere/ocean climate models using the Bryan-Cox formulation show peak oceanic poleward heat flow remains roughly constant and $\leq$ 1 PW as the CO_2 concentration of the atmosphere increases from 1/2 to 8 times present levels (Manabe and Bryan, 1985, Fig. 16). These calculations

were done at coarse resolution and for idealized geography of the continents, but even the eddy-resolving realistic geography model of Semptner and Chervin (1992) produced similar peak oceanic heat flows of only ~ 1 PW (Figure 10, right panel).

We conclude from our studies that the thermocline depth scale, global upwelling and poleward ocean heat flow are all controlled by small-scale turbulent eddy diffusivity -- an input to OGCMs. But what controls small-scale turbulence in stably-stratified ocean interiors? Certainly not the explicitly resolved large-scale current field. The most likely answer is the intermittent steepening of internal gravity waves as localized shear flows, and their subsequent breaking. What pumps the internal wave field? And where does the energy come from ultimately dissipated by small-scale turbulence? As far back as the first upwelling-diffusion model Munk (1966) observed the similarity between the amount of energy required for ocean mixing and the dissipation rate of lunar tides (Garrett, 1979). It is a sobering thought that the thermocline depth so critical to the evolution of climate and life on Earth may depend critically on whether our planet captured (or ejected by impact) its massive satellite, the Moon.

CONCLUDING REMARKS

Clearly, it is important to have useful and credible tools for policy analysis before climate itself changes. Thus climate system models employed in our research at NYU and by the Intergovernmental Panel on Climate Change (IPCC) are motivated, at least in part, by the desire to make timely predictions of anthropogenic climatic impacts from greenhouse gas and aerosol emissions across the chain of causality from emissions to impacts.

The Navier-Stokes equations describing the behavior of the Earth's atmosphere and oceans have been known since the nineteenth century. However, until the advent of high speed digital computers these equations could not be solved. Even now, and even with the most advanced supercomputers with parallelized architectures, it is only possible to compute present climate and climatic change approximately. Although the physics of the Earth's climate system is in principle known there is no "standard model" that can answer all of the questions of interest to policy makers. The problem is not so much with the underlying physics as with complex interactions among components beyond the range of explicit computability. It is a fact of life that quantitative analysis of climate change is virtually impossible without computers. But limitations on computer speed and memory make modelling as much an art as a science; albeit an art subject to the strict discipline of observation and experiment.

Clearly, there are advantages and disadvantages to climate system models of varying levels of complexity. Complex models provide more detail; simple models more insight. An important concept in climate system

modelling is the notion of a hierarchy of models of differing levels of complexity, dimensionality and spatial resolution, each of which may be optimum for answering different questions. It is not meaningful to judge one level better or worse than another independent of the context of analysis.

Ideally, one seeks a balance with each component of the climate system represented at an appropriate level of detail. How to do this is the modeler's "art". Their is no methodological crank to turn, although some overall principles are clear: For example, it would be an inefficient use of computer resources to couple a detailed model for some part of the system with little effect on the result to one with crudely represented physical processes that dominates the result. Einstein once quipped that "everything should be as simple as possible, but no simpler." Generations of modelers have agonized over what "no simpler" means. And this has been a particularly important issue for assessments of anthropogenic change climate change conducted by the IPCC.

The most general computer models for climate change employed by the IPCC are the general circulation models (GCMs), which solve the equations of the atmosphere and oceans approximately by breaking their domains up into volumetric grids, or boxes, each of which is assigned an average value for properties like velocity, temperature, humidity (atmosphere) and salt (oceans). The size of the box is the models' spatial resolution. The smaller the box, the higher the resolution. An assumption of GCM research is that the realism of climate simulations will improve as the resolution increases.

In practice, the speed of supercomputers is insufficient to resolve important subgrid process. Phenomena occurring over length scales smaller than those of the most highly resolved GCMs, processes that cannot be ignored, include cloud formation and cloud interactions with atmospheric radiation; sulfate aerosol dynamics and light scattering; ocean plumes and boundary layers; subgrid turbulent eddies in both the atmosphere and oceans; atmosphere/biosphere exchanges of mass, energy and momentum; terrestrial biosphere growth, decay and species interactions; and marine biosphere ecosystem dynamics -- to cite a few. Mismatches between the scale of these processes and computationally realizable grid scales in global models is well-know problem of Earth System Science.

To account for subgrid climate processes the approach in has been to "parameterize" -- that is, to use empirical or semi-empirical relations to approximate net effects at the resolution scale of the model. It is important to stress that all climate system models contain empirical parameterizations; and that no model derives its results entirely from first principles. The main conceptual difference between simple and complex models is the hierarchical level at which the empiricism enters.

It essential, for example, to account for the heat and carbon that enter the oceans as the climate warms from the greenhouse effect of CO_2 emitted by

fossil fuel burning. The internal mixing and transport in the oceans of this energy and mass invading at the air/sea interface are key processes to be represented in any model predicting future carbon dioxide, climate and sea level variations. The rate at which heat and dissolved carbon penetrate the thermocline (roughly the first kilometer of ocean depth) controls how much global warming is realized for a given radiative forcing, and how much CO_2 remains in the atmosphere. In principle, these processes could be computed by coupled atmosphere/ocean (A/O) GCMs. But A/O OGCMs are too time-consuming to run on computers for many different emissions scenarios; and do not necessarily yield more realistic results than simpler models -- provided, of course, that the simpler models are artfully designed. For this reason the IPCC scenario analyses were done with far simpler and computationally much faster models than GCMs, many of which were pioneered by our group at NYU.

How does one judge the most appropriate level for climate system modlling? This is partly art, and partly science. Consider a parameterization of the world's oceans employed in several IPCC carbon cycle, climate, sea level calculations: the one-dimensional (1D) upwelling/diffusing ocean introduced by Hoffert et al. (1980, 1981) and subsequently developed by many other researchers.

In this paradigm, the three-dimensional world oceans are replaced by a single horizontally-averaged column in which carbon concentration and temperature vary with depth. The column exchanges mass and energy at its top with a well-mixed ocean surface layer, and at its bottom is fed by cold water from a downwelling polar sea. This 1D paradigm works well at simulating historical climate and carbon cycle variations. To simplify further by replacing the column with a zero-dimensional well-mixed box would make it too simple. A well-mixed box can't capture that the mixing time of the oceans is long compared to the rate that carbon emissions and radiative forcing at the surface are changing. The result would be incorrect rates of heat and mass uptake over time. Things are already "as simple as possible" with a 1D ocean, so we stop there.

Another frequently asked question is: How do we know if model predictions are credible? Science today recognizes that there is no way to prove the absolute truth of any hypothesis or model, since it is always possible that a different explanation might account for the same observations. In this sense, even the most-well established physical laws are "conditional." Rather, the test is whether a theory or model is false. The more independent challenges that a theory or model passes successfully, the more confidence ought we to have in it. Indeed, the testability of a conjecture has become a necessary condition for it to be considered in the domain of science. As Sir Karl Raimund Popper, philosopher of science and developer of the doctrine

of falsifiabity, put it: "Our belief in any particular natural law cannot have a safer basis than our unsuccessful critical attempts to refute it (Popper, 1969)."

The application of the falsifiability rule can be seen in the values of climate sensitivity, equivalent to equilibrium temperature change for a CO_2 doubling, estimated by the IPCC in the range of 1.5 to 4.5 degrees Celsius. Climate sensitivity is explicitly computed by GCMs, but is specified as input in simple ocean/climate models. At least four independent methods have been used to estimate the climate sensitivity: [1] from the slope of the Earth's radiation to space measured by satellites versus the underlying surface temperature (Warren and Schneider, 1979); [2] from three-dimensional equilibrium atmospheric GCMs (Cess et al., 1989); [3] from reconstructions of radiative forcing and climate response of ancient (paleo-) climates (Hoffert and Covey, 1992) and [4] from comparisons of ocean/climate model runs with historical global temperature records (Schlesinger et al., 1992; Santer et al., 1996). Each method has unique disadvantages and uncertainties. However, all of these independent results give results consistent with the IPCC range 1.5 to 4.5 degrees C for doubling CO_2, and inconsistent with values an order of magnitude lower hypothesized by some (Lindzen, 1995).

In principle, a climate sensitivity outside the IPCC range derived by any independent method could "falsify" it, and render questionable IPCC predictions of climatic change in the next century. However, all methods give results consistent with the IPCC range of warming for a CO_2 doubling -- with most of the uncertainty from uncertainties in cloud feedback. Since the methods are independent, the IPCC range can be attributed a reasonably high confidence.

On the other hand, for some unlikely but potentially important effects like the shutoff of deepwater circulation in the oceans or the sudden breakup of ice sheets, it is almost impossible to get probabilities from observations or first principles; and numerical simulations are at best suggestive. One technique that has been explored in the case of long-term sea level rise has been to use simple climate system models in conjunction with parameter probabilities assigned by "expert opinions" (Titus and Narayanan, 1996). Again, most of these sea level calculations were done with ocean carbon cycle models of the type developed at NYU; and the comparisons were coordinated by Vijay Narayanan, a former student in our Earth System Science program.

Finally, simple climate system models appear to have the drawback dealing only with global averages, whereas regional variations of temperature and precipitation change are needed to complete the link in integrated assessments from emissions to impacts. But in practice many present-day integrated assessments are conducted with models whose core transient climate calculations are done with simple ocean/climate models using regional distributions of temperature and precipitation scaled to global mean temperature distribution (Santer et al, 1990; Hulme et al, 1994).

The foregoing considerations are meant to explain the rationale underlying our focus on schematic climate system models in our own research, by others for integrated assessments of global climate change, and by the IPCC. We are not suggesting that a particular modelling methodology or level is inherently superior for climate system analysis for all time. Indeed, the consensus of the climate modelling community is that detailed three dimensional models of atmosphere and ocean dynamics, and correspondingly highly resolved models of the Earth's terrestrial and marine biota, are the long-term goals of Earth System Science. These modelling efforts need to proceed in parallel, and mutually reinforce, the more idealized system models we have developed for scenario analysis and climate policy-related work.

BIBLIOGRAPHY

Barron, E.J., and WM.. Washington (1985) In ET. Sundquist and W.S. Broecker (eds.),*The Carbon Cycle and Atmospheric CO_2: Natural Variations Archean to Present, Geophys. Mongr. Ser. 32,* pp. 546-553, AGU, Washington, DC.

Barron, E.J., et al. (1995) A "simulation" of mid-Cretaceous climate. *Paleoceanography,* **10** (5), 953-962.

Benkovitz, C.M. (1994) *Development and Testing of a High- Resolution Model for Tropospheric Sulfate Driven by Observation-Derived Meteorology.* Ph.D. dissertation, Department of Applied Science, New York University.*

Broccoli, A. (1994) Learning from past climates. *Nature,* **371,** 282.

Broecker, W. (1996) Glacial climate in the tropics. *Science,* **272,** 1902-1904.

Bryan, F. (1987) Parameter sensitivity of primitive equation ocean general circulation models. *J. Phys. Oceanogr.,* **17,** 970-985.

Carissimo, B.C., A.H. Oort and T.H. Vonder Haar (1985) Estimating the meridional energy transports in the atmosphere and ocean. *J. Phys. Oceanogr.* ,**15,** 82-91.

Cess, R.D., et al. (1989) Intercomparison of cloud-climate feedback as produced by 14 atmospheric general circulation models. *Science,* **245,** 513-516.

Collins, D. (1996) *Tropical Temperature Stability in the Paleorecord.* Ph.D. dissertation, Department of Applied Science, New York University.*

Covey, C., L. Cirbus Sloan and M.I. Hoffert (1996) Paleoclimate data constraints on climate sensitivity. *Climatic Change,* **32,** 165-184.*‡

Crowley, T.J., and G.R. North (1991) *Paleoclimatology.* Oxford University Press, NY, 339 pp.

Fu, R., et al. (1992) Cirrus-cloud thermostat for tropical sea surface temperature tested using satellite data. *Nature* **358,** 394-397.

Gargett, A.E. (1989) Ocean turbulence. *Ann. Rev. Fluid Mech.,* **21,** 419-451.

Garrett, C. (1979) Mixing in the ocean interior. *Dyn. Atmos. Oceans*, **3**, 239-265.

Gleckler, P.J., et al. (1995) Cloud-radiative effects on implied oceanic energy transport as simulated by atmospheric general circulation models. *Geophys. Res. Lett.*, **22**, 791-794.

Hansen, J., et al. (1995) Low-cost long-term monitoring of global climate forcings and feedbacks. *Climatic Change*, **31**, 247-271.

Hoffert, M.I. (1990) Climatic change and ocean bottom water formation: Are we missing something? In M.E. Schlesinger (ed.), *Climate-Ocean Interactions*, pp. 295-317, Kluwer Academic Publishers, the Netherlands, 385 pp.*‡

Hoffert, M.I. (1991) The effects of solar variability on climate. In M.E. Schlesinger (ed.), *Greenhouse-Gas-Induced Climatic Change*, 267-285, Elsevier Science Publishers B.V., Amsterdam, 615 pp.*‡

Hoffert, M.I. (1992) Climate sensitivity, climate feedbacks and policy implications. In I.M. Mintzer (ed.), *Confronting Climate Change: Risks, Implications and Responses*, pp. 33-54, Cambridge University Press, New York, 282 pp.*‡

Hoffert, M.I. (1993) Poleward heat flow feedback implied by paleoclimate reconstructions. *EOS, Trans. Am. Geophys. Union*, **74**, 50.*

Hoffert, M.I. and C. Covey (1992) Deriving global climate sensitivity from palaeoclimate reconstructions. *Nature* **360**: 573-576.*‡

Hoffert, M.I., and C. Covey (1993) Response to Lindzen (1993). Scientific Correspondence. *Nature* **363**, 26.

Hoffert, M.I., and B.P. Flannery (1985) Model projections of the time-dependent response to increased carbon dioxide. In M.C. MacCracken and F. Luther (eds.), *Projecting the Climatic Effects of Increasing Carbon Dioxide, DOE/ER 0237*, pp. 141-190, US Department of Energy, Washington, DC.

Hoffert, M.I., and S.D. Potter (1996) Energy supply. In R.G. Watts (ed.) *Engineering Response to Global Climate Change*, pp. 205-260, Lewis Publishers, Boca Raton, FL (in press).*‡

Hoffert, M.I., A.J. Callegari and C.T. Hsieh (1980) The role of deep sea heat storage in the secular response to climatic forcing. *J. Geophys. Res.*, **85**, 6667-6679.

Hoffert, M.I., A.J. Callegari and C.T, Hsieh (1981) A box-diffusion model carbon cycle model with upwelling, polar bottom water formation and a marine biosphere. In B. Bolin et al. (ed.) *Carbon Cycle Modelling, SCOPE 16*, pp. 287-305, John Wiley, New York, 390 pp.

Hoffert, M. I., Y.C. Wey, A.J. Callegari and W.S. Broecker (1978), Atmospheric response to deep sea injections of fossil fuel carbon dioxide, *Climatic Change*, **2**, 53-68.

Hoffert, M.I., B.F. Flannery, A.J. Callegari, C.T. Hsieh and W. Wiscombe (1983) Evaporation-limited tropical temperatures as a constrain on climate sensitivity. *J. Atmos. Sci.*, **40**, 1659-1668.

Hoffert, M.I., I.L. Karol, K.I. Kobal and D. Wuebbles (1990) Changes in atmospheric composition. In M.C. MacCracken, M.I. Budyko, A.D. Hecht and Y.A. Izrael (eds.) *Prospects for Future Climate.* Chapter 4, pp. 65-110, Lewis Publishers, Chelsea, MI, 270 pp.*

Houghton, J.T. et al. (eds.) (1990) *Climate Change: The IPCC Scientific Assessment.*, Cambridge University Press, New York, 364 pp.

Houghton, J.T. et al. (eds.) (1996) *Climate Change 1995: The Science of Climate Change: Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change,* Cambridge University Press, New York, 584 pp.

Hulme, M., S.C.B. Raper and T.M.L. Wigley, 1994. An integrated framework to address climate change (ESCAPE) and further developments of the global and regional climate modules (MAGICC). In *Integrative Assessment of Mitigation, Impacts, and Adaptation to Climate Change;* N.Nakicenovic, et al. (eds.), International Institute for Applied Systems Analysis, Laxenburg, Austria, pp. 289-308.

Jain, A.K., H.S. Kheshgi, M.I. Hoffert and D.J. Wuebbles (1995) Distribution of radiocarbon as a test of global carbon cycle models. *Global Biogeochemical Cycles,* **9** (10), 153-166.*‡

Kattenberg, A., et al. (1996) Climate models -- projections of future climate. In J.T. Houghton et al. (eds.) *Climate Change 1995: The Science of Climate Change: Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change,* pp. 285-357, Cambridge University Press, New York, 584 pp.

Kheshgi, H.S., and A.G. Lapenis (1996) Estimating the accuracy of Russian paleotemperature reconstructions. *Palaeogeography, Paleoclimatology, Palaeoecology,* **121,** 221-237.*‡

Kheshgi, H.S., B.P. Flannery and M.I. Hoffert (1991) Marine biota effects on the compositional structure of the world oceans. *J. Geophys. Res.,* **96** (C3), 4957-4969.*‡

Kheshgi, H.S., B.P. Flannery, M.I. Hoffert and A.G. Lapenis (1994) The effectiveness of marine CO_2 disposal. *Energy,* **19**, 976-974.*‡

Kraus, E.B. (1990) Diapycnal mixing. In *Climate-Ocean Interaction.* In M.E. Schlesinger (ed.), *Climate-Ocean Interactions,* pp. 269-293, Kluwer Academic Publishers, the Netherlands, 385 pp.

Lapenis, A.G., and M.V. Shabalova (1994) Global climate changes and moisture conditions in the intercontinental arid zones. *Climatic Change,* **27**, 283-297.*‡

Ledwell, J.R., A.J. Watson and S.L. Law (1993) Evidence for slow mixing across the pycnocline from an open ocean tracer release experiment. *Nature*, **364**, 701-703.

Levitus, S. (1982) *Climatological Atlas of the World Ocean*, NOAA Professional Paper 13, Geophysical Fluid Dynamics Laboratory, Princeton, NJ, 173 pp. plus 17 microfiches.

Lindzen, R.S. (1993) Palaeoclimate sensitivity. Scientific Correspondence. *Nature* **363**, 25-26.

Lindzen, R.S. (1995) How cold would we get under CO_2-less skies? *Physics Today*, **48**(2), 78-80.

McKay, D.S., et al. (1996) Search for past life on Mars: Possible relic biogenic activity in Martian meteorite ALH84001. *Science*, **273**, 924-930.

Maier-Reimer, E., and K. Hasselmann (1987) Transport and storage of CO_2 in the ocean -- an inorganic ocean-circulation carbon cycle model. *Clim. Dyn.*, **2**, 63-90.

Manabe, S., and K. Bryan Jr., 1985: CO2-induced change in a coupled ocean-atmosphere model and its paleoclimatic implications. J. Geophys. Res., **90** (C6), 11,689-11,707.

Monin, A.S., and R.V. Ozmidov (1985) *Turbulence in the Ocean*, Reidel, Boston, 1985, 247 pp.

Moum, J.N., and T.R. Osborn (1986) Mixing in the main thermocline. *J. Phys. Oceanogr.*, **16**, 1250-1259.

Munk, W.H. (1966) Abyssal recipes. *Deep Sea Res.*, **13**, 707-730.

North, G.R., R.F. Cahalan and J.A. Coakley (1981) Energy balance climate models. *Rev. Geophys. Space Phys.*, **19**, 91-121.

Oort, A.H. (1983) *Global Atmospheric Circulation Statistics*, 1958-1973. NOAA Professional Paper 14, NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, 180 pp. plus 47 microfiches.

Pedlosky, J. (1986) Thermocline theories. In H.D.I Abarbanel and W.R. Young (eds.), *General Circulation of the Ocean*, pp. 55-101, Springer-Verlag, 291 pp.

Popper, K.R., 1969. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge, 439 pp.

Potter, S.D. (1993) *Optimization of Microwave Power Transmission from Solar Power Satellites*. Ph.D. dissertation, Department of Applied Science, New York University.*

Raval, A. and V. Ramanathan (1989) Observational determination of the greenhouse effect. *Nature* **342**, 758-761.

Rind, D., and M. Chandler (1991) Increased ocean heat transports and warmer climate. *J. Geophys. Res.*, **96** (D4), 7437-7461.

Santer, B.D., T.M.L. Wigley, M.E. Schlesinger and J.B.F. Mitchell (1990) Developing Climate Scenarios from Equilibrium GCM Results. Max Planck Institute for Meteorology Report 47, Hamburg, Germany.

Schlesinger, M.E., X. Jiang and R.J. Charlson (1992) Implications of anthropogenic sulfate for sensitivity of the climate system. In L. Rosen and R. Glasser (eds.), *Climate Change and Energy Policy;* American Institute of Physics, pp. 75-108, New York.

Schopf, J.W. (ed.) (1983) *Earth's Earliest Biosphere: Its Origin and Evolution.* Princeton University Press, New Jersey, 543 pp.

Semtner, Jr., A.J., and R.M. Chervin (1992) Ocean circulation from an eddy resolving model. *J. Geophys. Res.*, **97** (C4), 5493-5550.

Siegenthaler, U. and F. Joos (1992) Use of a simple model for studying oceanic tracer distributions and the global carbon cycle. *Tellus*, **44B**, 186-207.

Stephens, G.I., G.G. Cambell and T.H. Vonder Haar (1981) Earth radiation budgets.

Thompson, S.L., and E. Barron (1981) Comparison of Cretaceous and present Earth albedos: Implications for the causes of paleoclimates. *J. Geology,* **89**, 143-167.

Titus, J.G., and V. Narayanan (1996) The risk of sea level rise. *Climatic Change*, **33**, 151-212*.

Toole, J.M., K.L. Polzin and R.W. Schmitt (1994) Estimates of diapycnal mixing in the abyssal ocean. *Science*, **264**, 1120-1123.

Trenberth, K.E., and A. Solomon (1994) The global heat balance: heat transports in the atmosphere and ocean. *Clim. Dyn.*, 10, 107-134.

UNEP/WMO (1992) *United Nations Framework Convention on Climate Change (Text).* UNEP/WMO, Climate Change Secretariat, Palais des Nations, 1211 Geneva 10, Switzerland.

Volk, T. (1984) Multi-property modeling of the marine biosphere in relation to global climate and carbon cycles. Ph.D. dissertation. New York University.

Volk, T. and M.I. Hoffert (1985) Ocean carbon pumps: Analysis of relative strengths and efficiencies in ocean-driven atmospheric CO_2 changes. In E.T. Sundquist and W.S. Broecker (eds.),*The Carbon Cycle and Atmospheric CO_2: Natural Variations Archean to Present, Geophys. Mongr. Ser.* 32, pp. 91-110, AGU, Washington, DC.

Waliser, D.E. (1996) Some considerations on the Thermostat Hypothesis. *Bull. Am. Meteorology. Soc.*, **77** (2), 357-360.

Warren, S.G., and S.H. Schneider (1979) Seasonal simulation as a test for uncertainties in the parameterizations of a Budyko-Sellers zonal climate model. *J. Atmos. Sci.* **36**, 1377-1391.

Washington, W.M., and C. Parkinson (1986) *An Introduction to Three-Dimensional Climate Modelling.* Oxford University Press, New York, 422 pp.

Wigley, T.M.L., and S.C.B. Raper (1987) Thermal expansion of sea water associated with global warming. *Nature*, **330**, 127-131.

Wigley, T.M.L., and S.C.B. Raper (1992) Implications for climate and sea level of revised IPCC emissions scenarios. *Nature*, **357**, 293-300.

Wright, D.G., and T.F. Stoker, 1991: A zonally averaged ocean model for the thermohaline circulation, Part I: Model development and flow dynamics. *J. Phys. Oceanogr.*, **21**, 1713-1739.

---

*Research supported partly or wholly by this grant.

‡Reprint appended to this report.

Climatic Change and Ocean Bottom Water Formation: Are We Missing Something?

MARTIN I. HOFFERT
Earth Systems Group
Department of Applied Science
New York University
New York, NY 10003
U.S.A.

ABSTRACT. The global thermohaline circulation is driven by negative buoyancy sources associated with surface cooling and salinity enhancement, with subsequent upwelling over all the world's oceans. Relatively shallow, thermally-driven convective overturning occurring at high latitudes which arises from heating of the atmosphere by the oceans is fairly well represented in current numerical ocean models, but deep convection from surface salinity sources including North Atlantic Deep Water (NADW) and, particularly, Antarctic Bottom Water (AABW) is not. A schematic model is developed for AABW which incorporates brine rejection during the seasonal sea ice cycle and sinking in boundary layers along the continental shelf and slope of Antarctica. This model predicts the formation of 30–40 Sv of bottom water from an annual sea ice freezing rate of order 1 Sv. These values and the predicted top-to-bottom potential density increment in the Southern Ocean are in reasonable accord with observations. When the amount of sea ice freezing is expressed as a function of hemispheric mean sea surface temperature (calibrated from the seasonal cycle), the model predicts the possible cutoff of AABW and most of the thermohaline circulation for a four degree Celsius global mean warming, an event which could be triggered in the next century by anthropogenic greenhouse gases.

1. INTRODUCTION

The sinking of dense, cold, salty and well-ventilated water from the surface to the bottom of the world's oceans not only drives the global thermohaline circulation, but has a profound effect on the ability of the ocean to mix heat and dissolved constituents. A sharp change or cutoff of global bottom water would almost certainly trigger transient climate feedbacks on decadal to millennia time scales of oceanic turnover, influence the carbon cycle of the seas, and could even produce anoxia in abyssal regions. It would therefore seem a matter of some concern to understand better bottom water formation in the present ocean, bottom water changes which may have occurred in the geologic past,

M. E. Schlesinger (ed.), Climate-Ocean Interaction, 295–317.
© 1990 *Kluwer Academic Publishers. Printed in the Netherlands.*

and whether anthropogenic climate changes could initiate bottom water changes in the foreseeable future.

A very comprehensive survey and assessment of bottom water research in oceanography to the the beginning of this decade is given by Warren (1981). This descriptive work is invaluable, but does not in itself provide a basis for estimating changes in deepwater circulation until it is incorporated into climate models, however simplified. More recent (and more controversial) ideas on bottom water formation and its possible variation during glacial-interglacial transitions can be found in Kellogg (1987) and Broecker (1987). Despite observational gaps associated with the difficulty of making ship measurements near sea ice where high-latitude deepwater forms, there appears to be sufficient information to do a better job of modelling these processes than is being done at present. I will argue here that new classes of models which couple seasonal brine rejection by sea ice, turbulent gravity layers, and possibly entrainment along continental slopes, are needed to capture Antarctic Bottom Water formation more realistically in numerical ocean circulation models. A relatively simple, schematic model will also be developed here that suggests Antarctic Deep Water formation is vulnerable to a global greenhouse warming from fossil fuel combustion and trace gas emissions.

There is good evidence that the thermal stratification of the deep ocean has varied markedly over earlier geological epoches (Berger, 1981; Savin, 1977). It is important in modelling changes in bottom water to understand the nature of this evidence. One extremely useful method for estimating ocean paleotemperature histories employs mass spectrometric techniques to measure anomalies in the oxygen-18 to oxygen-16 isotope ratios in calcium carbonate ($CaCO_3$) shells in sediment cores drilled out of the sea floor. The empirical proxy temperature scale in the range of 0°C to 30°C is based on the temperature dependence of isotopic fractionation during $CaCO_3$ precipitation on the shells of organisms (Schopf, 1980),

$$T = 16.9 - 4.38(\delta_c - \delta_w) + 0.10(\delta_c - \delta_w)^2 \ [°C] \quad ,$$

where the isotopic deviations (delta's) of carbonate and water are defined in relation to standard isotopic ratios,

$$\delta = \left[\frac{(^{18}O/^{16}O)_{sample} - (^{18}O/^{16}O)_{standard}}{(^{18}O/^{16}O)_{standard}} \right] \times 10^3 \quad .$$

Since planktonic forams live in the surface ocean, and benthic forams near the bottom, their delta's contain a record of surface and deep-water temperatures during their lifetimes when $CaCO_3$ was precipitated on their shells. (The delta's of seawater also come into play, so corrections are needed for glaciations which fractionate isotopes on the water side.) The top-to-bottom temperature difference of the tropical ocean derived from deep sea sediment cores over the past 80 million years is shown in Fig. 1. The record indicates deep water temperatures were only about 10°C below the tropical surface temperature until about 30 million years ago, after which the temperature difference rose progressively to some 26°C today (Savin, 1977). Since surface temperatures of the tropics are apparently constrained to within a few degrees of today's as a result of equatorial evaporative damping (Hoffert *et al.*, 1983), the increasing temperature difference between the top and bottom of the world's oceans is largely the result of high-latitude cooling and the corresponding development of the deep "cold water sphere" in all the world's oceans.

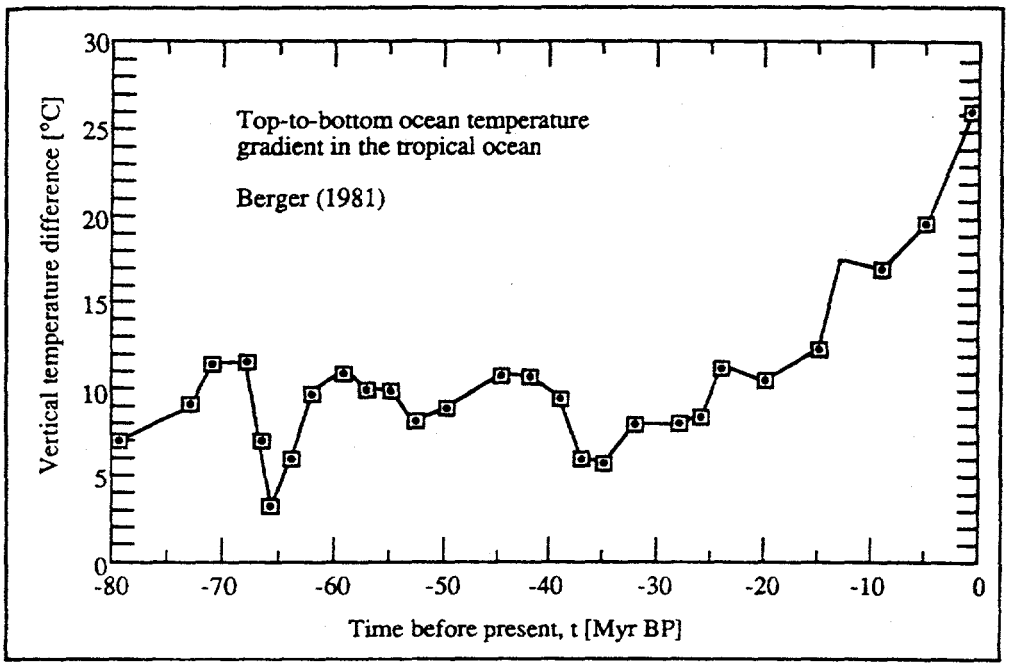


Figure 1. Variation of the vertical temperature difference of the tropical ocean over the past 80 million years deduced from the $^{18}O/^{16}O$ ratios in the calcium carbonate shells of planktonic (surface) and benthic (bottom) dwelling foraminifera (Berger, 1981). The paleotemperature record deduced from deep sea sediment cores shows a dramatic increase in the stability of the ocean over the past 30 million years as deep polar convection generated the "cold water sphere."

How could this happen? Continental drift was almost certainly a factor, particularly, the progressive physical isolation and polar location of the Antarctica continent since the final break-up and dispersal of supercontinent Pangea some 150 Myr BP. By the late Cretaceous (80 Myr BP), both East and West Antarctica stood on the same continental shelf roughly over the pole; by 30 Myr BP the continental separation had increased to the point where the Drake passage was open, and Antarctica had become a roughly circular continent at the South Pole surrounded by a "Southern Ocean" communicating deep water with the Atlantic, Pacific and Indian Oceans (Weller *et al.*, 1987). At the same time the presence of continental land masses near the poles in both hemispheres permitted the growth of glacial ice sheets, and the planet as a whole entered the ice ages of the past 50 Myr. The presence of high-latitude continental glaciers in the north which could accumulate land snow and ice over continental areas from one season to the next became the backdrop against which astronomical variations of the earth's orbital elements on 20–100 kyr time scales would produce glacial-interglacial transitions ("Milankovitch cycles"). But the most fundamental influence in forming cold global bottom water may have been the gradual build-up of sea ice in the Southern Ocean around the glaciated Antarctic continent as the global climate cooled.

2. BOTTOM WATER TODAY

The latitudinal asymmetry of the present continental distribution results in a dramatic difference in the way bottom water forms in the two hemispheres. In contrast to the "open" Southern Ocean around Antarctica, the main site of seasonal sea ice formation in the Northern Hemisphere is the Arctic ocean, a landlocked waterbody entirely inaccessible to the North Pacific and communicating weakly, if at all, with the North Atlantic. It is presently believed that the source of North Atlantic Deep Water (NADW) is the permanent pycnocline of the Iceland and Greenland Seas (Kellogg, 1987), a region of high salinity created by the climatological excess of evaporation less precipitation in that part of the North Atlantic. The current picture is that saline North Atlantic water enters the region through the Norwegian sea and destabilizes the water column making it susceptible to convection to intermediate depths. As NADW moves southward down the Atlantic it flows beneath the high-salinity (but warmer) Mediterranean outflow, and never does attain the density needed to sink to the bottom of the Atlantic. But on reaching the Southern Ocean, NADW mixes to the surface and is cooled in the Antarctic Circumpolar Current where it becomes part of the source water for the densest, coldest and deepest bottom water feeding all the world's oceans: Antarctic Bottom Water (AABW).

While the precise mechanisms for forming AABW are controversial, virtually all authors recognize the critical role of sea ice and brine rejection, particularly in the semi-enclosed Weddell and Ross Seas around Antarctica. A detailed discussion of observations and theories of sinking around Antarctica is given by Warren (1981). He emphasizes that in accounting for the formation of Weddell Sea Bottom Water and, by extension all Antarctic Bottom Water on the continental shelf and slope, it is necessary to represent three distinct processes: (1) how the density of the shelf water is increased to the point where it can sink through the relatively warm saline water at mid-depths; (2) how the water moves off the shelf, and (3) how it descends with entrainment to the sea floor. Whereas brine rejection by sea ice interacting with frictional and gravity boundary layers is normally thought to be the critical process, Gordon (1982) has speculated that episodic open areas in sea ice ("polynyas") generated by open-ocean convection during certain years can short-circuit the usual shelf-mixing process and thereby produce AABW directly. The sinking of AABW is thought to play a major role in driving the global thermohaline circulation. In particular, the cold deep water formed around Antarctica returns to the surface as global upwelling over all the world's oceans (Hoffert *et al.*, 1980).

3. OCEANIC CONVECTION

The formation of oceanic bottom water (and vertical mixing in general) is intimately related to the process of buoyant convection or "overturnings" in density-stratified fluids. In light of different interpretations in the literature regarding the role of convective motion in bottom water formation, I ask the reader to indulge me a brief review.

Convection normally refers to vertical motion induced by a gravitational instability in which heavier (denser) fluid overlies, or is adjacent to, lighter fluid. It is a distinctly different process from so-called "vertical advection" in which vertical motion results from other than gravitational buoyancy forces. But there is a caveat: Although all fluids (including seawater) are to some extent compressible, density changes associated with changes in hydrostatic pressure (altitude or depth) do not generate or suppress convection. It is therefore useful in assessing buoyancy effects to work not with the physical density, but with the *potential density*, ρ, of a fluid parcel, that is, the density the parcel would have were it brought adiabatically to sea level pressure (p_o = 1 atm). Likewise, *potential*

temperature is the temperature a parcel at some altitude or depth would hypothetically attain were it adiabatically compressed (atmosphere) or expanded (ocean) to p_o.

Bryan and Cox (1972) have developed a polynomial curve fit to the equation of state of seawater for use in numerical ocean models. The Bryan-Cox state equation evaluated at sea level pressure (Häkkinen, 1987) can be used to approximate potential density as a function of potential temperature and salinity,

$$\rho(T, S) = 1024.458 + a_T\Delta T + a_S\Delta S + a_{T2}\Delta T^2 + a_{S2}\Delta S^2 + a_{TS}\Delta T\Delta S$$

$$+ a_{T3}\Delta T^3 + a_{S3}\Delta S^3 + a_{T2S}\Delta T^2\Delta S + a_{S2T}\Delta S^2\Delta T [\text{kg m}^{-3}] \quad ,$$

where $\Delta T = T - 13.5$ [°C], $\Delta S = S - 32.6$ [‰], $a_T = -0.20134$, $a_S = 0.77096$, $a_{T2} = -0.49261 \times 10^{-2}$, $a_{S2} = 0.4609 \times 10^{-3}$, $a_{TS} = -0.20105 \times 10^{-2}$, $a_{T3} = 0.36597 \times 10^{-4}$, $a_{S3} = 0.65493 \times 10^{-5}$, $a_{T2S} = 0.37735 \times 10^{-4}$ and $a_{S2T} = 0.47371 \times 10^{-5}$. Here T is potential temperature in the units of °C and S is the salinity in units of ‰ (equivalent to parts per thousand, psu's or "per mille"). As a rule of thumb, oceanic potential density variations are dominated by potential temperature in the warm, upper thermocline, and by salinity variations at polar latitudes.

Since the general behavior of the state equation is not easy to visualize from the raw equations, I have presented a 3-D graphics view in Fig. 2 of the potential density surface versus T and S (the shaded surface) computed from the Bryan-Cox state equation. The solid curve shows the path in potential temperature-salinity-density space of the zonal mean southern hemisphere surface ocean from the equator to −72.5° latitude, this high latitude falling inside the Weddell and Ross Seas and thus being representative of Antarctic Surface Water (AASW). Also shown is the nearly isothermal descent of this water by seasonal salinity enhancement to 4000 meter depth as Antarctic bottom water (AABW) is formed (TS Data from Levitus, 1982). It will be assumed here that AABW is pumped downward by seasonal freezing of sea ice in the Southern Ocean as brine rejected by the ice enhances salinity and density of the liquid water phase, and causes sinking and flow down the Antarctic continental shelf and slope.

A much more idealized concept of bottom water formation as open ocean convection is *de facto* embodied in numerical ocean models. In the most general case, the buoyant stability of fluid columns is measured by their vertical potential density gradient. Stable columns have potential density increasing with decreasing elevation (light fluid above); vice versa for unstable columns (heavy fluid above). Stable columns suppress vertical mixing by turbulent eddies, unstable columns enhance it. Often the presence of an unstable gradient is presumed to increase vertical mixing so vigorously that a nearly instantaneous "convective overturning" occurs which wipes out the instability, that is, ρ becomes nearly constant with height (neutral stability). When this process is imposed in a numerical ocean model it is called a *convective adjustment*. Convective adjustments are "rapid," but it should be recognized that in the absence of negative buoyancy sources, stable columns too must eventually diffuse to neutrally stable profiles. The top-to-bottom temperature difference of the tropical oceans (Fig. 1) is very likely maintained by salinity-buoyancy pumps near the poles.

The low-albedo sea surface is normally warmer than the overlying atmosphere at any latitude so that surface sensible and latent heat heat fluxes are upwards, whereas net absorbed radiation flux is downward. The imbalance of opposing surface heat and radiation fluxes as a function of latitude is such that the oceans are an energy source to the atmosphere at high latitudes, and a sink in the tropics, with poleward meridional heat transfer by the upper ocean making up the difference. As depicted in Fig. 18 of Schlesinger and Jiang (1988), state-of-the-art numerical ocean models supply heat to the atmosphere at

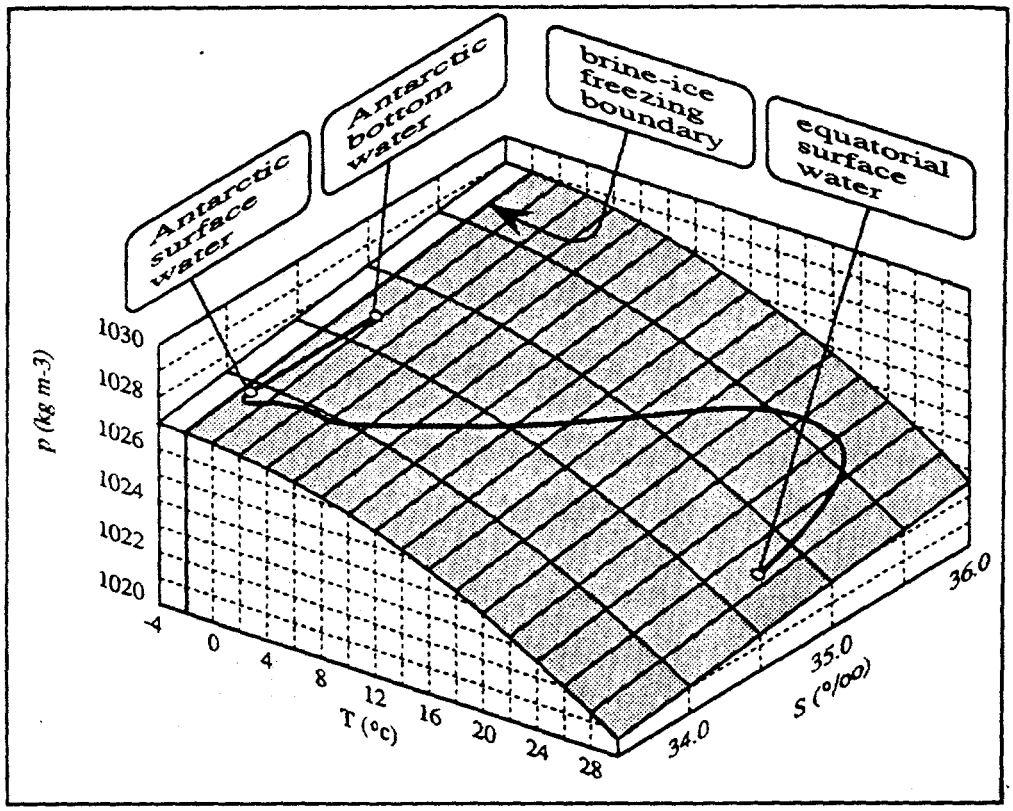


Figure 2. A three-dimensional view of the potential density of seawater as a function of potential temperature, T, and salinity, s. The slice of density surface shown extends over the T-S regime of the world's oceans as computed from the Bryan and Cox (1972) equation of state. The brine-sea boundary in this regime is also shown.

high latitudes from relatively shallow (0–500 m) depths by internally generating thermal convection ($\partial T/\partial z > 0$ and $\partial \rho/\partial z < 0$, with z positive downwards). An interesting effect has emerged from transient studies using three-dimensional coupled atmosphere-ocean climate models (Schlesinger and Jiang, 1988; Bryan and Spelman, 1985). In contrast to 1-D models which include only vertical heat fluxes by diffusion and upwelling (Hoffert *et al.*, 1980), the vertical mixing of heat (and density) in 3-D models is the sum of vertical thermal diffusion, vertical thermal advection (upwelling) and *vertical convection.* When an increase in atmospheric greenhouse gases causes global mean temperatures to climb, the 3-D models generally predict the suppression of convective cooling of the oceans at high latitudes such that downward diffusion minus advection dominates the global mean vertical heat flux. Although there are important model-to-model variations, one result is that high-latitude convective suppression is a major pathway for the penetration of heat to deeper levels during climatic transients.

These results underscore the fact that the upper and lower ocean are presently coupled at high latitudes largely by *thermal convection* in ocean circulation models. In contrast, Fig. 2 clearly indicates the high potential density AABW is generated nearly isothermally by *saline convection* most likely as a result of seasonal brine rejection by sea ice. As

Schlesinger and Jiang (1988) point out, the current generation of numerical ocean models do not generally include the effects of brine rejection from evaporation or the seasonal sea ice cycle as sources of convective overturning. On these grounds alone, they are missing some potentially important physics. As depicted by Broecker (1987, Fig. 2), a large-scale salt transport system apparently operates in today's ocean associated with the evaporative salinity pump of NADW. In this system water vapor flowing from the Atlantic to the Pacific in the atmosphere produces a counterpart salt current flowing down the Atlantic, around Africa, through the Southern Ocean and finally back northward to the Pacific. General circulation models allowing surface salinity sources to affect density through open ocean saline convection suggest this system is self-stabilizing. But even a salinity-pumped convective overturning model with interactive sea ice may be inadequate for calculating AABW because in reality the dense fluid is confined to thin boundary layers along the continental shelf and slope.

This is not to say that high-latitude thermal convection does not occur in the upper ocean, or that models incorporating it cannot predict deep sea warmings from CO_2-increases in the right general direction. Apart from the lack of a seasonal sea ice cycle, the Geophysical Fluid Dynamics Laboratory (GFDL) coupled atmosphere-ocean model of Bryan and Spelman (1985) did not have realistic geography or continental and sea floor orography, and does not include either the annual solar insolation cycle or cloud feedbacks. Manabe and Bryan (1985) were nonetheless able to use essentially the same model to show an increase in tropical surface temperature and a decrease in the top-to-bottom temperature difference of the ocean as a function of increasing atmospheric CO_2 in qualitative agreement with the paleoceanographic record (Fig.1). However, the magnitude of the simulated present top-to-deep temperature difference shown is less than half the observed value, and the bottom water flow rate is too small by a factor of five or more. To model the present Antarctic Bottom Water and changes therefrom realistically, it is necessary to know first how the sea ice freezing rate (and salt rejection) vary with climate, and secondly, how freezing rate changes induce changes in AABW density and flow rates.

4. THE SEASONAL SEA ICE CYCLE IN THE SOUTHERN OCEAN

The prediction of seasonal and climatological variations of sea ice using both stand-alone sea ice models and models coupled to atmosphere-ocean general circulation models is presently an active area of climate research. A very up-to-date review from the standpoint of 3-D climate models is given by Washington and Parkinson (1986). Briefly, the existing models range from single-layer thermodynamic energy balances (Semtner, 1976) to rather detailed multi-layer (including snow) models which include the effects of wind and currents on sea ice dynamics (Maykut and Untersteiner, 1971; Parkinson and Washington, 1979). These models all predict seasonal variations of sea ice thickness, area and volume over specified grids under the influence of specified heat fluxes and fluid stresses. However, it is the sea ice area which is best constrained observationally.

The advent of the satellite era has markedly improved our knowledge of sea ice area changes during seasonal solar heating cycles. For example, Untersteiner (1984) believes the best data set for Antarctic sea ice since 1972 is from passive microwave satellite images—primarily the ESMR instrument on NIMBUS 7. The transition from ice-covered to ice-free water in this data is relatively sharp, albeit still somewhat ambiguous, with concentrations $\approx$ 15% normally cited as the ice edge.

Existing sea ice models can generally be tuned to more or less recover the observed area changes, although many cycles are usually needed to insure seasonal repeatability.

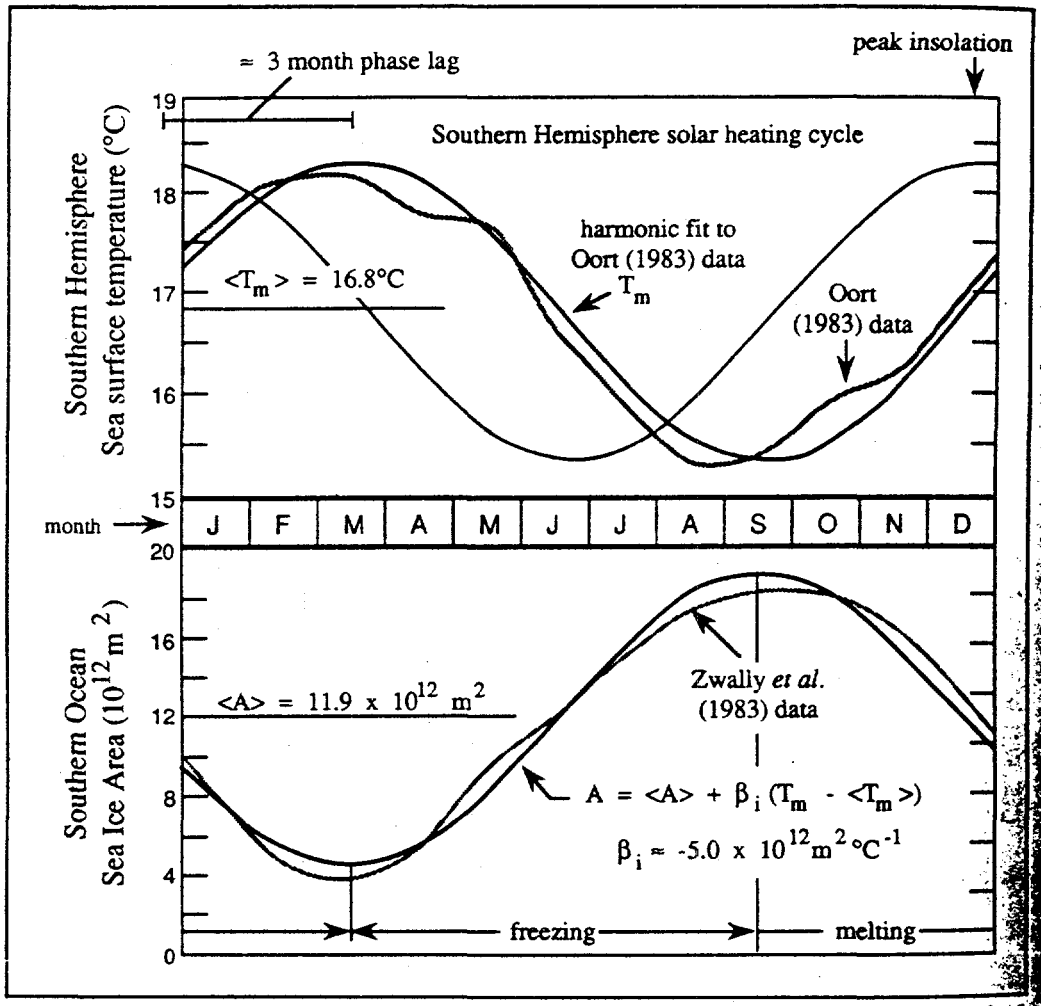


Figure 3. Seasonal sea ice cycle in the Southern hemisphere. (Top) climatological southern hemisphere average of sea surface temperature versus time of year (Oort, 1983), and a harmonic fit to this data of amplitude 1.5°C, are compared with the southern hemisphere insolation cycle approximated for circularized orbit. (Bottom) Climatological satellite-derived sea ice area data (Bottom) Climatological satellite-derived sea ice area data (Zwally *et al.*, 1983) versus time of year are shown along with a linear anticorrelation for sea ice area.

To determine whether the sea ice area cycle is empirically correlated with some simple measure of the annual cycle, I have compared the climatological mean seasonal cycle of sea surface temperature averaged over the Southern Hemisphere (Oort, 1983, Table 25a) with the climatological mean of the seasonal sea ice area cycle derived from the NIMBUS 7 data of Zwally *et al.*, (1983). The result is shown in Fig. 3. The Oort southern hemisphere sea surface temperature data vary nearly harmonically with time over the annual cycle, and can be fit to a good approximation by a sine wave of amplitude 1.5°C

phase-lagged by slightly under three months relative to the southern hemisphere mean solar heating cycle (which peaks in late December). It can be shown that this surface thermal response of the hemisphere as a whole is entirely consistent with the thermal inertia of an oceanic mixed layer some 75–100 m thick.

The amplitude of seasonal temperature variations at latitudes poleward of about 30°S is larger than 1.5°C, but has approximately the same phase lag because the maximum lag possible (for a slab of infinite thermal inertia subject to periodic thermal forcing) is 1/4 year. If the sea ice boundary is constrained by the seawater freezing isotherm, (say) −1.8°C, we might expect the sea ice area also to vary harmonically with a 3-month phase lag, at least so long as the ice extends beyond the physical boundary of Antarctica. As shown in the lower panel of Fig. 3, this is very nearly what is observed. As sea surface temperature decreases from its March peak to its September minimum, sea ice area correspondingly increases by freezing from its March minimum to its September maximum. The result is that sea ice area is "instantaneously" anticorrelated with the southern hemisphere mean sea surface temperature, such that a 1°C cooling can be considered to freeze $\approx 5 \times 10^{12}$ m^2 (5 million square kilometers) of sea ice. Accordingly, $\approx 15 \times 10^{12}$ m^2 of sea ice is produced over the March-to-September freezing part of the cycle in the Southern Hemisphere as the mean oceanic surface temperature cools a total of ≈ 3.0°C.

There seems to be considerable uncertainty as to the thickness of the sea ice packs. Numerical models indicate annual mean values in the range of ≈ 1–3 m (Washington and Parkingson, 1986). Recognizing that the mean sea ice thickness is not very well constrained observationally (it varies with season as well as location), we estimate the annual volumetric increase of Antarctic sea ice in the range

$$Q_i' \approx \frac{(15 \times 10^{12}\ \mathrm{m^2\ yr^{-1}}) \times (2 \pm 1\mathrm{m})}{3 \times 10^7\ \mathrm{syr^{-1}}} \approx (1.0 \pm 0.5) \times 10^6\ \mathrm{m^3\ s^{-1}} = 1.0 \pm 0.5\ \mathrm{Sv} \quad .$$

An interesting problem which is addressed in the next section is how one Sverdrup of pure water freeze-distilled from the salty surface ocean can produce on the order of 30–40 Sverdrups of dense bottom water.

In light of the central role played by the annual freezing rate in the bottom-water theory to be developed, we also want an estimate of how Q_i' can change as a function of global climate. This can be done in principle using a realistic coupled atmosphere-ocean-sea ice model. Since one was not available (and may not exist) I took the empirical approach of assuming that the sea ice area $A_i(T_m)$ varies linearly during the seasonal cycle according to the expression in the lower panel of Fig. 3, so long as a positive area is computed; however, if the hemispheric mean temperature T_m becomes sufficiently high to produce a negative area (sea ice boundary inside continental Antarctica), then $A_i(T_m) = 0$. The ratio of volumetric sea ice changes during an annual cycle relative to that for today's annual cycle (denoted by $_o$) is then approximated by,

$$\frac{Q_i'}{Q_{io}'} = \frac{A_i(< T > -\Delta T) - A_i(< T > +\Delta T)}{A_i(T_o - \Delta T) - A_i(T_o + \Delta T)} \quad ,$$

where $< T >$ is the annual mean hemispheric temperature, T_o is the present hemispheric mean temperature, and ΔT is the amplitude of the hemispheric mean seasonal temperature cycle.

This empirical approach predicts that the amount of sea ice frozen during the annual cycle remains constant so long as the amplitude of the seasonal temperature excursion remains constant and the mean water temperature is sufficiently cold to allow sea ice

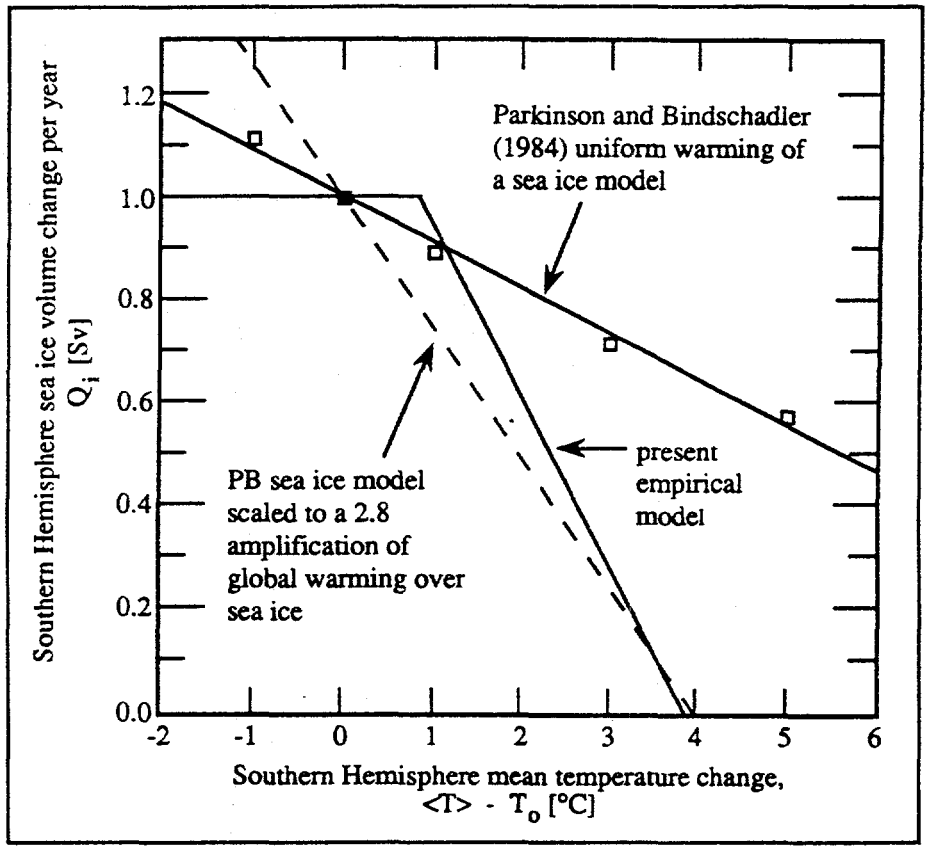


Figure 4. Amplitude of annual sea ice volume change versus change in southern hemisphere mean temperature. An empirical model based on the annual sea ice cycle is compared with the sea ice model predictions of Parkinson and Bindschadler (1984). In the case of global warming, a 2.8 polar amplification of hemispheric mean temperature brings the Parkinson-Bindschadler curve close to the empirical model (see text). All cases are normalized to an annual sea ice freezing rate of 1.0 Sv.

off Antarctica during the warmest month (March). However, when the warming becomes sufficient to make coastal waters ice-free during part of the year ($< T > -T_o \geq 1°C$), the amount of ice formed decreases linearly until the Southern Ocean becomes entirely ice-free the year round ($< T > -T_o \geq 4°C$). This is the empirical model curve of Fig.4.

At first glance the sensitivity of sea ice to temperature change predicted empirically looks much greater than that calculated by Parkinson and Bindschadler (1984) from a detailed sea ice model driven by perturbed climatological air temperatures, the results of which are also shown in Fig. 4. The Parkinson-Bindschadler (PB) calculations were driven by uniform temperature changes superimposed on present climatology over sea ice. But it seems highly unlikely that high-latitude warming will equal hemispheric-mean warming. For example, current "equilibrium" climate models predict high-latitude warming is amplified by a factor of 2–3 relative to the global-mean warming (Grotch,

1988), which would bring the PB curve more in line with the empirical one. Figure 4 shows this is the case when the PB temperature change is amplified by a factor of 2.8. Note that polar amplification may be reduced or absent for a cooling, and that different physics may need to be invoked to analyze that condition. We will refer to the PB uniform temperature perturbation case later as a reference (if unrealistic) scenario for very weak sea ice sensitivity.

5. A SCHEMATIC MODEL FOR ANTARCTIC BOTTOM WATER

Apart from the question of the amount of sea ice frozen (and brine rejected) during the Austral winter, a central issue for climate modelling is the link between the rate of freezing in the Southern Ocean and the flow of dense, cold, saline water to the bottom. Although a number of observations (Gill, 1973; Foster and Middleton, 1980; Lennon *et al.*, 1987) and some prior models (Killworth, 1973; Melling and Lewis, 1982) indicate that turbulent Ekman and gravity layers along the sloping sea floor are favored routes for transporting this dense water to the bottom of the world's oceans, the precise role of brine rejection in driving these flows remains obscure. It is therefore not too surprising that these bottom water processes have not yet been incorporated in global numerical models of ocean circulation (Schlesinger and Jiang, 1988).

Under normal conditions the oceans beneath the seasonal mixed layer are stably stratified ($\partial\rho/\partial z > 0$). As discussed previously, the usual method for forming bottom water in ocean general circulation models (OGCMs) is by so-called convective adjustment. The "overturning-water-column" (OWC) paradigm for bottom water formation is pervasive in the GCM community, and is the usual conceptual model offered for induction of deep convection at all ocean latitudes, including that induced around Antarctica by the rejection of salt during sea ice formation (cf., Fig. 2.22, p. 56 of Washington and Parkinson, 1986). Huang and Bryan (1987) recently emphasized the importance of doing convective adjustment in a manner which conserves both total volume and buoyancy. But the more fundamental question in the present context is whether high-latitude deep convection occurs in narrow boundary layers too small to be explicitly resolved by the OGCMs' grid volumes or uniformly over grids hundreds of kilometers square, as in the OWC model. If convective penetration to abyssal regions is confined to narrow isolated zones, its realistic depiction by OGCMs will require either a significant increase in horizontal and vertical resolution to capture these zones or the parameterization of these effects based on turbulent convection in negatively buoyant gravity layers.

I want now to propose an idealized model for Antarctic Bottom Water formation based on the idea that density enhancement by brine-rejecting sea ice causes bottom water to sink along the Antarctica continental shelf and slope in thin benthic boundary layers of the order of a hundred meters in thickness (Fig. 5). This is a markedly different approach from the bulk column overturnings used in coarse-resolution OGCMs. Benthic gravity layers with locally high density have been observed in oceanographic field studies as well as in the laboratory, and can be clearly seen, for example, in hydrographic data sections in the Weddell Sea off the Antarctica continent (Foster and Middleton, 1980, Fig. 2). The variables and parameters of our model are summarized in Table 1.

As described below, the bottom-water forming process is modelled as the result of two interactive processes: (1) the dilution of rejected salt (and its density anomaly) by bottom water and (2) the sinking of bottom water along benthic boundary layers driven by a sea-ice-freezing-induced density anomaly.

Table 1. Variables and parameters of bottom water model

Variables

v_b = velocity in benthic boundary layer [m s^{-1}]
u^* = $(\tau_b/\rho_b)^{1/2}$ = friction velocity of benthic layer [m s^{-1}]
ρ_s = Antarctic surface water potential density [kg m^{-3}]
ρ_b = Antarctic bottom water potential density [kg m^{-3}]
S_s = Antarctic surface water salinity [‰]
S_b = Antarctic bottom water salinity [‰]
T_s = Antarctic surface water potential temperature [°C]
T_b = Antarctic bottom water potential temperature [°C]
τ_b = turbulent shear stress of benthic layer [N m^{-2}]
n = co-ordinate normal to bottom topography [m]
h = bottom boundary layer thickness [m]
Q_i' = Annual volume of sea ice frozen in Southern Ocean [m^3 s^{-1}]
Q_i = $(\rho_i/\rho_b)Q_i'$ = density-corrected sea ice freezing rate [m^3 s^{-1}]
Q_s = surface inflow rate to sea ice margin [m^3 s^{-1}]
Q_b = $v_b hc$ = bottom water flow rate [m^3 s^{-1}]

Parameters

β = $[(\rho_b-\rho_s)/\rho_s]/[(S_b-S_s)/S_s]$ = constant in state equation ≈ 0.0249
a = earth radius $\approx 6.37 \times 10^6$ m
ϕ_s = latitude of Antarctica shelf break $\approx -76°$
c = $2\pi a\cos(\phi_s)$ = circumference around shelf break $\approx 9.68 \times 10^6$ m
κ = von Kármán's constant ≈ 0.40
c_d = $\tau_b/(\rho_b v_b^2) = (u^*/v_b)^2$ = benthic layer drag coefficient $\approx \kappa^2 \approx 0.16$
f = Coriolis parameter $\approx 1.4 \times 10^{-4}$ radians s^{-1}
g = gravitational acceleration ≈ 9.8 m s^{-2}
θ = Antarctica continental slope ≈ 1 to 2×10^{-3} radians
A_o = surface area of world's oceans $\approx 3.34 \times 10^{14}$ m^2
ρ_i = Antarctic sea ice density ≈ 900 kg m^{-3}

5.1. Dilution Effects

Referring to Fig. 5, we assume first that the volumetric flow rates of water and salt between the surface inflow, sea ice freezing and the induced bottom water flow rate are conserved, and are in a long-term steady state. While Q_i and Q_b are produced only during the March-to-September freezing cycle, their values are presumed to be averages over the annual freezing cycle in the steady-state model. Since interannual changes in annual mean sea ice volume are slow compared to seasonal variations, the sea ice volume melted in Austral spring/summer is approximately equal to that frozen in the fall/winter seasons (Fig. 3); but melting does not induce a sinking flow since surface density and salinity decrease relative to their annual means, ρ_s and S_s. No sinking occurs during the September-to-March melting.

Under these long-term steady state conditions, conservation of water mass between the surface inflow, sea ice formation and the induced bottom water outflow implies

$$\rho_s Q_s = \rho_i Q_i' + \rho_b Q_b \quad .$$

The freeze-distillation of seawater to water ice is normal accompanied by trapped brine bubbles and subsequent crystallization of sea ice as an ice/salt solution of solid phases

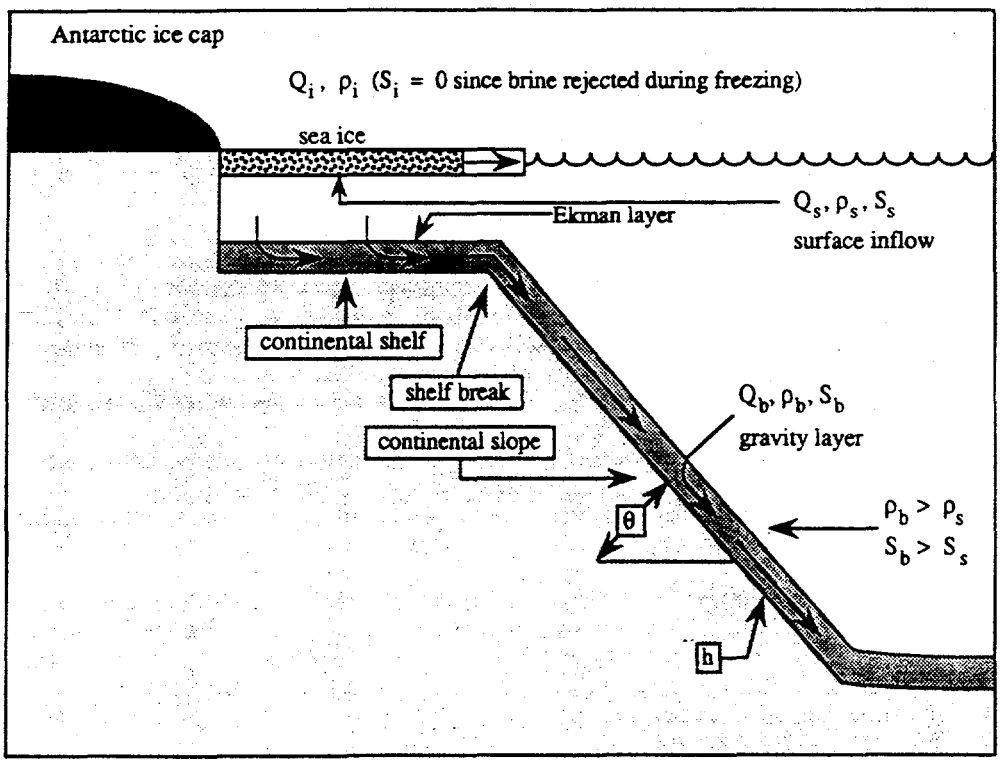


Figure 5. Schematic diagram of the formation of bottom water around the Antarctic continent. The present model assumes high density water formed by brine rejection is entrained by an Ekman boundary layer on the continental shelf, and subsequently flows down the continental slope in a gravity layer.

(eutectic). The salinity of sea ice is nonetheless much less than that of the source seawater, in the range of 15–30%, depending on the air temperature at which freezing occurs (Neumann and Pierson, 1966, pp. 82–83). Thus 70–85% of the salt present in seawater is rejected when sea ice freezes. This is accounted for by writing the salt flux to sea ice as $\rho_i \gamma S_S Q_i'$, where $\gamma = S_i/S_S (\approx 0.15\text{–}0.30)$ is the ratio of the salinity of sea ice to that of the surface water. The balance between the salt influx in surface water and its outflux to sea ice and sinking bottom water is then

$$\rho_S S_S Q_S = \rho_i \gamma S_S Q_i' + \rho_b S_b Q_b \quad .$$

It is convenient to introduce the density-corrected sea ice freezing rate,

$$Q_i = \frac{\rho_i (1-\gamma)}{\rho_b} Q_i' \approx 0.7 Q_i' \quad ,$$

and the ρ–S state equation for the potential density increase of bottom water at sea ice latitudes from freezing-induced brine intensification of seawater,

$$\frac{\rho_b - \rho_s}{\rho_s} = \beta \left[\frac{S_b - S_S}{S_S} \right] \quad .$$

Combining these four equations then gives a simple expression for the difference between the potential density of the bottom water entrained by the Ekman and gravity layers and that of the surface inflow ($\Delta\rho = \rho_b - \rho_s$) in terms of the reduced freezing rate and the bottom water formation rate,

$$\frac{\Delta\rho}{\rho_b} \approx \frac{\Delta\rho}{\rho_S} \approx \frac{\beta Q_i}{Q_b} \quad . \tag{1}$$

Admittedly, this equation must be regarded as very approximate at this time insofar as it neglects the possibility of further entrainment (and density dilution) along the continental slope, but nonetheless it is quite useful for illustrating certain critical points.

Perhaps the most obvious question we can ask of Eq. (1) is whether it is consistent in an order of magnitude sense with the Antarctic Bottom Water density enhancement, bottom water flow rates and sea ice freezing rates in the present world ocean.

Typical salinities and potential temperatures of Antarctic waters are (Levitus, 1982; cf., our Fig. 2): S_S = 33.9 ‰, Ts = −1.46°C at the surface and S_b = 34.7 ‰, T_b = 0.31°C at 4 km depth. The corresponding potential densities from the full Bryan and Cox (1972) state equation are ρ_S = 1027.29 kg m^{-3} and ρ_b = 1027.88 kg m^{-3}, giving a typical "observed" density enhancement of $\Delta\rho/\rho_b \approx 0.0006$. A rough estimate of total Antarctic Bottom Water is the flow rate needed to produce the global upwelling of 1-D upwelling-diffusion models which reproduce the observed temperature structure of the world ocean thermocline ($w \approx 4$ m yr$^{-1} \approx 1.27 \times 10^{-7}$ m s^{-1}; Hoffert *et al.*, 1980), that is, $Q_b \approx wA_o \approx 42$ Sv. With $\beta \approx 0.025$ (estimated from the Bryan-Cox state equation for Antarctic conditions), Eq. (1) requires a corrected sea ice freezing rate of $Q_i = (\Delta\rho/\rho_b)Q_b/\beta \approx 1.0$ Sv. This is close to the present volumetric change of sea ice during the annual cycle estimated earlier, even accounting for the lower ice density which gives the uncorrected freezing rate as $Q_i' \approx 1.0/0.9 \approx 1.1$ Sv.

I want to emphasize that for any given freezing rate, Eq. (1) represents the degree to which the bottom water density is diluted by the flushing effect of bottom water flow. It predicts that at constant Q_i, the density of bottom water *decreases* with *increasing* Q_b, as shown by the family of hyperbolic curves in Fig. 6. This seems surprizing at first glance because one intuitively expects the density anomaly to increase as bottom water is more vigorously pumped. It will be shown shortly that bottom density does indeed increase with Q_b because Q_i and Q_b are not independent. To determine where the bottom water flow rate and density enhancement points lie along each of these curves, an additional constraint associated with the flow dynamics is needed. The closure of this problem based on a consideration of the flow dynamics of the Ekman and gravity benthic boundary layer shown in Fig. 5 is developed next.

5.2. Boundary Layer Effects

Our model allows three kinds of force per unit volume to act along the benthic boundary layer: (1) a turbulent ("viscous drag") force, $F_t = \partial\tau_b/\partial n \approx (\rho_b v_b^2 C_d)/h \approx (\rho_b C_d Q_b^2)/(c^2 h^3)$, (2) a Coriolis ("Ekman") force, $F_c = \rho_b v_b f \approx (\rho_b f Q_b)/(ch)$, and (3) a negative buoyancy ("gravity") force, $F_g = \rho_b - \rho_S)g \sin\theta \approx \Delta\rho g\theta$. The regions in which these forces act are determined by the subsurface topography off the Antarctic continent as depicted schematically in Fig. 5.

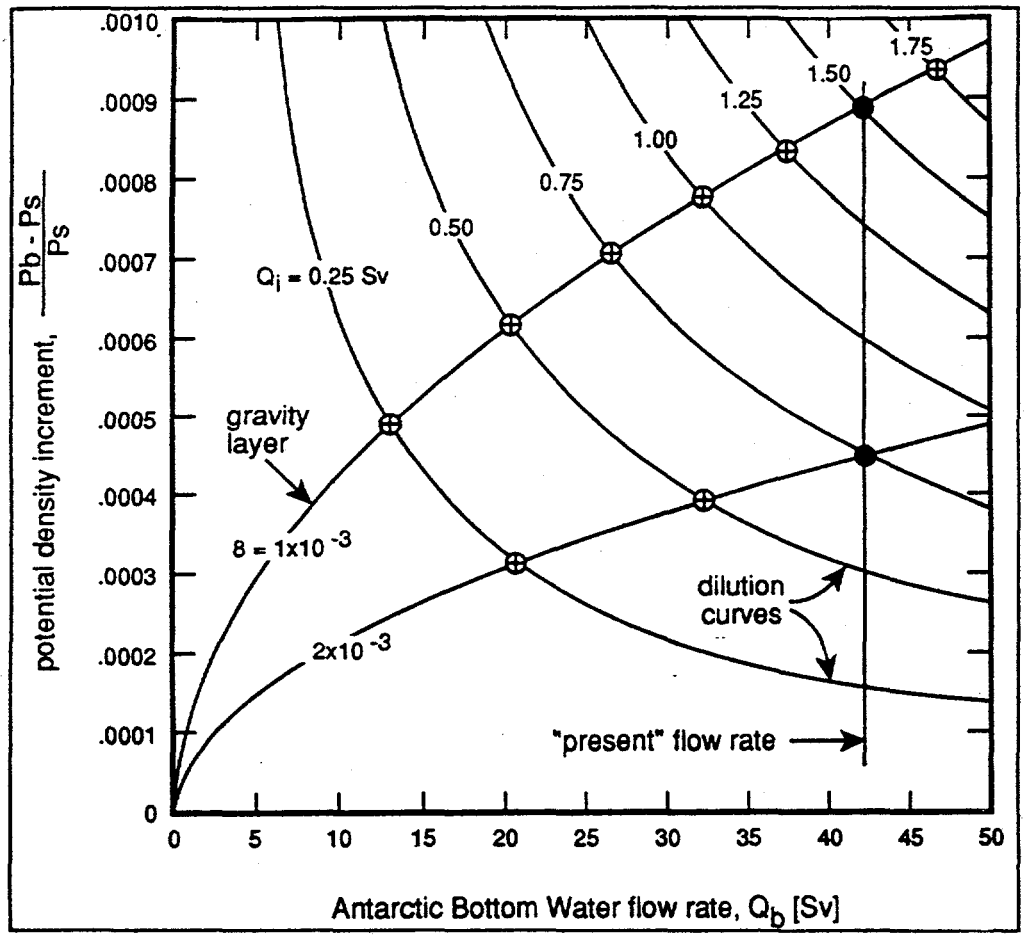


Figure 6. Bottom water potential density increment normalized to surface density versus the flow rate of Antarctic Bottom Water from the present model. The dilution curves are shown for several sea ice freezing rates from Eq. (1); the gravity layer curves for two continental slopes from Eq. (2). Intersections of dilution curves with gravity layer curves define Antarctic bottom water densities and flow rates as a function of sloppe angles and freezing rates (crossed dots). Q_b scales with the $q \times Q_i$ product such that the two intersections denoted by plack dots both give the present flow rate. But the density increment pumped downward is different, with the two cases bracketing the observed value ≈ 0.0006. The implied annual freezing rate of sea ice is in the range 0.75–1.50 Sv, which also tallys with observations.

It can be shown for typical Antarctic $\Delta\rho$'s that on the relatively flat continental shelf ($\theta << 10^{-4}$ radians) the viscous drag along the bottom is balanced mainly by Coriolis forces ($F_t \approx F_c$), in which case the benthic layer thickness is given by

$$h_c \approx \left\{ \frac{C_d Q_b}{fc} \right\}^{1/2} .$$

Notice that we could have alternately written the Coriolis (Ekman) boundary layer thickness on the continental shelf in terms of the friction velocity (instead of the bottom water flow rate through the benthic layer) as $h_c \approx C_d v_b f^{-1} \approx C_d^{1/2} u^* f^{-1}$. Our choice for the drag coefficient as the square of von Kármán's constant, $C_d \approx \kappa^2$, is then seen to give the classical boundary layer thickness scale for a neutrally stable benthic layer (Turner, 1981), $h_c \approx \kappa u^* f^{-1}$. We assume the drag coefficient remains roughly constant across the shelf break, but recognize that a different force balance applies in the layer after the shelf break transition, namely, that the turbulent drag force along the continental slope is balanced mainly by the component of negative buoyancy acting along the gravity layer ($F_t \approx F_g$) such that

$$h_g \approx \left\{ \frac{\rho_b C_d Q_B^2}{\Delta\rho g \theta c^2} \right\}^{1/3} .$$

Equivalently, the gravity layer thickness in terms of the boundary layer velocity is $h_g \approx (\rho_b C_d v_b^2)/(\Delta\rho g\theta)$. Solving this for the current in the gravity layer gives the so-called Chezy equation, $v_b = [g(\Delta\rho/\rho_b)h_g\theta/C_d]^{1/2}$. The flow of density-driven gravity boundary layers over sloping surfaces is known to be described by the Chezy equation in laboratory experiments (Turner, 1981). This equation was recently found by Lennon *et al.* (1987) to also describe gravity currents produced by evaporative salt enhancement measured in the Spencer Gulf off South Australia. In commenting on Lennon *et al*'s. (1987) findings, Whitehead (1987) observes that these measurements have implications beyond the local offshore oceanography considered, and that they "may even enable us to understand more clearly the situation in the Weddell Sea." We are aiming precisely in that direction, although we also embody the earlier suggestion of Killworth (1973) that the entire coastline of Antarctica, and not just the coastline of the Weddell Sea, should be considered as a source of Antarctic Bottom Water.

If one accepts the ideas that the Ekman layer on the continental shelf is the main entrainment site for high-salinity, dense water produced by the seasonal sea ice freezing cycle, and that the thickness of the layer containing bottom water is continuous across the transition to a gravity layer at the shelf break, it follows (from $h_c = h_g$ = constant) that the bottom water density anomaly as a function of its flow rate is

$$\frac{\Delta\rho}{\rho_b} \approx \frac{f^{3/2}}{g\theta C_d^{1/2}} \left\{ \frac{Q_b}{c} \right\}^{1/2} . \qquad (2)$$

This expression for bottom density provides "closure" of the bottom water model insofar as it can be combined with the dilution Eq. (1) to set the density and flow rate of Antarctic Bottom Water as a function of the sea ice freezing rate. For example, Eq. (2) is plotted in Fig. 6 as the "gravity layer" curve along with the dilution curves of Eq. (1) discussed earlier. The intersection of these curves in density-bottom water space implicitly defines the bottom water flow rate as a function of the sea ice freezing rate. It is also possible to eliminate explicitly the bottom density between Eqs. (1) and (2), and thereby get the bottom water flow rate, the bottom density anomaly, and the boundary layer thickness as analytic functions of Q_i and the specified model parameters:

$$Q_b(Q_i) \approx \frac{(C_d c)^{1/3}(\beta g \theta Q_i)^{2/3}}{f} , \qquad (3)$$

$$\Delta\rho(Q_i) \approx \frac{\rho_b f}{(g\theta)^{2/3}} \left\{ \frac{\beta Q_i}{C_d c} \right\}^{1/3} , \quad (4)$$

$$h(Q_i) \approx \frac{C_d^{2/3}}{f} \left\{ \frac{\beta g \theta Q_i}{c} \right\}^{1/3} . \quad (5)$$

These results indicate that bottom water flow rate scales with the two-thirds power of the freezing rate, whereas the bottom density anomaly and boundary layer thickness scale with the one-third power. With an annual freezing rate of $A_{bo} = 1.0$ Sv and $\theta = 2 \times 10^{-3}$, Eqs. (3)–(5) give the following reference values: $Q_b \approx 42$ Sv, $\Delta\rho_o/\rho_b \approx 0.00044$ and $h_o \approx$ 70m, all of which are in reasonably good agreement with observations. The power-law scaling can also be used to estimate the effects of different freezing rates. For example, $Q_{bo} = 0.5$ Sv gives a bottom water flow rate of $Q_{bo} \approx 32$ Sv, which is more in accord with a model in which (say) 10 Sv of bottom water is formed from North Atlantic Deep Water.

Some important points which emerge from this analysis are that the Antarctic Bottom Water flow rate induced by the seasonal sea ice cycle is of the order of thirty times the sea ice freezing rate, and that the benthic boundary layer—a weak (1/3 power) function of the freezing rate—is of the order of a hundred meters thick. Although the model predicts that the negative buoyancy flux of AABW is determined by the sea ice freezing rate from simple dilution considerations [$\Delta\rho Q_b \approx \rho_S \beta Q_i$ from Eq. (1)], the determination of its flow rate required a consideration of benthic boundary layer dynamics—a feature not normally modelled by OGCMs. Our picture of "thin" high-density layers along the Antarctic continental slope feeding bottom water to the floor of all ocean basins thus underscores the probable inability of coarse vertical resolution OGCMs to resolve explicitly these processes. This factor in combination with the lack of seasonal sea ice and salinity-rejection cycles may explain why present-day ocean circulation models do not produce realistic deep upwelling rates. The implications of these processes for climatic transitions involving the possible shutoff of Antarctic Bottom Water are discussed in the following section.

6. THE CO_2 GREENHOUSE AND CHANGES IN BOTTOM WATER CIRCULATION

How vulnerable is the circulation of cold, deep water to the sea floor to global greenhouse warming of the surface in light of the foregoing analyses?

We saw in Fig. 1 that the paleoceanographic record contains evidence of progressive cooling of abyssal waters as the planet entered the global ice ages. Conversely, a warmer climate could render the high-latitude oceans ice-free, and thus shut down bottom water formation, and consequently the global upwelling of the thermohaline circulation. As discussed earlier, AABW formation results not from the direct effect of cold surface temperatures at high latitudes but, instead, from the indirect effect of the seasonal sea ice salinity pump. According to the picture developed here, the AABW pump acts as a half-wave rectifier allowing dense water formed in Austral winter by salt rejection to sink gravitationally along the slope boundary layer, while the light water formed during melting remains at the surface. In a fundamental sense we can say that AABW is a function of the configuration of the continents and oceans, which determines whether paths exist from the slopes of polar continents to the bottom of the world's oceans; on the earth's spin axis tilt (obliquity), which creates the seasonal cycle; and on the global mean temperature, which determines the extent of high-latitude sea ice frozen during the seasonal cycle.

Although continental geography and planetary tilt will almost certainly remain constant over the 10–100 years timeframe of anthropogenic climate change, the effects of anthropogenic greenhouse gases on global temperatures can be very significant. Current equilibrium GCM predictions are in the range of a 2.8–5.2°C global mean warming for a CO_2 doubling alone (see below), with perhaps an additional 50% warming from chlorofluorocarbons (CFCs), CH_4, N_2O, and other trace gases. Figure 7, based on our bottom water model, indicates that for our empirical sea ice sensitivity, a global mean warming of this magnitude could not only render the Southern Ocean ice-free the year round, but could effectively shut down the seasonal sea ice pump driving AABW and the thermohaline circulation. Implicit in this conclusion is the phenomenon of high southern latitude amplification of the warming, as the uniform-with-latitude warming of the Parkingson-Bindschadler (PB) curves show a much weaker response of bottom water to global warming. (The effect of a global cooling is another story, and will not be pursued here.)

Figure 8 compares the predicted increase in zonal mean surface air temperature versus latitude for a carbon dioxide doubling of four atmospheric general circulation models (GCMs) during the southern hemisphere winter and summer seasons (Grotch, 1988). The intercomparison shows generally similar trends, with differences in detail. The GISS model, for example, predicts twice the tropical warming the other models do. Nevertheless, a distinct high-latitude amplification of the mean global warming over the southern hemisphere sea during the critical Southern Hemisphere winter (JJA; when sea ice freezing occurs) is exhibited by all models. At the latitudes of Southern Ocean sea ice, surface temperature rises are in the range of 6–12°C—some three times the global mean warming. A corresponding correction for high-latitude amplification of the global warming would increase the negative slope of the PB sea ice-temperature curve of Fig. 7 by comparable factors, and make it nearly the same as our empirical sensitivity curves. The point is that, to the extent these GCM results realistically represent the surface temperature distribution toward which the planet is headed, they imply that a greenhouse-triggered cutoff of AABW is possible, even likely, sometime in the next century.

The scenario of transient climate change during and immediately following such a cutoff of AABW is extremely interesting, and needs to be explored in more detail. For one thing, it is important to determine how heat is taken up by the oceans on a latitude-by-latitude basis to estimate the time scales over which the high latitudes will warm. This will require more research into the relative roles of saline and thermal convection. The nearly isothermal (and hence less stable) nature of high-latitude ocean columns suggests heat will be mixed to the bottom more rapidly than at middle and tropical latitudes. (Enhanced mixing at high latitudes is already observed for geochemical tracers.) Because vertical mixing of heat brings into play a deeper column of fluid, it increases the local thermal inertia and could actually retard high-latitude surface warming. A high latitude thermal lag is consistent with the fact that Zwally *et al.* (1983) were not able to to detect a secular decrease in Antarctic sea ice extent since 1973, although the planet as a whole has been warming. Hansen and Lebedeff (1988) find the rate of warming of the earth's surface over the past hundred years ($\approx$ 0.5°C per century) is accelerating, with an increase in global mean temperature of $\approx$ 0.3°C from 1975 to 1987 ($\approx$ 2.5°C per century)—perhaps signalling the emergence of the global greenhouse. If this warming continues to accelerate into the next century, the establishment of a new global climate along the lines predicted in Fig. 8 is a distinct possibility, although the transient may be such that high southern latitudes will be the last to feel it. Indeed, Hansen and Lebedeff observe that the recent warming has been most pronounced in the tropics, and speculate that this is because low latitudes are where the ocean mixed layer is thin and exchange

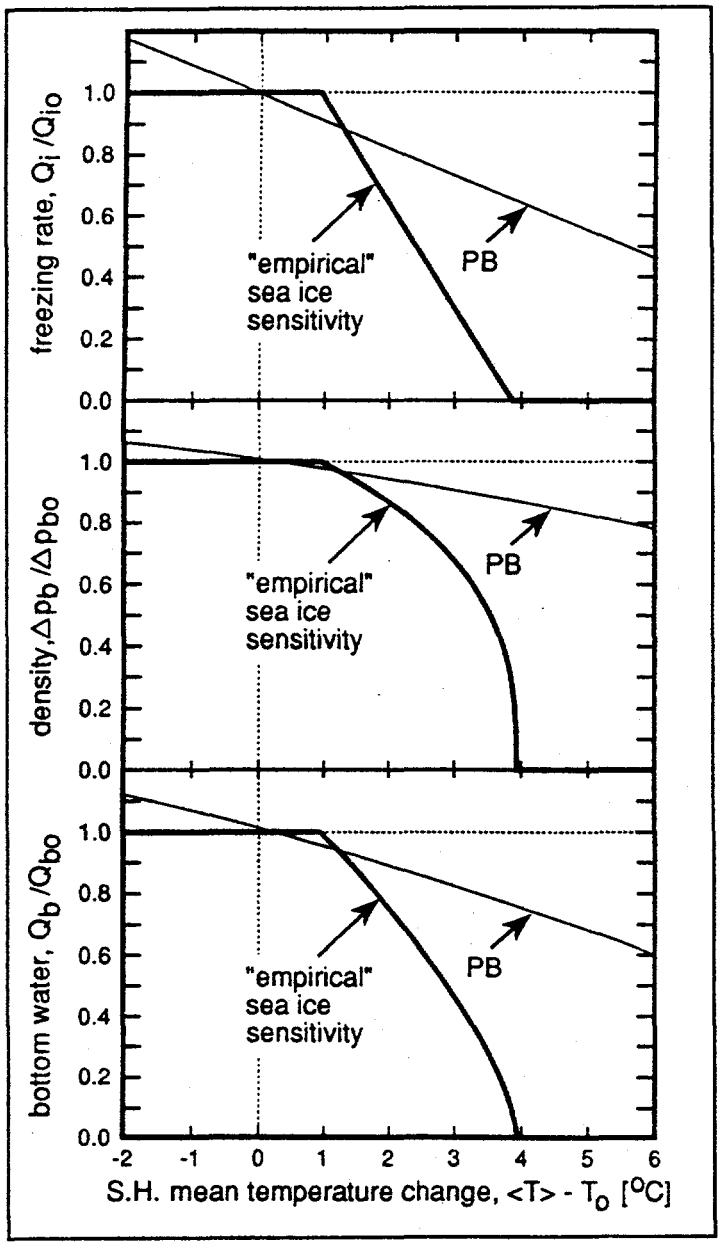


Figure 7. Variation of normalized Antarctic sea ice freezing rate (Top), bottom-to-top potential density difference (Middle), and bottom water flow rate (Bottom), versus Southern Hemisphere mean temperature change from present bottom water model. Results are shown for both the empirical and Parkinson-Bindschadler (PB) sea ice sensitivity.

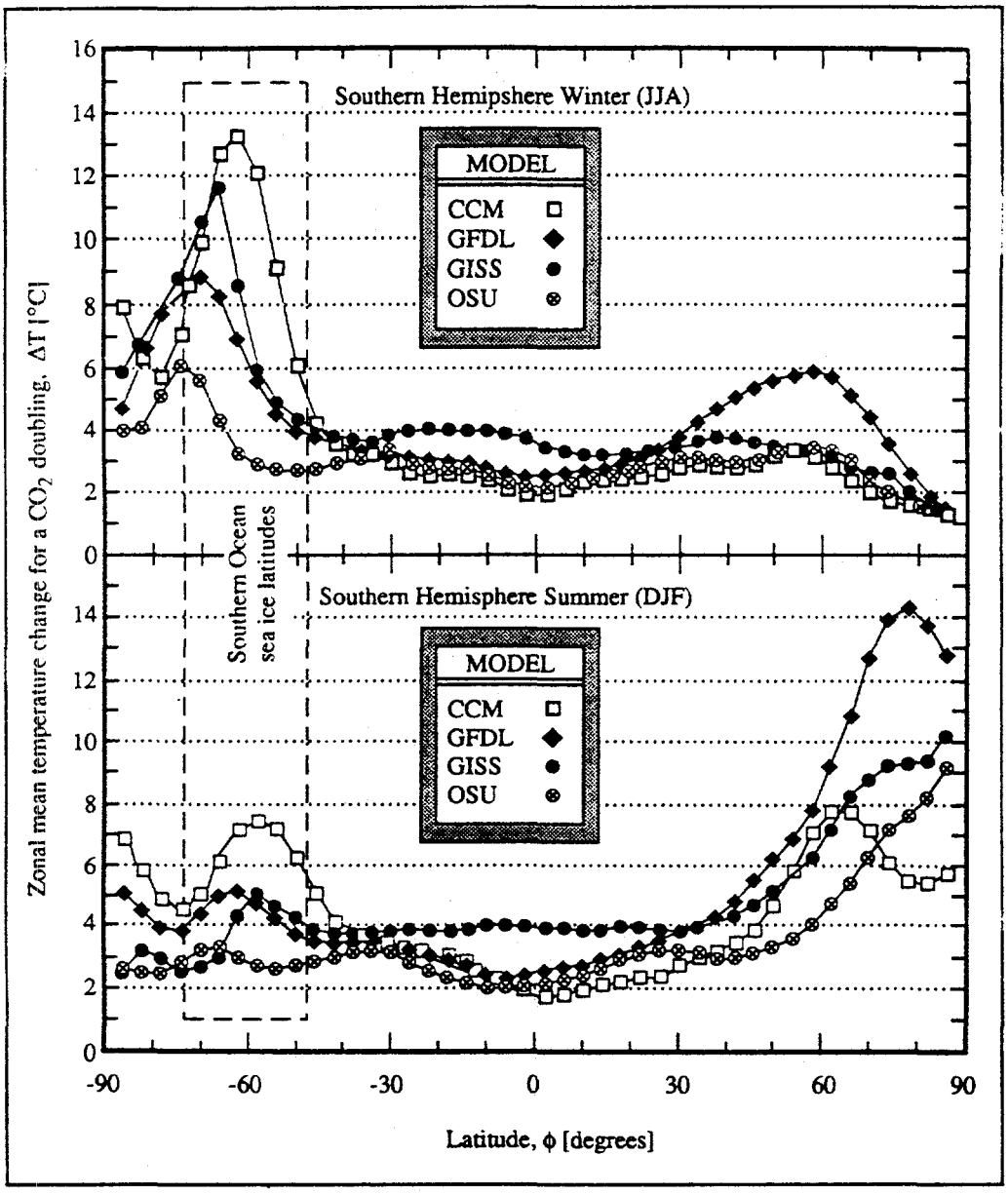


Figure 8. Pole-to-pole distribution of the long-term equilibrium zonal mean surface air temperature change from a CO_2 doubling as computer by four atmospheric GCMs (Grotch, 1988): CCM = Community Climate Model of the National Center for Atmospheric Research; GFDL = Geophysical Fluid Dynamics Laboratory Modsl; GISS = NASA/Goddard Institute for Space Studies Model; OSU = Oregon State University Model.

with the deeper ocean is inhibited by stable stratification. A high-latitude thermal lag would delay in time, but not avoid, the shut down of the Antarctic sea ice pump. If the high southern latitudes surface waters approach the greenhouse warming predicted by GCMs sometime in the next century, the effects could be both profound and "rapid."

In a recent study using a simple upwelling-diffusion transient climate model, Watts and Morantine (1988) found a sudden shutoff of thermohaline upwelling would paradoxically trigger a transient global cooling, because warming the deep sea by vertical eddy diffusion only would require the import of heat from the surface to conserve energy. Eventually the surface would return to the predicted steady state on the 1000-year time scale of ocean mixing. But the cooling transient would be more rapid (on decadal to century time scales). In light of the sea ice-global temperature link, a surface cooling in the wake of a bottom water shutoff from a surface warming could produce even more complex feedbacks. If the re-cooling were sufficient to reintroduce sea ice, bottom water might reform, leading to another warming, and so on. The net effect could be a damped oscillatory approach to the long-term steady state, or even an unanticipated nonlinear transient behavior.

These very preliminary considerations suggest the rich variety of ocean physics that can come into play during climatic transition involving changes in world ocean bottom water. Their implications for the future environmental state of the earth should be more than sufficient motivation to study them. The development of better models for deepwater formation from saline sources at the surface, perhaps along the lines sketched out here, seems a good starting point.

ACKNOWLEDGMENTS

I thank Stan Grotch of Lawrence Livermore National Laboratory for suppling the GCM intercomparison data of Fig. 8, and Mike Schlesinger and Tyler Volk for helpful discussions. This work was supported by the Carbon Dioxide Research Division, Office of Basic Energy Sciences, of the United States Department of Energy under contract DE-FG02-85ER60350 to New York University.

REFERENCES

Berger, W. H., 1981: Paleoceanography: The deep-sea record. In *The Oceanic Lithosphere; The Sea, Vol. 7*, C. Emiliani (ed.), Wiley-Interscience, NY, pp. 1437–1519.

Broecker, W. S., 1987: Unpleasant surprises in the greenhouse? *Nature*, **328**, 123–126.

Bryan, K., and M. D. Cox, 1972: An approximate equation of state for numerical models of ocean circulation. *J. Phys. Oceanogr.*, **2**, 510–514.

Bryan, K., and M. J. Spelman, 1985: The ocean's response to a CO_2-induced warming. *J. Geophys. Res.*, **90**, 11,679–11,688.

Foster, T. D., and J. H. Middleton, 1980: Bottom water formation in the western Weddell Sea. *Deep-Sea Res.*, **27A**, 367–381.

Gill, A. E., 1973: Circulation and bottom water production in the Weddell Sea. *Deep-Sea Res.*, **20**, 111–140.

Gordon, A. L., 1982: Weddell Deep Water variability. *J. Mar. Res.*, **40**, suppl., 199–217.

Grotch, S. L., 1988: Regional Intercomparisons of General Circulation Model Predictions and Historical Climate Data, Report DOE/NBB-0084, Atmospheric and Geophysical Sciences Division, Lawrence Livermore National Laboratory, Livermore, CA, 291pp.

Häkkinen, S., 1987: A coupled dynamic-thermodynamic model of an ice-ocean system. *J. Geophys. Res.*, **92**, 9469–9477.

Hansen, J., and S. Lebedeff, 1988: Global surface air temperatures: Update through 1987. *Geophys. Res. Lett.*, **15**, 323–326.

Hoffert, M. I., A. J. Callegari and C. T. Hsieh, 1980: The role of deep sea forcing in the secular response to climatic forcing. *J. Geophys. Res.*, **85**, 6667–6679.

Hoffert, M. I., B. P. Flannery, A. J. Callegari, C. T. Hsieh and W. Wiscombe, 1983: Evaporation-limited tropical ocean temperatures as a constraint on climate sensitivity. *J. Atmos. Sci.*, **40**, 1659–1668.

Huang, R. X., and K. Bryan, 1987: A multilayer model of the thermohaline and wind-driven ocean circulation. *J. Phys. Oceanogr.*, **17**, 1909–1924.

Kellogg, T. B., 1987: Glacial-interglacial changes in global deepwater circulation. *Paleoceanography*, **2**, 259–271.

Killworth, P. D., 1973: A two-dimensional model for the formation of Antarctic Bottom water. *Deep-Sea Res.*, **20**, 941–971.

Lennon, G. W., D. G. Bowers, R. A. Nunes, R. D. Scott, M. Ali, Cai Wenju, M. Herzfeld, G. Johansson, S. Nield, P. Petrusevics, P. Stephenson, A. A. Suskin and S. E. A. Wijffels, 1987: Gravity currents and the release of salt from an estuary. *Nature*, **327**, 695–697.

Levitus, S., 1982: Climatological Atlas of the World Ocean. NOAA Professional Paper 13, National Oceanic and Atmospheric Administration, Rockville, MD, 173 pp.

Manabe, S., and K. Bryan, 1985: CO_2 induced changes in a coupled ocean-atmosphere model and its paleoclimatic implications. *J. Geophys. Res.*, **90**, 1689–1707.

Maykut, G. A., and N. Untersteiner, 1971: Some results from a time-dependent thermodynamic model of sea ice. *J. Geophys. Res.*, **76**, 1550–1575.

Melling, H., and E. L. Lewis, 1982: Shelf drainage flows in the Beaufort Sea and their effect on the Arctic Ocean pycnocline. *Deep-Sea Res.*, **29**, 967–985.

Neumann, G., and W. J. Pierson, Jr., 1966: Principles of Physical Oceanography. Prentice-Hall, Inc., Englewood Cliffs, NJ, 545 pp.

Oort, A. H., 1983: Global Atmospheric Circulation Statistics 1958–1973. NOAA Professional Paper 14, National Oceanic and Atmospheric Administration, Rockville, MD, 180 pp. + 47 Microfiches.

Parkinson, C. L., and R. A. Bindschadler, 1984: Response of Antarctic sea ice to uniform atmospheric temperature increases. In *Climate Processes and Climate Sensitivity*, J.E. Hansen and T. Takahashi (eds.), Maurice Ewing Ser. 5, Geophys. Monograph. 29, American Geophysical Union, Washington, DC, pp. 254–264.

Parkinson, C. L., and W. M. Washington, 1979: A large-scale numerical model of sea ice. *J. Geophys. Res.*, **84**, 311–337.

Savin, S. M., 1977: The history of the earth's surface temperature during the past 100 million years. *Ann. Rev. Earth Planet. Sci.*, **5**, 319–355.

Schlesinger, M. E., and X. Jiang, 1988: The transport of CO_2-induced warming into the ocean: An analysis of simulations by the OSU coupled atmosphere-ocean general circulation model. *Clim. Dyn.*, **3**, 1–17.

Schopf, T. J. M., 1980: Paleoceanography. Harvard University Press, Cambridge, MA, 341 pp.

Semtner, A. J., Jr., 1976: A model for the thermodynamic growth of sea ice in numerical investigations of climate. *J. Phys. Oceanogr.*, **6**, 379–389.

Turner, J. S., 1981: Small scale mixing processes. In *Evolution of Physical Oceanography*. B. A. Warren and C. Wunsch (eds.), MIT Press, Cambridge, MA, pp.236–262.

Untersteiner N., 1984: The cryosphere. In *The Global Climate*, J. T. Houghton, (ed.), Cambridge University Press, NY, pp. 121–140.

Warren, B. A., 1981: Deep circulation in the world ocean. In *Evolution of Physical Oceanography*, B. A. Warren and C. Wunsch (eds.), MIT Press, Cambridge, MA, pp.6–41.

Washington, W. M., and C. L. Parkinson, 1986: An Introduction to Three-Dimensional Climate Modeling. Oxford University Press, 422 pp.

Watts, R. G., and M. Morantine, 1988: Rapid climatic change and the deep ocean. Climatic Change (in press).

Weller, G., C. R. Bentley, D. H. Elliot, L. J. Lanzerotti and P. J. Webber, 1987: Laboratory Antarctica: Research contributions to global problems. *Science*, **238**, 1361–1367.

Whitehead, J. A., 1987: Dense water off continents. *Nature*, **327**, 656.

Zwally, H. J., C. L. Parkingson and J. C. Comiso, 1983: Variability of Antarctic sea ice and changes in carbon dioxide. *Science*, **220**, 1005–1012.

M.E. Schlesinger (Editor), Greenhouse-Gas-Induced Climatic Change:
A Critical Appraisal of Simulations and Observations
© 1991 Elsevier Science Publishers B.V., Amsterdam

The Effects of Solar Variability on Climate*

M. I. HOFFERT
Earth Systems Group
Department of Applied Science
New York University
New York, NY 10003
U.S.A.

ABSTRACT. It has been hypothesized for at least a century that some of the observed variance in global temperature records arises from variations in solar output. Theories of solar-variability effects on climate could not be tested directly prior to satellite measurements because uncertainties in ground-based measurements of solar irradiance were larger than the solar variations themselves. Measurements by the Active Cavity Radiometer (ACRIM) onboard the Solar Max satellite and by the Earth Radiation Budget (ERB) instrument onboard Nimbus 6 are now available which indicate solar-constant variations are positively correlated with solar activity over an 11-yr solar cycle, and are of order ± 1.0 W m^{-2} relative to a mean solar constant of S_o = 1367 W m^{-2}, $\Delta S/S_o \approx \pm 0.07\%$. For a typical climate sensitivity parameter of $\beta = S_o \partial T/\partial S \approx 100°C$, the corresponding variations in radiative *equilibrium* temperature at the Earth's surface are $\Delta T_e \approx \pm 0.07°C$. The realized temperature variations from solar forcing, ΔT, can be significantly smaller because of thermal damping by the ocean.

I consider effects of solar variability on the observed and projected history of the global temperature record in light of this data using an upwelling-diffusion ocean model to assess the effect of ocean thermal inertia on the thermal response. The response to harmonic variations of the 11-yr sunspot cycle is of order $\Delta T \approx \pm 0.02°C$, though the coupling between response and forcing is stronger for long-term variations in the envelope of the solar cycle which more nearly match the thermal response time of the deep ocean (e.g., the 80-year Gleissberg cycle). Nonetheless, solar variability effects are estimated to have been small compared with the 0.5°C warming observed over the past century and the increased rates of global warming observed in recent years. It can be concluded from this analysis that variations in solar output are unlikely to significantly alter the warmings projected for the next century from anthropogenic greenhouse gases.

1. INTRODUCTION

Global climatic change is thought to arise primarily from four factors (Hoffert and Flannery, 1985): (i) variations in solar luminosity; (ii) variations in planetary albedo associated

*This report is a revised and condensed version of a paper originally published as Hoffert, M. I. *et al.*, *Climatic Change* **13** (1988), 267-285.

with changing amounts of aerosols or dust, surface reflectivity and cloudiness distributions; (iii) variations in amounts of infrared absorbing gases in the atmosphere (H_2O, CO_2, O_3 and various trace gases); and (iv) internal feedbacks among elements of the climate system. This report will focus on what can be learned from direct satellite measurements of solar irradiance fluctuations from 1980-1984 about the contribution of solar variability to the global surface temperature history of the Earth.

2. GLOBAL TEMPERATURE HISTORIES AND SOLAR VARIABILITY

Considerable effort has been expended in recent years to describe the temperature history of the Earth from instrumental records over the past century. Figure 1, for example, illustrates two recent reconstructions of surface temperature anomalies (relative to the year 1980) extending back in time over a century which have been developed independently by Jones *et al.* (1986) at the University of East Anglia, UK, and Hansen and Lebedeff (1988) at the Goddard Institute for Space Studies in New York. The Jones *et al.* curve is based on area-weighted averages of both land and sea records, with sea surface temperatures from the Comprehensive Ocean-Atmosphere Data Set (COADS) corrected for the transition from the sailing ship "bucket" temperatures to the water inlet temperatures of steamships; 1980-1984 temperatures are from NOAA observations adjusted for compatibility with earlier data. The Hansen and Lebedeff curve is based on slightly different data sources and methodologies, but exhibits essentially the same trends: a systematic global warming of some 0.5°C per century superimposed on considerable variability on interannual to decadal time scales. Both analyses indicate the rate of global warming has increased over the past several years, perhaps signaling the emergence of the fossil fuel greenhouse effect.

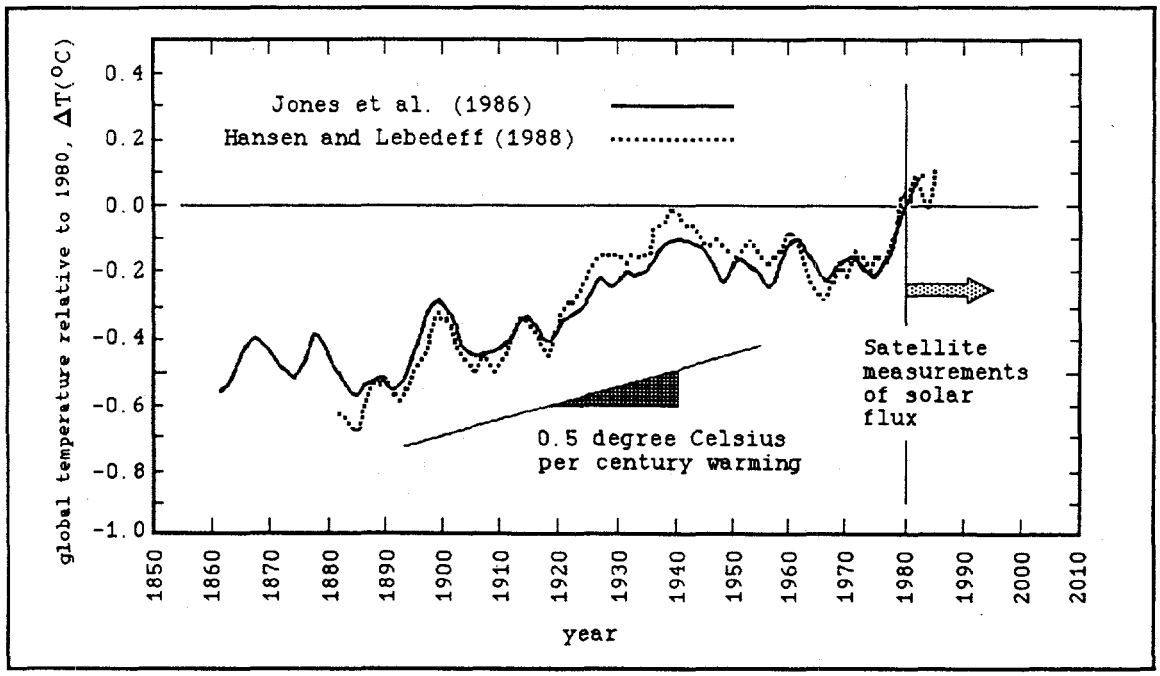


Figure 1. Smoothed global-mean surface air temperatures: 1861-1984.

To establish whether, and how much, of this variation is due to anthropogenic greenhouse gas emissions, it is necessary to know the relative contributions of other factors. As discussed below, space-based irradiance observations are necessary for an accurate assessment of solar variation effects on surface temperature. Figure 1 shows that direct measurements of solar irradiance incident on the Earth from platforms above the atmosphere are very recent in comparison with the instrumental record. The question addressed here is the degree to which one can attribute the historical variance of global-mean temperature to solar variability in light of the relatively small interval of the historical temperature record over which space-based irradiance measurements are available.

A number of modeling studies have appeared in recent years aimed at explaining global temperature records in terms of multiple driving mechanisms (Schneider and Mass, 1975; Robock, 1978, 1979; Vinnikov and Groisman, 1979; Hansen *et al.*, 1981; Gilliland, 1982; Gilliland and Schneider, 1984; Reid, 1987). All of these to some extent allow for an effect from solar luminosity variations using assumed correlations with ground-based observables such as sunspots. Nevertheless, the contribution to the temperature signal from variations in the "solar constant" S_o remains controversial. Despite longstanding proposals that solar variability associated with sunspot cycles has a significant impact on climate over the past 100 years, some recent assessments dispute this strongly (Pittock 1978, 1983; Budyko *et al.*, 1986). Interestingly, the measurements of solar irradiance from satellites which bear directly on this issue have not yet been incorporated in transient climate models.

The main problem with surface-based radiometers is that owing to uncertainties in atmospheric scattering and absorption effects they can measure variations in the solar constant only to an accuracy of 1-2% (Newkirk, 1983). The change in global-mean surface temperature corresponding to a solar irradiance anomaly $\Delta S/S_o$ persisting long enough for the planet to come into a new radiative equilibrium – perhaps decades, as a result of the ocean's thermal inertia – can be estimated from the steady-state climate sensitivity parameter, $\beta \equiv S_o \partial T/\partial S \approx 100°C$ (Hoffert and Flannery, 1985). Thus an uncertainty of $\Delta S/S_o \approx \pm 1\text{-}2\%$ corresponds to an upper bound uncertainty in the effect of solar variations on global temperature anomalies of $\Delta T \approx \beta \Delta S/S_o \approx \pm 1\text{-}2°C$. This is much greater than the variance of the instrumental global temperature record from all physical mechanisms over the past 100 years (Fig. 1). Thus ground-based measurements clearly have inadequate resolution to assess solar-fluctuation effects on climate.

The fault is not with the instruments. Present-day radiometers can measure solar irradiance with a precision of $\pm 0.002\%$ and a long-term accuracy better than 0.1% (Willson, 1984), the range needed for solar variability studies. But to exploit that capability one has to get the detector above the sensible atmosphere, point it in the right direction, and be confident that its calibration will not drift for many years. Certainly, applications to understanding climatic change have been a major motivation for recent solar flux monitoring programs from spacecraft. The past decade has seen the beginnings of extraterrestrial long-term solar monitoring programs including the Earth Radiation Budget (ERB) instruments onboard the NIMBUS 6 and 7 satellites (Hickey *et al.*, 1981), and more recently the Active Cavity Radiometer Irradiance Monitor (ACRIM) onboard the Solar Maximum Mission (SMM) satellite (Willson *et al.*, 1981; Willson and Hudson, 1981; Willson, 1984).

3. SPACE-BASED IRRADIANCE MEASUREMENTS AND SOLAR ACTIVITY

Apart from short-duration NASA sounding rocket and Skylab experiments (Eddy, 1979), continuous data sets of solar flux measured from above the atmosphere are mainly from the NIMBUS 6 and 7 ERB satellite experiments, launched in mid-1975 and late-1978,

respectively, and from the Solar Maximum Mission ACRIM instrument in orbit since early 1980 (Willson, 1984). Some problems with other instruments on "Solar Max" were corrected in orbit in April 1984 by astronauts from the ill-fated Space Shuttle Challenger, but a fairly continuous ACRIM record exists from 1980 onwards.

The ERB/NIMBUS 6 was a simple detector comprised of a blackened flat plate attached to a thermopile incapable of electrical self-calibration. It relied on prelaunch calibrations to relate its observations to SI units (W m^{-2}), and had too wide a field of view to resolve the solar disk to better than 4°. Willson (1984) estimates ERB/NIMBUS 6 measurement uncertainties of $\Delta S/S_o \sim 0.2\%$, too large for reliable estimates of irradiance fluctuation effects on climate. The ERB radiometer on the follow-on NIMBUS 7 launched in late 1978 is a superior detector capable of self-calibration. But the most accurate irradiance monitoring in space to date is from the self-calibrating SMM/ACRIM, with a long-term accuracy estimated by Willson (1984) as $\Delta S/S_o < \pm 0.1\%$. In a flight test, three ACRIM sensors agreed to within 0.04% of their average result. The precision of the data, about ±0.002%, is higher than its accuracy. The time resolution of ACRIM raw data samples (< 1 s) is much shorter than what is required for climate analysis, so some level of averaging is needed for data analysis.

While speculations on sunspot effects on irradiance and climate have been made for hundreds of years (Lamb, 1972), the SMM/ACRIM data permit realistic assessments for the first time of correlations between irradiance and solar surface features observable from the Earth's surface (Hoyt and Eddy, 1982, 1983). The most studied features are *sunspots* - dark, cool regions on the sun's visible surface, or photosphere, whose numbers have varied with an ≈ 11-year cycle since they have been observed continuously by telescope beginning in the early 17th Century (Eddy, 1979). Indices of sunspot activity include the so-called Wolf daily sunspot number, N_W, and the number of daily sunspot groups, N_G. The number of sunspots/group is of the order of 10 (Hoyt and Eddy, 1983); $N_G \sim N_W/10$. The Zurich sunspot number, which is dominated by the number of sunspot groups, is also used.

The various sunspot indices tend to trend together; they approach a maximum when 100 or more individual spots are found on a solar hemisphere at one time, and a sunspot minimum, or quiet sun, when few or none are seen for months at a time. The last sunspot maximum was around 1980, hence the term Solar Maximum Mission (SMM) for the Sun-observing satellite launched that year. The fact that ERB/Nimbus 7 began monitoring irradiance two years before the sunspot maximum permits an assessment of possible irradiance-sunspot correlations. An overlapping record of ACRIM and ERB measurements is available beyond 1980.

In analyzing the early ACRIM data, Willson *et al.* (1981) focused on short-term solar irradiance fluctuations over the first 153-day period in 1980. The major new finding from early ACRIM data was that reductions in solar constant as much as $\Delta S/S_o \sim -0.2\%$ were found over timescales of 5-8 days as sunspot groups passed over the solar disk. This short-term anticorrelation is opposite in sign to the usual assumptions made by climate modelers that long-term luminosity variations are positively correlated with sunspots (Robock, 1979). Physical models for the short-term irradiance deficit are based on the idea that the "dark" sunspots create temporary blockage of emerging solar flux, which must be re-radiated over timescales of a month or more (Hoyt and Eddy, 1982). We show next that short-term anticorrelations of irradiance with sunspot groups may reverse in sign on monthly timescales when the five-year ACRIM record is used.

Figure 2a shows monthly averages of the first five years of ACRIM irradiance data. In contrast to the short-term anticorrelation of irradiance with the area of sunspot groups crossing its surface, the monthly mean ACRIM data trends downward along with sunspots since the sunspot maximum in 1980. The negative, deviant irradiance point at about April

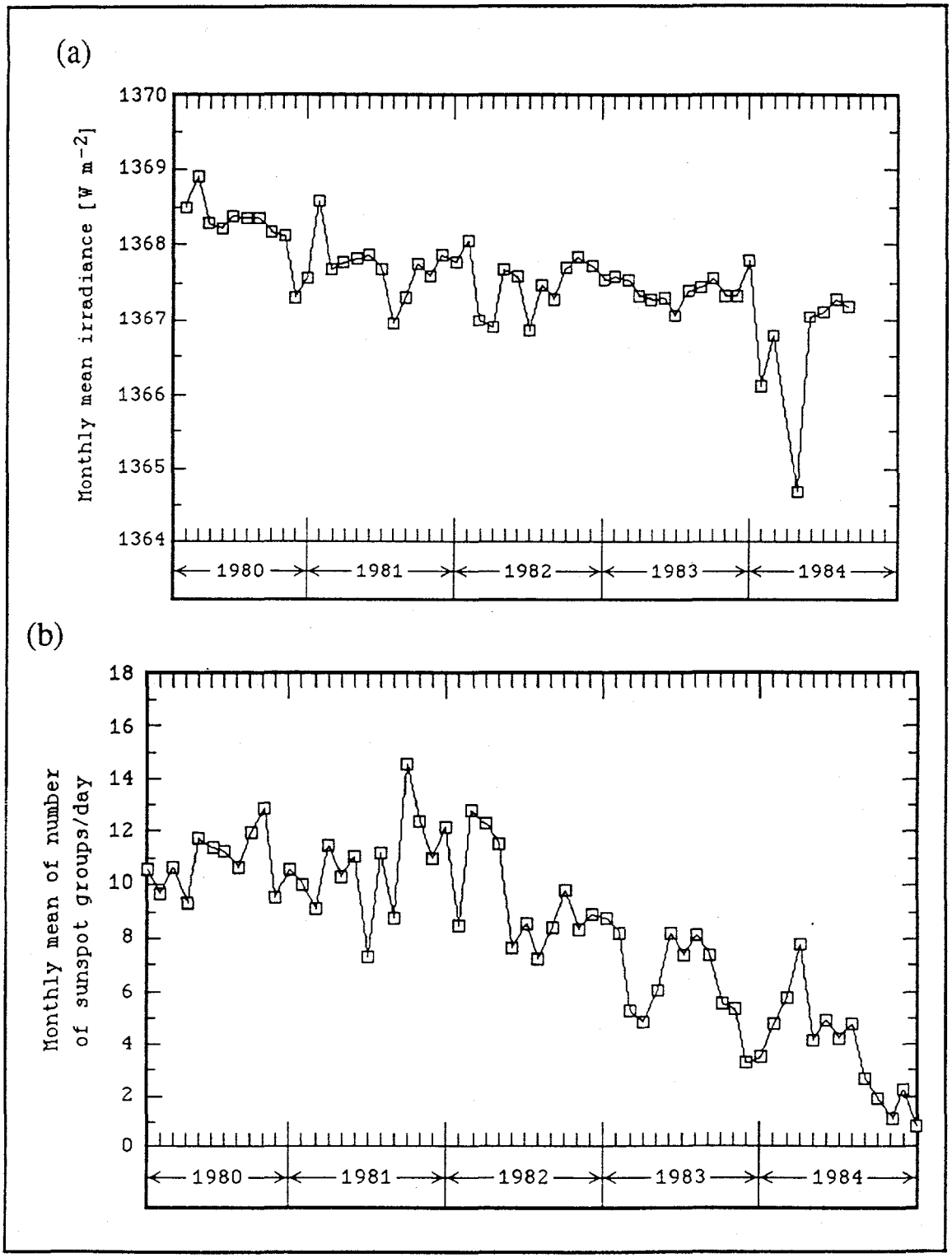


Figure 2. (a)SMM/ACRIM Irradiance Data : 1980-1984 (Willson, 1985). (b) NCAR sunspot group data: 1980-1984 (Gilliland, 1985).

1984 is probably insignificant; it occurred at the time of SMM repair and represents little real data. (Note also that only 11 months of data were available in 1980.) Sunspot numbers alone are often used in irradiance correlations for climate models. Robock (1979), for example, employed the positive correlation $\Delta S/S_o \sim 0.0052\langle N_W\rangle$ (%) to drive a climate model for the "Little Ice Age" in the northern hemisphere extending back to the sixteenth century, where $\langle N_W\rangle$ is the Wolf sunspot number of the 11-year cycle smoothed out. We now know, however, that such relations are based on surface irradiance observations of insufficient accuracy.

Other potentially relevant observables on the solar disk are available for the past 100 years from routine observations published by solar observatories. For example, the sunspot *umbra* is its dark, central core whose mean brightness is ~ 0.25 of the surrounding photosphere, and the *penumbra*, a somewhat less dark region surrounding the umbra with a brightness ~ 0.75 that of the surrounding photosphere. The relative contrast (brightness - 1) of these zones is ~ -0.75 for the umbra and ~ -0.25 for the penumbra. Hoyt (1979) proposed, on largely empirical grounds, a possible correlation between the umbra/penumbra ratio and the historical northern hemisphere surface temperature record from 1881 to 1980 of Jones *et al.* (1982). Also distinguishable from the darker photospheric background are irregular, bright patches, called *faculae* or *plage*. These "anti-sunspots" emit energy fluxes higher than the background levels of solar radiation. Their mean facular contrast is quite low, ~ +0.03 (Hoyt and Eddy, 1982), which makes the bright zones less obvious than sunspots, although their areal extent is greater. Both types of features are related to the underlying solar magnetic field activity through fluctuations in magnetic field lines, or "tubes," which penetrate the visible surface.

Based on corrected area-weighted contributions of light and dark areas on the photosphere, an expression can be written for solar irradiance incorporating projected surface areas of the umbra and penumbra of sunspot groups, u_i and p_i, the facular area f, and a correction factor for photospheric limb darkening, $C(\theta) = 0.36 + 0.84\cos\theta - 0.20\cos^2\theta$, where θ is the angle between the radius vector to the central point of the Sun and the line of sight to the spot group (Hoyt and Eddy, 1982),

$$\frac{\Delta S}{S_o} = 0.03f - \sum_{i=1}^{N_G} C(\theta)(0.75u_i + 0.25p_i) \quad .$$

From this formula one expects an anticorrelation between sunspots and irradiance if the area-weighted cooling from dark sunspots dominates the heating effect from faculae area. If, on the other hand, the bright faculae dominate, and if the facular area scales with sunspot area, then a positive sunspot/irradiance correlation should result. However, if the bright faculae and dark sunspots have comparable opposing magnitudes over given averaging periods, the net sunspot/irradiance correlation could very well be weak or insignificant, particularly so if there are unrelated factors influencing brightness variations. When Hoyt and Eddy (1982) applied their model to corrected umbral, penumbral and facular areas observed over the past 100 years spanning 10 solar cycles, the computed irradiance curve showed minima during sunspot maxima, suggesting a long-term *anticorrelation*. This is opposite to the long-term positive correlation assumed between sunspots and irradiance assumed by some climate modelers, and is opposite to the secular trend of the ACRIM data itself (see below).

4. STATISTICAL ANALYSIS OF ACRIM/NCAR DATA

Figure 3 is a scatter diagram showing the monthly mean ACRIM irradiance data of Fig. 2a in W m^{-2} versus the monthly mean sunspot group number $<N_G>$ of Fig. 2b. As a first step in the statistical analysis, the regression line through the data shown in Fig. 3 was computed using a least square best-fit routine. (The April 1984 point, while shown in Fig. 3, is not used in computing the regression line.)

A positive irradiance/group number correlation was found for the monthly data, albeit with appreciable scatter around the trend line. The standard deviation of irradiance was ± 0.56 W m^{-2}, corresponding to only $\sim$ 16% of the irradiance variation predictable by $<N_G>$. The group, rather than the Wolf, sunspot number was used for consistency with Willson *et al.'s* (1981) findings on short-term variability. More elaborate correlations such as the Hoyt and Eddy (1982) one based on umbral, penumbral and facular areas would presumably do better. Our objective at this point, however, was simply from gross statistical analysis to determine whether a time scale could be found at which the bulk correlation of irradiance with sunspot groups switches from negative to positive, perhaps arising from re-radiation after some time lag τ_{lag}.

To do this we computed the cross-correlation coefficients, R, which compare deviations about the means of the irradiance and group number time series. A value of R = +1 implies that the relative magnitudes and signs of deviations of one time series can be used to predict the behavior of the second time series; a value of R = -1 implies that deviations in one data set are comparable in magnitude but opposite in sign to the other,

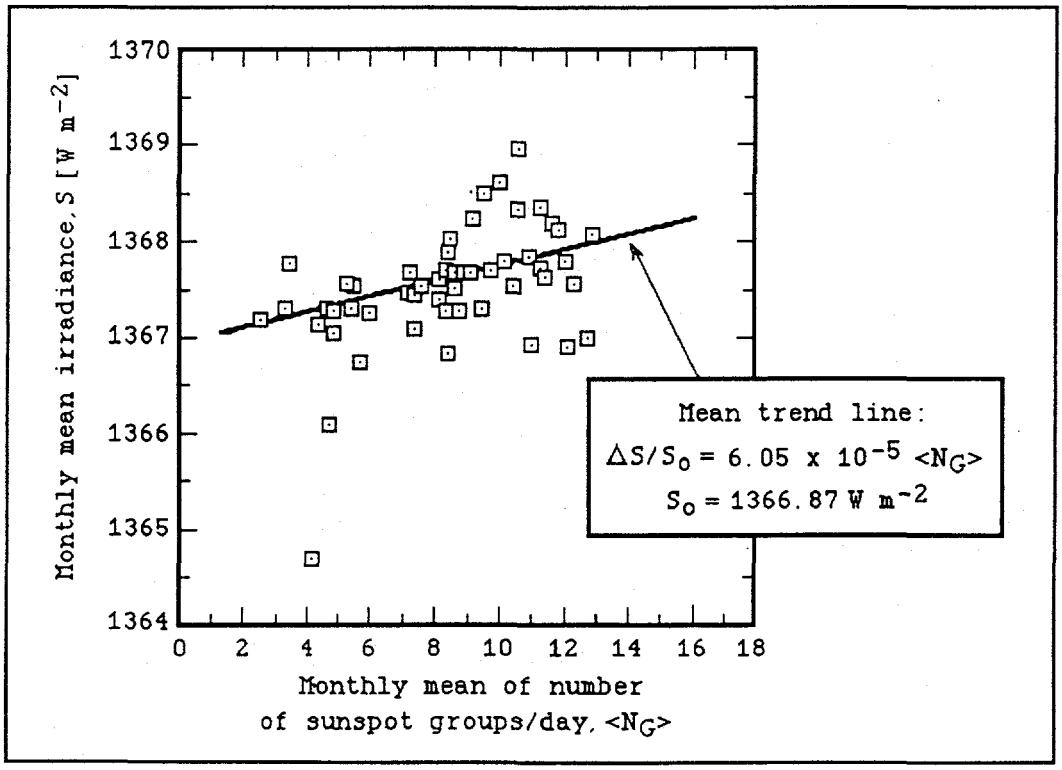


Figure 3. ACRIM monthly irradiance data versus NCAR sunspot groups: 1980-1984.

that is, they are anticorrelated. One can predict the behavior of one time series from the other with a confidence level of (R^2x100)%. To isolate possible effects of re-radiation from facular areas correlated with sunspots at earlier times, we introduced a time lag variable, τ_{lag}, such that $R^2(\tau_{lag})$ is the confidence level with which we can predict the behavior of one time series at time t + τ_{lag} from the behavior of the other time series at time t. The cross correlation coefficients versus lag time over the 1980-1984 time frame varied smoothly in the range of -14 months < τ_{lag} < +14 months, with a peak at τ_{lag} ~ 6 months. Throughout this range the formal predictability was only $100R^2$ < 25%, too small to indicate a significant lagged correlation of irradiance with earlier sunspots.

The whole-sun radiance measured calorimetrically by ACRIM is the integral of the incident component of energy flux emanating from all areas of the solar surface. Our statistical analysis did not include effects of bright features which may vary *independently* of present and past sunspots; this may explain the low correlation.

Foukal and Lean (1988) studied the slow changes in solar irradiance using ERB and ACRIM data. They found a low-amplitude (0.04% - 0.07%) variation over time-scales of 4 to 9 months which is well-correlated with changes in total facular radiations from an 81-day smoothing of daily data. They attribute the observed dimming of the Sun mainly to a decreasing contribution to the irradiance by a network of unbalanced bright photospheric magnetic elements. However, the contributions of sunspots and *large* faculae are nearly balanced, that is, the blocking effect of sunspots is compensated by the enhanced irradiance of "large" faculae. Our lack of a significant lagged correlation discussed above is consistent with their contention that "the excess radiation from these bright magnetic elements cannot represent the re-radiation of blocked sunspot heat flux." On the other hand, there is no firm basis for back calculations of solar effects on climate prior to the satellite era unless the faint-network radiation can be related to a pre-satellite observational data base. This will be discussed again later.

The downward irradiance trend of the ACRIM data is consistent with similar findings of solar cooling at about the same rate by the NIMBUS 7 ERB instrument during this period. (Both data sets are plotted versus time in Fig. 3 of Fröhlich, 1987). In the 1980-1984 period both the measured SMM/ACRIM and the measured NIMBUS 7/ERB irradiance data trend downward at a rate of about 0.01-0.02%/year (Willson, 1984). During the same epoch all indices of solar activity, including counts of groups (Fig. 1, middle panel) also decline, since the period embraces the declining phase of an 11-year solar cycle whose sunspot maximum was a wide peak spanning 1980-82. Thus one expects an apparent, if possibly accidental, positive correlation. The question is whether this reflects the behavior of the Sun over longer time frames. The answer is problematical because ERB/NIMBUS 7 irradiance values have decreased monotonically since early 1979, well before the peak of the 11-year activity cycle (Hickey *et al.*, 1981). That is, the NIMBUS data decline more or less monotonically while the sunspot numbers rise to a maximum (1980) and then fall, apparently refuting a straightforward correlation between measured irradiance and any simple sunspot index.

These uncertainties underscore that models for solar climatic effects based on correlations between irradiance and surface observables prior to 1980 should be viewed with caution, particularly since a number of correlations that have appeared previously in the literature are now known to be inconsistent with the space-based observations. We will review a number of these in the concluding section, where we also consider the implications for transient climate change of a new irradiance correlation developed by Schatten (1988) which *does* appear consistent with both solar physics and satellite data. But first we consider what can be learned about the global thermal-response to the early ACRIM measurement themselves using a simple transient-response climate model.

In addition to mean values, tables of solar observations from 1874 to 1981 (Hoyt and Eddy, 1982) are available which give the number of sunspot groups/day (N_G) and Wolf sunspot number (N_W), as well as the projected and corrected umbral area (u), whole spot area (w = u + p) and facular area (f). Areas are normally given in units of 10^{-6} of the solar disk. R. Gilliland (1985) of the National Center for Atmospheric Research has supplied us with monthly mean values of these parameters from 1980-1984, enabling an analysis over the ACRIM time frame. Figure 2b shows the variation of monthly mean sunspot groups/day, $\langle N_G \rangle$, from Gilliland's NCAR data over the same time frame as the ACRIM data of Fig. 2a. The declining trend of both monthly mean irradiance and monthly mean sunspots from 1981-1984 is quite evident.

5. CLIMATE MODEL RESPONSE TO SOLAR FORCING: 1980-1984

The world's oceans exert a kind of "thermal-flywheel" effect on all external climatic forcing including solar irradiance fluctuations. To study the influence of ACRIM irradiance data on the response of global-mean surface temperature $T_s(t)$ we used the upwelling-diffusion, one-dimensional ocean/transient climate model of Hoffert *et al.* (1980). The salient features of the model are reviewed below.

A useful reference condition for transient climate studies is the *equilibrium temperature*, T_e, corresponding to the steady-state surface temperature at solar flux S, planetary absorptance a, and atmospheric carbon dioxide concentration c. At the reference values of S_o, a_o and c_o, $T_o \equiv T_e$. An increase in any of the forcing parameters by $\Delta S = S - S_o$, $\Delta a = a - a_o$ or $\Delta c = c - c_o$ tends to create a new equilibrium surface temperature,

$$T_e = T_o + \beta_T \left[\frac{\Delta S}{S_o} + \frac{\Delta a}{a_o}\right] + \beta_c \ln\left[1 + \frac{\Delta c}{c_o}\right] \quad ,$$

where $\beta_T \approx 108°C$ and $\beta_c \approx 3.6°C$ are climate sensitivity parameters (Hoffert and Flannery, 1985). For a planet with zero thermal inertia, $T_s(t) = T_e(t)$. In the real world the $T_s(t)$ response is delayed and modified by oceanic mixing and storage in ways which depend on the $T_e(t)$ forcing.

The transient climate model used here computes heat capacity and internal mixing effects on $T_s(t)$ of an ocean mixed layer of depth h = 75 m and thermal relaxation time $\tau \approx 4$yr, overlying a deep ocean upwelling at $w \approx 4$ m yr^{-1} with eddy diffusivity $\kappa \approx 2000$ m^2 yr^{-1}. Considerations leading to these numerical values and to the model itself are discussed in Hoffert *et al.* (1980). The evolving surface temperature is computed by numerical solution of the differential equation

$$\frac{dT_s}{dt} = \frac{[T_e(t) - T_s]}{\tau} + \frac{1}{h}\left[\kappa\frac{\partial T}{\partial z} + w(T - T_p)\right]_{z=h} \quad ,$$

where T_p is the temperature of polar bottom water. The last term in large brackets on the right-hand side is the rate of temperature change from heat transfer with the deep ocean at the mixed layer/thermocline interface. If the term is small, then heat is trapped in the mixed layer and only superficial heating of the oceans needs to occur for climate to re-equilibrate with an altered surface heat balance. If the term is large, then warming of the ocean's surface cannot occur until the ocean warms from top to bottom. To evaluate the complete ocean model in transient evolution, the T_s-equation is integrated numerically simultaneously with a coupled 1-D, upwelling-diffusion energy transport equation for T(z,t) in the deep ocean,

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial z}\left[\kappa\frac{\partial T}{\partial z} + w(T - T_p)\right] \quad ,$$

where z is depth below the mixed layer, and the bottom boundary condition on this equation is $\kappa\partial T/\partial z + wT = wT_p$ at $z = h_d$. The partial differential equation for oceanic internal temperature, T(z,t), is solved by finite differences over 81 fifty-meter-thick layers between the mixed layer/thermocline interface and the sea floor at $h_d = 4050$ m.

To calculate the irradiance effect from 1980-1984, we forced the system with $T_e(t) = T_o + \beta_T \Delta S(t)/S_o$, where t is the time from a hypothetical initial state in 1980, T_o is the 1980 surface temperature, $S_o = 1368.4$ W m^{-2} is the monthly mean ACRIM value for early 1980, and $\Delta S(t)/S_o$ was specified by the monthly mean ACRIM irradiance data of Fig. 2a. The polar sea temperature was held constant at $T_p = 1°C$ during the run. The initial vertical ocean temperature profile used to start the calculation was specified by

$$T(z,0) = T_p + (T_o - T_p)e^{zw/\kappa} \quad ,$$

that is, a pre-existing oceanic steady state was assumed in 1980. This assumption was imposed by lack of prior data, but should be viewed with caution in light of a possible "history effect" (see below).

The surface temperature response is illustrated in Fig. 4. Also shown is the equilibrium temperature forcing and the global surface temperatures from 1975-1985 blown up from Fig. 1. It is evident in Fig. 4 that for the climate sensitivity parameter used here ($\beta_T \sim 108°C$), the ACRIM irradiances correspond to short-period (monthly) fluctuations of equilibrium temperature of order $\sim 0.1°C$. The predicted longer-term trend over the five years is cooling, but considerably smoothed relative to the forcing and with fluctuations damped to the $\sim 0.01°C$ level by the ocean's thermal inertia. Interestingly, although the smoothed Jones *et al.* (1986) temperature data indicate a warming in the early 1980s, the more recent 5-year running-mean global temperatures of Hansen and Lebedeff (1988) show a pronounced cooling dip from 1982-1984, followed by a very recent global warming.

The magnitude of the solar-induced effect is generally smaller than the observed temperature signal, but could still contribute a significant part of the signal over time. It is tempting, but probably premature, to attribute Hansen and Lebedeff's cooling dip to the irradiance decline.

The computed response reflects the interplay of oceanic mixing and storage to modulate the imposed solar signal. The total response might look different with a different initialization, as the memory of prior heat anomalies in the ocean is reflected in the superimposed transient solutions. But since the model equations are linear, our solar-induced transient solution initialized at a steady state would simply add to prior history, greenhouse-gas and albedo-change components.

Hansen *et al.* (1986) suggested that the solar cooling implied by satellite measurements might have been compensated for by anthropogenic greenhouse-gas warming over comparable time periods, and that such effects might even continue for some time. Wigley (1987, personal communication) estimates that for roughly the five-year period from 1979-1984, the total reduction in infrared cooling to space from the observed CO_2 increase of ~ 6 ppm, and from other anthropogenic greenhouse gases, was ~ 0.26 W m^{-2}. Since the observed irradiance decline over this period of $\sim 0.1\%$ corresponds to a planet-wide-average decrease of absorbed solar flux of ~ 0.24 W m^{-2}, a short-term cancellation of effects on the radiation balance is indeed possible, bearing in mind that the uncompensated effect over these 5 years is in any event small.

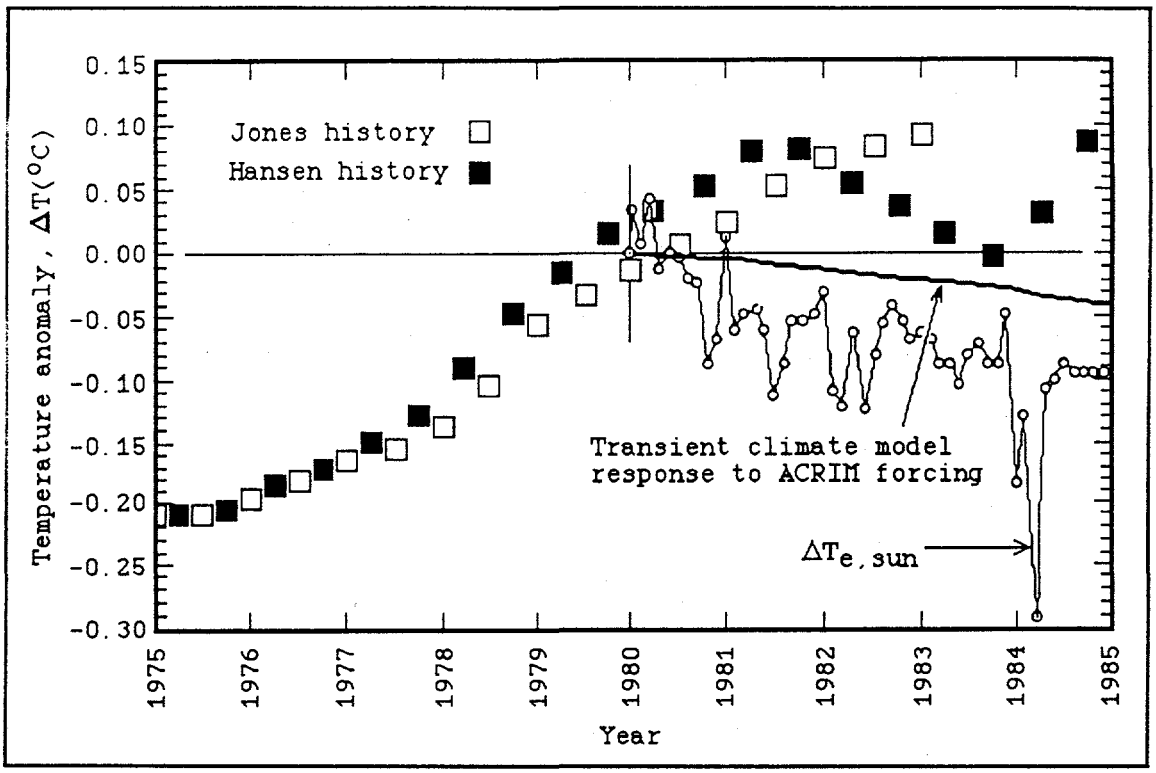


Figure 4. ACRIM equilibrium temperature forcing and model response initialized in 1980 compared with a blow-up of the overlapping temperature histories of Fig. 1.

6. IMPLICATIONS FOR TRANSIENT CLIMATE MODELS

I have already referred to the pre-ACRIM positive irradiance-sunspot-number correlations used for example by Schneider and Mass (1975) and Robock (1979) in climate models as being based on insufficiently accurate observations. Hansen *et al.* (1981) used the Hoyt (1979) umbra/penumbra ratio correlation for the solar component to improve predictions by their model of the local peak around 1940 of global temperature (see Fig. 1). This correlation has not to our knowledge been tested against extraterrestrial irradiance measurements, and Hoyt (1979) himself states: "The high cross-correlation between northern hemisphere temperature anomalies and the umbral/penumbral ratio may be a mathematical oddity without physical meaning." Gilliland and Schneider (1984) modeled the effects of solar forcing in a transient climate model with a sinusoidal term based on assumed solar radius cycle of 76-year period with phase and amplitude arbitrarily adjusted to fit temperature data. However, their "best fit" of solar forcing to surface temperature histories contradicts satellite observations, since Gilliland and Schneider (1984) show a rise in solar forcing over the 1980-1984 time frame when both SMM/ACRIM and ERB/NIMBUS 7 instruments measured declining irradiance trends.

It is now clear that correlations grounded in accurate (space-based) observations are needed to reliably assess solar-variation effects on climate, but some care in needed with regard to matching these to the time-scales of interest. At the time, the Hoyt and Eddy (1982) short-term anticorrelation seemed well-motivated by the ACRIM data. It produced a 90% correlation with sunspot blocking (10% short-term storage) based on ACRIM data

over the first year of Solar Max operation (Hoyt and Eddy, 1983). To extend it over long periods, one needs corrected umbral, penumbral and facular areas (these are available for the past ~ 100 years). But, when this extension was done by Hoyt and Eddy (1982) for the period April 1974-October 1981, their predictions of monthly mean solar irradiance (computed from their sunspot and facular radiation model and observed projected sunspot areas) had *minima* of order 0.1% during periods of *maximum* solar activity. This anti-correlation of long-term solar activity with irradiance is in the same direction as the short-term blockage effect of Willson *et al.* (1981), but opposite to the five-year trend of the ACRIM and ERB data! The discrepancy can only mean that other features on the solar surface correlated with the solar cycle brighten sufficiently to compensate for short-term sunspot blockage effects of Hoyt and Eddy (1982). The findings of Foukal and Lean (1988) suggest these features are faint-network faculae distributed over the large solar surface.

The blockage factor of Hoyt and Eddy has nonetheless proven useful in deciphering the ACRIM record. Livingston *et al.* (1988) showed that the ACRIM irradiance record corrected for sunspot blockage is well-correlated with ground-based irradiance spectrophotometry of Fraunhofer absorption lines of certain elements in the sun's atmosphere. In particular, the strength of the mid-photospheric manganese 0.5394 μm line tracks the corrected ACRIM irradiance almost perfectly. (Because line strength is determined relative to the local blackbody continuum, no correction for sunspot blocking is needed to show any global variability.) These comparisons independently confirm the variability of the corrected ACRIM signal, indicate that its source is (probably network) faculae, and indicate that the corrected irradiance follows the 11-year activity cycle.

Evidently, the key to success in modelling climatically significant irradiance variations over solar cycles is the way one handles the residual irradiance from network-facular radiation: A positive residual gives a brighter sun as the sunspot number increases. A recent solar irradiance model in this vein whose predictions overlap the ACRIM time frame was developed by Lean and Foukal (1988), who model irradiance variations due to solar magnetic activity during the past three solar cycles. They find irradiance residuals (irradiance minus the blocking effects of sunspots) to be well-correlated with the 10.7 cm microwave flux which represents radiation from bright facular elements. Assuming that this correlation holds over the past three solar cycles, and by using estimates of sunspot contrast and number in the past, they reconstruct the irradiance time series from 1954-1985. The match between the data trend and their model over the overlapping period appears to be quite good.

Because the 1-D ocean analysis of the previous section was intentionally limited to the ACRIM time frame, it did not include the history effect that would develop from a "running start" in which the memory of earlier forcing cycles was present. Prior simulations of the response of a coupled atmosphere/ocean transient model have shown that the amplitude of the global temperature response to harmonic radiative forcing of fixed amplitude increases as the period of the forcing period becomes large compared with the mixed layer thermal damping time (Hoffert *et al.*, 1980). Moreover, if one is aiming at a transient climate model on the time scale of the hundred-year instrumental record, it may be necessary to account for changes in the peak number of sunspots during a given cycle, a quantity which itself varies over still longer periods. Reid (1987), for example, has recently linked the variation in sea surface temperature (SST) to the *envelope* of the 11-year cycle of solar activity (this is called the Gleissberg cycle and has a period of ~ 80 years). The solar constant was varied in phase with the envelope of the sunspot number, with the amplitude of the solar constant assumed proportional to the sunspot number on the envelope. Using this variation of solar constant with time as the forcing term in the transient one-dimensional, upwelling-diffusion model of Hoffert *et al.* (1980), Reid reconstructs SST's from 1860-1980.

To assess the implications of these effects for transient climate models, we worked with a recent, and fairly comprehensive, model for solar irradiance variability developed by Schatten (1988). His model includes: (i) effects of solar active regions (sunspots and large faculae), (ii) effects of global faculae (polar and network) and (iii) long-term variations in the upper envelope of sunspot number. The latter effect sets the so-called "quiet sun" solar constant S_o associated with few or no active regions and little or no global field or faculae. The ratio $\Delta S/S_o$ (where ΔS is the deviation from S_o) is computed as the sum of two factors, one accounts for the blocking effect of sunspots and the other for the enhanced irradiance from faculae. Faculae are due to both active regions and *global* sources. The effect of the meridional drift of the active regions across the solar photosphere is considered when computing the irradiance variation. This model also predicts irradiance to be positively correlated with the 11-year cycle of sunspot activity. As shown in Fig. 5, the ACRIM irradiance trend from 1980-1984 is matched quite well. Moreover, Schatten (1988) used the sunspot prediction of Schatten and Sofia (1987) for the projected solar constant variation through 1997. Figure 5 also shows this prediction from 1987-1997 and its associated standard error.

We used Schatten's (1988) model of solar irradiance variation to drive our transient climate model from 1976 more than one solar cycle into the future. Figure 6 shows the equilibrium temperature forcing and the model's response. In general, the imposed solar

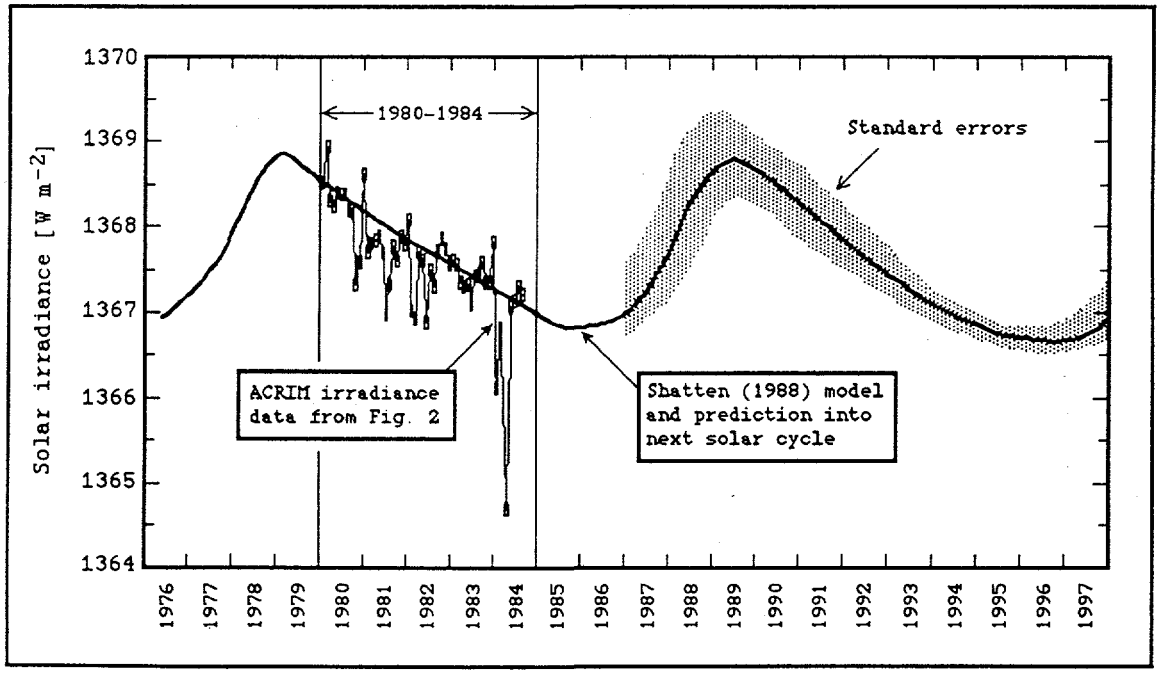


Figure 5. Solar irradiance model from 1976-1997 from a model incorporating effects of global faculae and active cavity regions as a function of solar latitude and time. The projection into the next sunspot cycle is based on Schatten and Sofia's (1987) prediction of sunspot number; the solar constant for zero solar activity was normalized to 1366.5 W m^{-2} for comparison with ACRIM data. The model curves were supplied by K. H. Schatten of NASA/Goddard Space Flight Center (personal communication), and differ from those in Schatten (1988) insofar as they include a more realistic and gradual change of polar faculae during the cycle.

signal is both damped and phase-lagged by the thermal inertia of the oceans, and the response exhibits an upward drift as it seeks to oscillate about the mean equilibrium temperature of the forcing. (This would occur asymptotically for a large enough number of identical cycles.) Hoffert *et al.* (1980) have shown that a 65-70% reduction of the amplitude of the response to identical 11-year sunspot cycles is caused by mixed layer damping. From Fig. 6 it is seen that the peak of the response is ~ 0.05°C which is consistent with this estimate. Since the amplitude of the response will increase if the forcing period becomes large (while keeping the forcing amplitude constant), a model driven by the ~ 80-year Gleissberg cycle will exhibit a temperature response that will be damped to a smaller extent.

Note that the "historical" temperatures of Jones *et al.* (1986) and Hansen and Lebedeff (1988) in Fig. 6 are smoothed over interannual variations and contain the effects of atmospheric aerosol and greenhouse-gas variations as well as solar forcing. The variance in the observed signal is substantially larger than that predicted by solar effects alone, but does not appear inconsistent with such effects. The (linear) thermodynamic upwelling-diffusion ocean model results discussed here indicate that currently available satellite data are sufficient to rule out a major solar-variation effect on surface temperature in the short term (but longer-term effects are still possible and may even be quite important). We should also emphasize that our 1-D ocean model may not be entirely appropriate for analyzing the local response to short time scale forcing, particularly in low thermal inertia continental interiors, and the possibility of localized nonlinear amplification cannot be ruled out. Moreover, we cannot exclude the possibility that nonlinearities in the climate system amplify and modulate imposed forcing in ways not captured by the current linear models (Gaffin *et al.*, 1986).

In light of these results, it seems extremely important to continue monitoring the variations of solar flux from space in the coming decades, perhaps from a dedicated instrument on the proposed NASA Earth Observing System (EOS). Since some transient

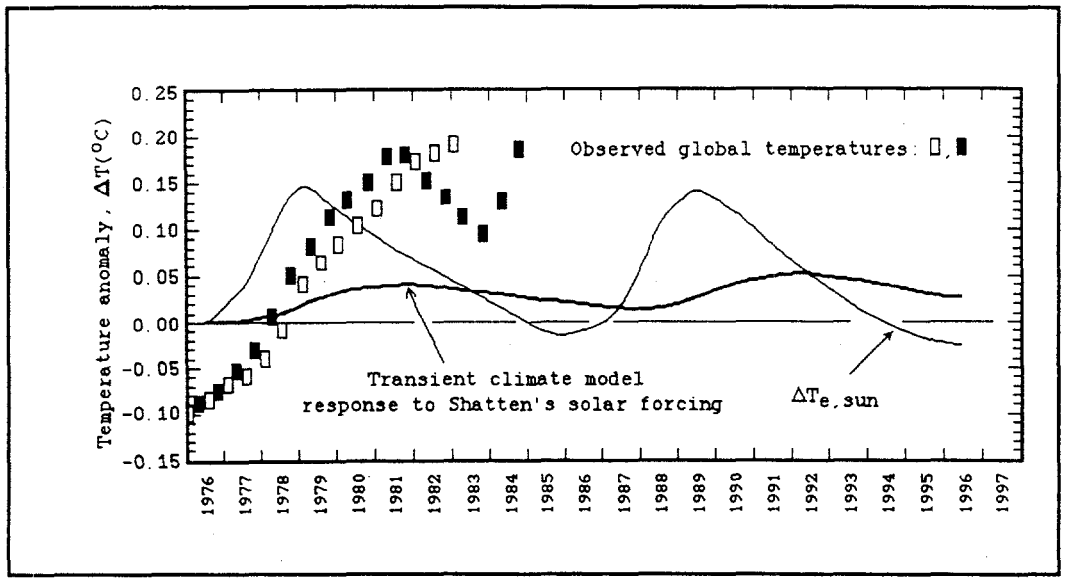


Figure 6. Equilibrium and transient global temperature response from the transient climate model for Schatten's (1988) solar forcing over the present and subsequent solar cycle.

climate models predict the emergence of a fossil fuel carbon dioxide warming signal in this time frame, it would be particularly useful to know the components of the signal that can be attributed to other effects, including solar irradiance changes. Hopefully, future generations will have the data needed to resolve the transient effects of our dynamic Sun on climate, as long time-series irradiance monitoring from space becomes an operational fact of life.

ACKNOWLEDGEMENTS

We thank Tom Wigley for drawing our attention to the "history effect" and its possible importance. This work was supported at New York University by the National Aeronautics and Space Administration (Solar Maximum Guest Investigator Program) under Grant NAG 5-503 and by the Department of Energy (Carbon Dioxide Research Division) under Grant DE-FGO2-85EK60350.

REFERENCES

Budyko, M. I., E. K. Byutner and K. Ya. Vinnikov, 1986: Detecting anthropogenic climate change. Presented at the USSR/USA Meeting of Experts on Causes of Recent Climate Change, Leningrad, July, 1986; to appear in *Meteorology and Hydrology*, **12** (in Russian); English version available from M. I. Hoffert.

Eddy, J. A., 1979: *A New Sun: The Results from Skylab*. NASA SP-402, Scientific and Technical Information Service, NASA, Washington, DC.

Foukal, P., and J. Lean, 1988: Magnetic modulation of solar luminosity by photospheric activity. *The Astrophysical Journal* (in press).

Fröhlich, C., 1987: Variability of the solar "constant" on time scales of minutes to years. *J. Geophys. Res.*, **92**, 796–800.

Gaffin, S. R., M. I. Hoffert and T. Volk, 1986: Nonlinear coupling between surface temperature and ocean upwelling as an agent in historical climate variations. *J. Geophys. Res.*, **91**, 3944–3950.

Gilliland, R. L., 1982: Solar, volcanic and CO_2 forcing of recent climatic changes. *Climatic Change*, **4**, 111–131.

Gilliland, R. L., 1985: Variations in sunspots and facular area from 1981–1985 (personal communication). National Center for Atmospheric Research, Boulder, CO.

Gilliland, R. L. and S. H. Schneider, 1984: Volcanic, CO_2 and solar forcing of Northern and Southern Hemisphere surface air temperatures, *Nature*, **310**, 38–41.

Hansen, J. E., D. Johnson, A. Lacis, S. Lebedeff, P. Lee, D. Rind and G. Russell, 1981: Climatic impact from increasing carbon dioxide. *Science*, **213**, 957–966.

Hansen, J. E., A. Lacis, D. Rind, G. Russell, I. Fung, P. Ashcraft, S. Lebedeff, R. Ruedy and P. Stone, 1986: The greenhouse effect: Projections of global climatic change. In *Effects of Changes in Stratospheric Ozone and Global Climate, Vol. 1: Overview*, U.S. Environmental Protection Agency, Washington, DC, August 1986.

Hansen, J. E., and S. Lebedeff, 1988: Global surface air temperatures: Update through 1987. *Geophys. Res. Lett.*, **15**, 323–326.

Hickey, J. R., B. M. Alton, F. J. Griffen, H. Jacobowitz, P. Pellegrino, E. A. Smith, T. Vonder Haar and R. H. Maschhoff, 1981: Solar variability indications from Nimbus 7 satellite data. In *Variations of the Solar Constant*, S. Sofia (ed.), NASA Conference Publication #2191, NASA, Washington, DC, pp. 59–72.

Hoffert, M. I., A. J. Callegari and C. T. Hsieh, 1980: The role of deep sea heat storage in the secular response to climatic forcing. *J. Geophys. Res.*, **85**, 6667–6679.

Hoffert, M. I., and B. P. Flannery, 1985: Model projections of the time-dependent response to increasing carbon dioxide. In *Projecting the Climatic Effects of Increasing Carbon Dioxide*, M. C. MacCracken and F. M. Luther (Eds.), DOE/ER-0237, United States Department of Energy, Carbon Dioxide Research Division, Washington, DC, pp.148–190.

Hoyt, D. V, 1979: Variations in sunspot structure and climate. *Climatic Change*, **2**, 79–92.

Hoyt, D. V., and J. A. Eddy, 1982: *An Atlas of Variations in the Solar Constant Caused by Sunspot Blocking and Facular Emissions from 1874 to 1981*. NCAR/TN194/STR, High Altitude Observatory, National Center for Atmospheric Research, Boulder, CO; available from National Technical Information Service, Springfield, VA.

Hoyt, D. V., and J. A. Eddy, 1983: Solar irradiance modulated by active regions from 1969 through 1981. *Geophys. Res. Lett.*, **10**, 509–512.

Jones, P. D., T. M. L. Wigley and P. M. Kelly, 1982: Variations in surface air temperature: Part I. Northern Hemisphere, 1881–1980. *Mon. Wea. Rev.*, **110**, 59–70.

Jones, P. D., T. M. L. Wigley and P. B. Wright, 1986: Global temperature variations between 1961 and 1984. *Nature*, **322**, 430–434.

Lamb, H. H., 1972: *Climate: Past, Present and Future*, Vol. 1, Methuen, London.

Lean, J. and P. Foukal, 1988: A model of solar luminosity modulation by magnetic activity between 1954–1984. *Science*, **240**, 906–908.

Livingston, W. C., L. Wallace and O. R. White, 1988: Spectrum line intensity as a surrogate for solar irradiance variations. *Science*, **240**, 1765–1767.

Newkirk, G., Jr., 1983: Variations in solar luminosity. *Ann. Rev. Astr. Astrophys.*, **21**, 429–467.

Pittock, A. B., 1978: A critical look at long-term sun-weather relationships. *Rev. Phys. Space Phys.*, **16**, 400–420.

Pittock, A. B., 1983: Solar variability, weather and climate, an update. *Quart. J. Roy. Meteorol. Soc.*, **109**, 23–57.

Reid, G. C., 1987: Influence of solar variability on sea surface temperatures. *Nature*, **329**, 142–143.

Robock, A., 1979: The "Little Ice Age": Northern hemisphere average observations and model calculations. *Science*, **206**, 1402–1404.

Schatten, K. H., 1988: A model for solar constant secular changes. *Geophys. Res. Lett.*, **15**, 121–124.

Schatten, K. H., and S. Sofia, 1987: Forecast of an exceptionally large even-numbered solar cycle. *Geophys. Res. Lett.*, **14**, 632–635.

Schneider, S. H., and C. F. Mass, 1975: Volcanic dust, sunspots and temperature trends. *Science*, **190**, 741–746.

Vinnikov, K. Ya., and P. Ya. Groisman, 1979: Empirical model of climatic change. *Meteorologiya i Gidrologiya*, **3**, 25–36 (in Russian).

Wigley, T. M. L., 1987: personal communication. Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, UK.

Willson, R. C., 1984: Measurements of solar total irradiance and its variability. *Space Science Rev.*, **38**, 203–242.

Willson, R. C., 1985: ACRIM data on solar irradiance 1980–1984, Jet Propulsion Laboratories, Pasadena, CA (personal communication).

Willson, R. C., and H. S. Hudson, 1981: Variations of solar irradiance. *Astrophys. J. Letters*, **244**, L185–L189.

Willson, R. C., S. Goulkis, M. Janssen, H. S. Hudson and G. A. Chapman, 1981: Observations of solar irradiance variability. *Science*, **211**, 700–702.

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 96, NO. C3, PAGES 4957–4969, MARCH 15, 1991

Marine Biota Effects on the Compositional Structure of the World Oceans

H. S. KHESHGI AND B. P. FLANNERY

Corporate Research Laboratories, Exxon Research and Engineering Company, Annandale, New Jersey

M. I. HOFFERT

Department of Applied Science, Earth Systems Group, New York University

The vertical structure of total carbon, alkalinity, nutrients, and dissolved oxygen in the world oceans is examined with a one-dimensional equatorial ocean/polar ocean box model. Photosynthesis/respiration cycles affect and are affected by fluxes in the ocean and the structure of the profiles. Marine biota produce by-products that lead to organic and inorganic (calcareous) sediments. In steady state, rates of phosphorous and alkalinity runoff from land are linked to surface nutrient supply, the rates of particulate rain, the degree of anoxia near sediments, the lysocline depth, and thereby control rates of sedimentation. These, in turn, are influenced by internal mixing dynamics and the action of the marine biota. The interdependence of ocean composition and rates of organic and inorganic carbon burial is found to be sensitive to the traits of the marine biosphere.

1. INTRODUCTION

Oceanic biological activity influences the chemical composition of the ocean and thereby affects where carbon resides on the Earth, yet the mechanisms by which the marine biosphere acts and reacts to its environment are far from being quantitatively understood. In this paper we hypothesize simple mechanistic reactions of marine biota to its environment (e.g., to nutrient levels). To study the biosphere's role in the ocean carbon cycle, we use a model for transport of dissolved chemical components and particulate biomass, production or decay/dissolution of biomass, and consumption/release of chemical components by the biomass. Our model, shown in Figure 1, approximates the world oceans as a one-dimensional (1D) upwelling ocean representing low-latitude regions, a polar surface water box representing downwelling high-latitude regions, and a bottom water box.

Classical box models treat carbon cycle species as well mixed in reservoirs which exchange fluxes at their boundaries (in both directions) at rates proportional to concentrations in adjacent boxes [*Keeling*, 1973; *Bacastow and Keeling*, 1973; *Bjorkstrom*, 1979; *Broecker and Peng*, 1982]. This type of transport is essentially diffusive. When many boxes are stacked vertically beneath a surface mixed layer, they make up a box-diffusion model whose vertical exchange coefficients are related to the vertical eddy diffusivity. *Oeschger et al.* [1975] demonstrated this limiting relationship between disaggregated and continuous models but did not include advective transport or a downward transport path from high-latitude outcrops ventilating deeper regions. A subsequent development was box-outcrop models with high-latitude water exchanging with deeper regions within the framework disaggregated boxes, some of which may only have one-way transport [*Sarmiento and Toggweiler*, 1984; *Lapenis and Kolomeitsev*, 1987; *Volk and Liu*, 1988]. The effect of isopycnal mixing can also be approximated by an outcrop-diffusion model in which mixing is allowed to occur along inclined constant potential density surfaces [*Siegenthaler*, 1983]. Some authors have also developed multi-ocean box models disaggregated according to individual basin data [*Broecker and Peng*, 1986; *Bjorkstrom*, 1986]. The main point is that the box models typically have exchange coefficients calibrated to recover observed steady state fields, with little or no quantitative input from the physical circulation and mixing of the ocean. So-called inverse box models (including dynamical constraints) are an exception, but for various reasons have not proven successful as yet [*Moore and Bjorkstrom*, 1986].

The alternative to box models is to compute interior concentration fields from spatially continuous conservation equations using specified ocean transport properties (currents and diffusivities) in one, two, and three dimensions. In one (vertical) dimension, the transport properties are vertical advection (normally upwelling) and vertical eddy diffusivity. One-dimensional upwelling-diffusion models for deep-sea temperature [*Munk*, 1966], dissolved oxygen [*Wyrtki*, 1962], and radiocarbon [*Craig*, 1969] are classical deep-sea "recipes" of oceanography. The upwelling-diffusion model was subsequently extended to the surface, and to include a polar sea box outcrop which closes the thermohaline circulation [*Hoffert et al.*, 1981; *Volk and Hoffert*, 1985]. This is the approach taken in the present model. Its advantage is that the depth dependence of the concentration of all relevant species can be linked to physical transport properties. The model provides a useful picture of coupled biology and chemistry in the world ocean with a relatively modest computational investment. It also permits testing and assessment of parameterizations for marine chemistry and biological effects that can be subsequently incorporated in more spatially resolved models whose internal current fields are derived from ocean circulation models [*Sarmiento*, 1986; *Maier-Reimer and Hasselmann*, 1987; *Bacastow and Maier-Reimer*, 1990].

In the present oceans, biomass is produced primarily by

Copyright 1991 by the American Geophysical Union.

Paper number 90JC02336.
0148-0227/91/90JC-02336$05.00

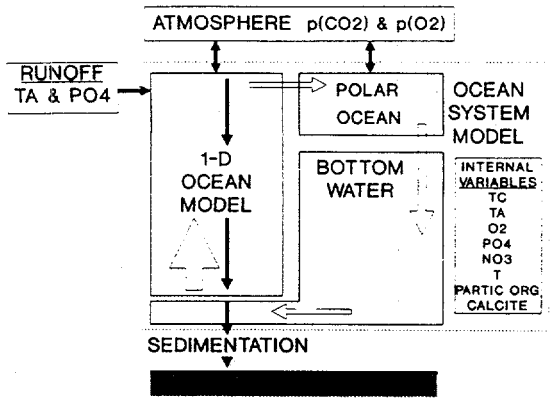


Fig. 1. Diagram of the ocean system model.

photosynthesis in upwelling oceans. The rate at which nutrients, namely, phosphate and nitrate, are transported to the thin phototropic layer on the ocean's surface limits the rate of biomass production. Photosynthesis in the mixed layer at the surface of the ocean also consumes carbon and releases oxygen. Eventually, planktonic biomass dies, and a fraction of the biomass falls from the mixed layer in a particulate form comprised of organic bodies along with calcareous and silicious shells. As the biomass particulates settle to the ocean floor, the organic bodies decay, releasing nutrients and carbon, and consuming oxygen at mid-ocean depths. A small fraction of the organic bodies reach the ocean floor, where some are buried in sediment to be enfolded into the Earth's crust. The calcium carbonate shells also fall to the ocean floor, where they can dissolve if they are at a depth below the lysocline, the depth below which the ocean is sufficiently undersaturated by calcium carbonate to permit extensive dissolution, or they too can be buried and rendered inaccessible to exchange with the ocean's waters.

An effect of particulate transport in the ocean made possible by biological activity is that carbon and nutrients are pumped down from the mixed layer to the deep ocean, thereby increasing concentrations in the deep ocean, whereas oxygen is consumed and its concentration is lowered in the abyssal ocean. In this way, biological activity affects the depth dependence or structure of ocean composition and does so on a time scale roughly equal to that required for a parcel of ocean water to upwell from the ocean floor to the ocean surface, about 1000 years. A model for the transport of dissolved and particulate components and for biological transformation of dissolved components to and from particulate components in an upwelling ocean is developed (in section 2) and compared to the composition depth profiles of the present world oceans.

Capture of particulate biomass in ocean sediments removes components from the ocean system, namely, carbon, phosphate, and the components that affect ocean alkalinity. For the ocean to be in long-term steady state, these components must be returned to the ocean system by runoff or, as with carbon dioxide, exchange with the atmosphere. Herein, the rate and composition of particulates captured in sediments are estimated, and the estimate is equated to specified runoff rates of components so that the model ocean system is in long-term steady state. The dependence of ocean steady states on ocean circulation, the traits of marine biota, and rates of runoff are reported in section 3.

2. 1D Model for Upwelling Oceans

2.1. *Model Description*

Our upwelling ocean model evaluates the distribution of temperature and several species, subject to transport as described in a 1D upwelling-diffusion model characterized by its vertical eddy diffusivity K and upwelling velocity w (see notation section below). To the upwelling ocean model we add an additional transport term to describe settling of biomass at velocity v_i. Temperature and species concentrations are the dependent variables c_i in the model. Specifically, these are temperature [T], total carbon [TC], oxygen [O_2], titration alkalinity [TA], phosphate [P], and nitrate [N], as well as particulate organic [BO] and inorganic [BI] biomass. The distribution of temperature and composition is controlled by the advection-diffusion equation

$$\frac{\partial c_i}{\partial t} = \frac{\partial}{\partial z}\left(K\frac{\partial c_i}{\partial z}\right) - \frac{\partial (v_i c_i)}{\partial z} + w\frac{\partial c_i}{\partial z} + S_i = 0 \qquad (1)$$

In (1), z is the depth ranging from 0 at the surface to B (6000 m) at the bottom, c_i is any of the dependent variables of the model, and S is the source term for net local production of component i. As written in (1), the coefficients K, v, w, and S can depend on depth and the local dependent variables and can differ for each component, e.g., $v = 0$ except for BO, BI. Internal to the model, $0 < z < B$, the net downward flux J_i for any component i is

$$J_i = -K\frac{\partial c_i}{\partial z} + (v_i - w)c_i \qquad (2)$$

Typical values for the transport coefficients are $K = 2000\ \mathrm{m^2\ yr^{-1}}$, and $w = 4\ \mathrm{m\ yr^{-1}}$ [*Hoffert et al.*, 1980], and $v = 32{,}000\ \mathrm{m\ yr^{-1}}$ only for particulates [*Volk*, 1984].

The characteristic steady state solution of (1) for a component, such as temperature, where the source term is zero and where the transport coefficients are constant, exhibits simple exponential decay:

$$T(z) = T_b + (T_0 - T_b)e^{-z/L} \qquad (3)$$

where T_0 is the temperature at the surface, and L ($= K/w = 500$ m) is a characteristic length for temperature change. Throughout the interior the "flux of temperature" is wT_b, where T_b is the temperature of upwelling bottom water.

Boundary conditions require that the flux J_i of species i must be continuous across both the top or bottom boundary. At the surface the model contains a well-mixed layer of depth H (100 m) in contact with the atmosphere. At the bottom the model is fed by upwelling water carrying specified amounts of each dependent variable. At the bottom boundary, $z = B$, the external flux is characterized by the value c_{ib} of species i that upwells:

$$J_{ib} = (v_i - w)\rho c_{ib} \qquad z = B \qquad (4)$$

At the air-sea interface, $z = 0$, the exchange nutrients and biomass is zero. Heat, CO_2, oxygen, and alkalinity transfer across the surface boundary.

For the gases,

$$J_0(\mathrm{CO_2}) = E(\mathrm{CO_2})\frac{p_a(\mathrm{CO_2}) - p_0(\mathrm{CO_2})}{340\ \mu\mathrm{atm}} \qquad (5)$$

$$J_0(O_2) = E(O_2)\frac{[O_2]_{sat} - [O_2]_0}{[O_2]_{sat}} \quad (6)$$

In the boundary conditions, $p_a(CO_2)$ is the partial pressure of CO_2 of the atmosphere far above the surface of the ocean and $p_0(CO_2)$ is the partial pressure of CO_2 in ocean water near the surface. The concentration of oxygen at the surface is $[O_2]_0$. The saturation concentrations for dissolved oxygen in equilibrium with the atmospheric concentration is $[O_2]_{sat}$. The air-sea exchange rates for CO_2 and O_2 are $E(CO_2)$ and $E(O_2)$. For $E(CO_2)$ we use the value 20 mol m^{-2} yr^{-1}, for $p_a(CO_2)$ = 340 μatm, given by *Broecker and Peng* [1982]. For $E(O_2)$ we use the value 382 mol m^{-2} yr^{-1} given by *Peng et al.* [1979]. For $[O_2]_{sat}$ we use the function of temperature given by *Weiss* [1970] with a salinity of 34.9‰. Finally, to describe the partial pressure of CO_2 dissolved in seawater, we use the relations for CO_2 chemistry described by *Peng et al.* [1987]. As a function of temperature, alkalinity, and salinity, these relations provide values for $p(CO_2)$ and several other important functions, such as the concentration of carbonate ion $[CO_3^{--}]$, that will be used to describe source terms, discussed below.

For the temperature boundary condition at the surface, we consider a linearized relation for the net flux of energy between atmosphere and ocean equal to $\lambda(T_a - T_0)$, where λ is the heat transfer coefficient (with units W $°C^{-1}$ m^{-2}). Note that our equation for temperature could be converted into an equation for heat by multiplying all temperature terms by a conversion factor ρC_p, the specific heat per m^3. Equivalently, we can express the boundary condition for air-sea heat exchange as follows:

$$J_0(T) = \left[\frac{\lambda(T_a - T_0)}{\rho C_p}\right] \quad (7)$$

For λ, a typical order of magnitude is λ = 2.2 W $°C^{-1}$ m^{-2} [see *Hoffert and Flannery*, 1985, and references therein].

The alkalinity boundary condition accounts for runoff of alkalinity J_r(TA) into the mixed layer:

$$J_0(TA) = J_r(TA) \quad (8)$$

a value determined by model calibration described in section 3.1.

Biospheric interactions occur as molecular species become incorporated into biomass during photosynthesis, redistribute themselves by transport processes, and decay by oxidation and dissociation. We distinguish two types of biomass: BO and BI represent organic soft tissue, and inorganic shell material fixed in particles that settle at velocity v through the water column. The molar amounts of nitrate, carbon, and oxygen fixed as organic biomass forms [BO] relative to that of phosphate are parameterized by a set of ratios, the Redfield ratios (RN, RC, and RO, respectively), that characterize their average concentrations in organic soft biomass. The Redfield ratios are not constant throughout the ocean but serve as a general guide to the relative effects on various chemical species. As an example, *Bolin* [1977] gives (RN, RC, RO) = (16, 106, −138). In this case the net photosynthesis reaction is

$$138O_2 + (CH_2O)_{106}(NH_3)_{16}(H_3PO_4) \rightarrow 106CO_2 + 122H_2O + 16HNO_3 + H_3PO_4 \quad (9)$$

and respiration is the reverse reaction expression.

In keeping with previous studies of the oceanic carbon cycle and biological cycle, not all nutrients upwelling in the ocean column reach the mixed layer and are incorporated into settling forms of biomass, BO and BI. A fraction of the nutrients is convected out of the ocean column before it reaches the mixed layer. This effect is an implication of the decrease in the upwelling velocity beneath the mixed layer which is reflected in the results of ocean general circulation models [*Hoffert*, 1989]. One simple model for this effect is the "Oeschger" model [see *Broecker and Peng*, 1982, p. 527], where a flow of water is taken out of the ocean column at an intermediate depth so as to fit measured depth composition profiles. Alternatively, *Joos and Siegenthaler* [1989] used diffusive coupling between the ocean column and high-latitude ocean boxes. Perhaps more direct would be to specify a variable upwelling velocity, so that its results would more clearly compare to the results of general circulation models. Nutrients that do arrive in the mixed layer can also form dissolved organic biomass [cf. *Sugimura and Suzuki*, 1988], some of which will reappear as decay products in high-latitude and bottom water. In this study we opt to retain the simplicity of a constant upwelling-diffusion model for the ocean column and account for these effects by introducing a sink of components in the mixed layer equal to parameter F times the sink due to photosynthetic production of particulate organic biomass, S_{Phot}.

We specify the source terms for BO as depth dependent rates for photosynthetic production, S_{Phot} and oxidation decay S_{Dec} of biomass BO, where these rates are scaled to units of equivalent moles of phosphate. For the inorganic biomass BI we assume that photosynthetic production occurs at a fixed ratio RI = BI/BO. Production of BI occurs by the reaction

$$Ca^{++} + CO_3^{--} \rightarrow CaCO_3 \quad (10)$$

Whenever the ambient concentration of carbonate ion $[CO_3^{--}]$ (which, like $p(CO_2)$, is a function of T, TA, and TC) falls below its critical value [*Broecker and Takahashi*, 1978],

$$[CO_3^{--}]_{crit} = 47.5\ \mu mol/kg\ e^{z/6250\ m} \quad (11)$$

then $CaCO_3$ (calcite) dissociates by the reverse of reaction (10). We, however, neglect the rate of dissociation during the time that BI particulates settle down through the water column.

Finally, we must specify the rates for production and oxidation of biomass. The rate at which organic biomass is formed in the mixed layer, measured in units of phosphate fixed, is taken to be proportional to both nutrients, phosphate and nitrate:

$$S_{Phot} = k_{Phot}[P][N] \quad z < H$$
$$S_{Phot} = 0 \quad z \geq H \quad (12)$$

where k_{Phot} is the rate coefficient for new production. The rate of decay S_{Dec} is equal to the concentration of organic biomass [BO] times k_{Dec}, the rate coefficient for decay:

$$S_{Dec} = k_{Dec}[BO] \quad (13)$$

Oxygen is consumed in the decay process. When, locally, all oxygen is consumed (anoxic conditions), decay occurs

through other mechanisms and, in this regime, model results will not be calculated.

From the assembled rates for photosynthetic production of biomass and decay by oxidation and dissociation, we can define the source terms for all components:

$$S_{BO} = S_{Phot} - S_{Dec} \qquad S_{BI} = RI \cdot S_{Phot}$$
$$S_P = -S_{Phot}(1 + F) + S_{Dec} \qquad S_{TC} = RC \cdot (S_P - S_{BI})$$
$$S_N = RN \cdot S_P \qquad S_{TA} = -S_N - 2S_{BI} \cdot RC$$
$$S_O = RO \cdot S_P \qquad S_T = 0 \tag{14}$$

The expression for change in alkalinity recognizes the effect of nitrate and Ca^{++} on TA [cf. *Broecker and Peng*, 1989].

2.2. *Model Calibration*

Working with a 1D model, we cannot hope to capture the full complexity that controls the actual distribution of species throughout the 3D world oceans. However, our intent is to capture generic behavior in a manner that allows us to gain insight and to analyze the role of biospheric interactions. Tuning of the model to reproduce average vertical profiles is carried out to put the model in the proper regime for analysis. Actual values for adjustable transport coefficients and rate parameters in the model are not meant to be quantitatively precise. The intent is to be accurate enough to draw useful conclusions about trends.

In the mixed layer, transport is dominated by eddy diffusivity and particulate settling, not upwelling. The upwelling velocity w equals zero. The eddy diffusivity is taken to be large enough, $K = 30{,}000\ m^2\ yr^{-1}$, so that there are not significant variations of concentrations or temperature through the mixed layer of depth $H = 100$ m.

Below the mixed layer, the upwelling velocity and vertical eddy diffusivity are approximated to be constant and have the values ($w = 4\ m\ yr^{-1}$ and $K = 2000\ m^2\ yr^{-1}$, respectively) suggested by *Hoffert et al.* [1980], who adjusted these parameters to fit globally averaged temperature profiles. In Figure 2*a*, model predictions (with these values of transport coefficients, an atmospheric temperature of 22°C and an upwelling water temperature of 1°C) are compared to *Levitus'* [1982] data. Upwelling bottom water and atmospheric concentrations of the trace species being followed are taken to be

$$p_a(CO_2) = 285\ \mu atm \qquad [TA]_b = 2410\ \mu mol\ kg^{-1}$$
$$[TC]_b = 2310\ \mu mol\ kg^{-1} \qquad [P]_b = 2.14\ \mu mol\ kg^{-1}$$
$$[O_2]_b = 210\ \mu mol\ kg^{-1} \tag{15}$$

which are comparable to world-averaged values. The remaining parameters, the kinetic constants for production (RI and k_{Phot}) and decay (k_{Dec}) of biomass and the bypassing of nutrients around the mixed layer (F), are adjusted so that model results match measured vertical profiles of trace chemical species.

Figure 2*b* shows the fit between phosphate and oxygen profiles to GEOSECS data [*Takahashi et al.*, 1981]. Upwelling concentrations listed above were taken from measured concentrations at 5500 m. The other parameters influencing the profiles are k_{Phot}, k_{Dec}, and F. The conversion of nutrients in the mixed layer is nearly complete; the production rate constant $k_{Phot} = 0.16\ \mu mol^{-1}\ kg\ yr^{-1}$ is chosen to cause this. The remaining two constants are chosen to match the minimum oxygen and maximum phosphate concentration magnitudes of the data. Values of $F = 1.04$ and decay rate constant of $k_{Dec} = 50.\ yr^{-1}$ are found to match these. These parameters also determine the vertical profile of particulate organic biomass, shown in Figure 2*d*, which is similar to the assumed exponential decay (with depth) of BO assumed by *Volk* [1984] and *Shaffer* [1989], and roughly matches the scattered data on organic particulate flux derived from sediment trap measurements which best fits a profile proportional to $1/z$ [*Berger et al.*, 1987].

Figure 2*c* shows the fit between titration alkalinity and total carbon profiles to GEOSECS data [*Takahashi et al.*, 1981]. The ratio of calcium carbonate to organic biomass production was adjusted to RI = 0.44 to match the alkalinity observed in the mixed layer, $[TA]_m = 2314\ \mu mol/kg$. The concentration of total carbon in the bottom water $[TC]_b$ was chosen to be about 25 $\mu mol/kg$ greater than that observed in order to lead to a lysocline depth of 3.4 km, in order to agree with the estimate of the fraction of the seafloor subject to calcite accumulation given by *Li et al.* [1969].

Figure 2*e* shows the carbonate concentration compared to the critical concentration. Even though the mixed layer of the ocean is supersaturated by more than a factor of 3, there is currently no significant abiotic precipitation of calcium carbonate in the open ocean. Interestingly, the high concentration of carbonate ion in the mixed layer is due to pumping of carbon from the mixed layer; if biota did not pump carbon deep into the ocean, the degree of supersaturation would not be as great. Perhaps the biota finds advantage in a high degree of supersaturation in the mixed layer.

3. MODEL RESPONSES OF THE WORLD OCEAN SYSTEM

The model for the ocean system is shown in Figure 1. The 1D upwelling model is linked with polar ocean and bottom water box models connected at the mixed layer and the ocean floor, where the conservation equations for the box models lead to upwelling concentrations of TA, TC, O_2, and P (to which N is made proportional). The box and 1D models together make up the ocean system model.

The world external to the ocean system communicates or influences the ocean system directly by runoff of P (and N) and TA, and by transport through the atmosphere of TC and O_2, and by, for example, climatic influence on ocean circulation (modeled here by simply an upwelling velocity w, an eddy diffusivity K, and the fraction F of nutrients that bypasses the mixed layer) and atmospheric temperature T_a. Internal to the model are parameters which determine the strength and direction of biological activity in the simple model of marine biota. Among these parameters are the Redfield ratios (RN, RO, and RC), the fraction of calcium carbonate secreted per unit of particulate organic biomass (RI), kinetic rate constants for the photosynthetic production and biological decay of biomass (k_{Phot} and k_{Dec}), and the settling velocity of particulate biomass (v). Since the Redfield ratios and fraction RI are presumably encoded in the DNA of marine organisms, it is possible that these biological

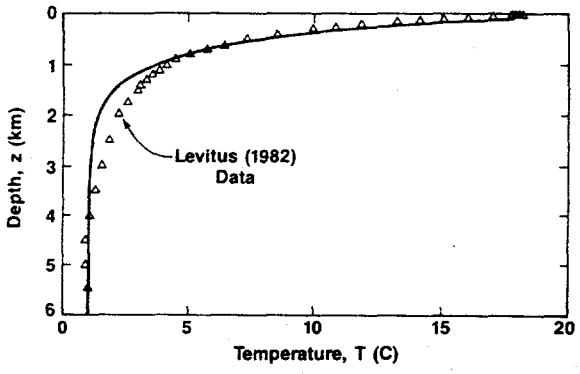


Fig. 2a

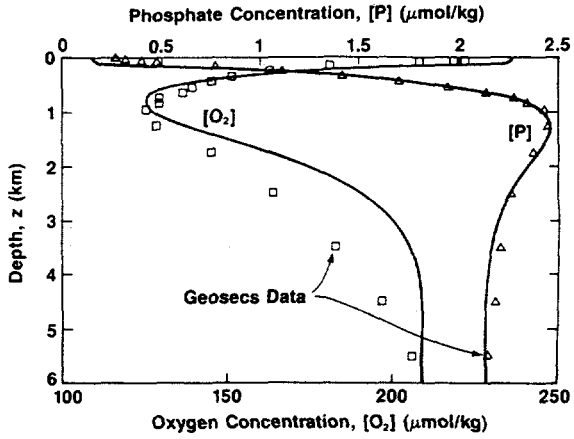


Fig. 2b

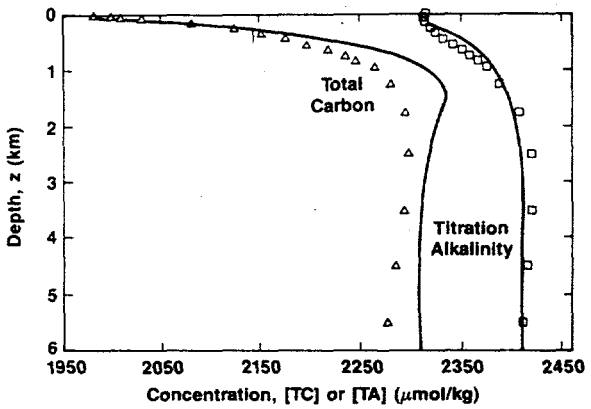


Fig. 2c

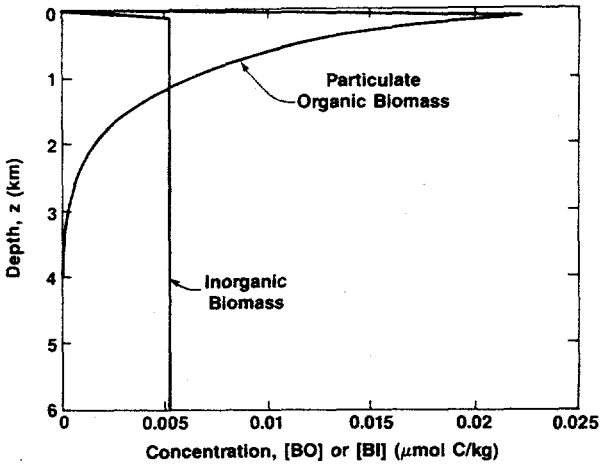


Fig. 2d

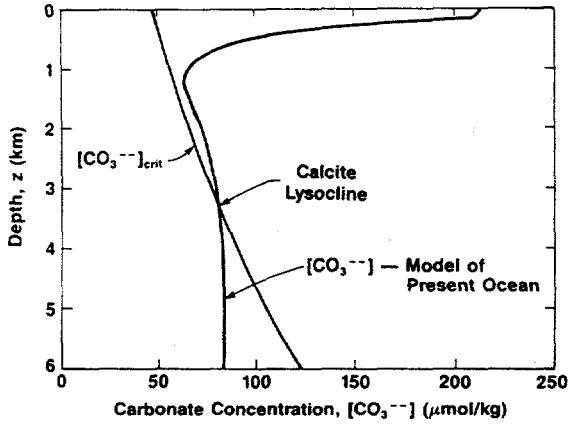


Fig. 2e

Fig. 2. Comparison of model profiles for (*a*) temperature, (*b*) phosphate and oxygen concentrations, (*c*) and total carbon and titration alkalinity to measured profiles. The 1D model implies the profile of (*d*) organic and inorganic particulate biomass and (*e*) carbonate ion concentration.

parameters can be altered by natural selection. If so, the potential exists for the marine biosphere to play an active role in regulating the compositional structure of the oceans and atmosphere [cf. *Lovelock*, 1988].

3.1. *Model Description and Calibration*

The sedimentation rate of calcite particulates $J_{sed}(BI)$ is approximated to be the net flux of inorganic biomass BI to the ocean floor over the region Ω in the depth dimension over which the carbonate ion concentration is above its critical value:

$$J_{sed}(BI) = \int_{\Omega} RC \cdot [BI] \rho v \, \partial A = \int_{\Omega} RC \cdot [BI] \rho v \frac{\partial A}{\partial z} \, dz \quad (16)$$

If there is a single lysocline at depth $z = z_l$, then the region Ω would be that above the lysocline, $0 < z < z_l$. The depth distribution $A(z)$ of the ocean floor is approximated by the

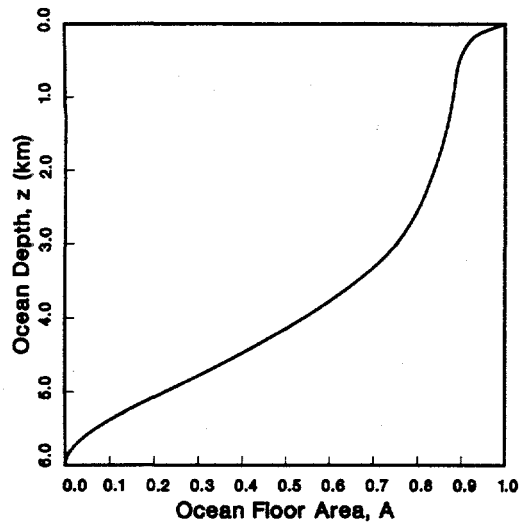


Fig. 3. Hypsometric profile of the model ocean floor.

function shown in Figure 3, which passes through the ocean hypsometry data given by *Menard and Smith* [1966]. The implied sedimentation rate of inorganic carbon is equal to $J_{sed}(BI)$. The implied sedimentation rate of alkalinity, due to sedimentation of calcium, is equal to twice $J_{sed}(BI)$.

The sedimentation rate of organic biomass particulates $J_{sed}(BO)$ is proportional to the net flux of organic biomass BO to the ocean floor:

$$J_{sed}(BO) = \gamma \int_0^{6000\text{ m}} [BO]\rho v\, \partial A$$

$$= 0.1 \int_0^{6000\text{ m}} [BO]\rho v \frac{\partial A}{\partial z}\, \partial z \qquad (17)$$

The proportionality constant $\gamma \simeq 0.1$ [*Berger et al.*, 1987] represents the fraction of biomass that is not consumed by benthic respiration. An alternate assumption used in 1D models without hypsometry is to assume the highly reactive fraction is entirely consumed, and that some specified part of the low-reactivity component is buried [*Shaffer*, 1989]. Thus the general consistency of our model with observations of organic carbon burial rates needs to be viewed with caution, as the precise mechanism of this burial is still at issue. In any case, the flux $J_{sed}(BO)$ to the ocean floor is, in oceans similar to the present one, small compared to the upwelling fluxes of components in the model. Organic biomass sedimentation, nonetheless, accounts for roughly half the rate of phosphate removal from the ocean system, and is shown below to strongly influence biological activity. The other half of the removal rate is by incorporation of a small concentration of phosphate on inorganic biomass particulates [*Broecker and Peng*, 1982, p. 29]. The implied sedimentation rate of inorganic carbon is equal to $J_{sed}(BI)$. The implied sedimentation rates of organic P, N, TC, and O are equal to $J_{sed}(BO)$ times the component's Redfield ratios.

The only route for alkalinity and phosphate (and nitrate again tracks phosphate) into the ocean system is by runoff from the external world. In steady state this is equal to the components' sedimentation rates. The upwelling concentrations $[P]_b$ and $[TA]_b$ are varied to satisfy the equations

$$J_r(P) = J_{sed}(P) = J_{sed}(BO) \qquad (18)$$

$$J_r(TA) = J_{sed}(TA) = -RN\, J_{sed}(BO) + 2\, J_{sed}(BI) \simeq 2\, J_{sed}(BI) \qquad (19)$$

For the present ocean conditions, to which the 1D model was calibrated in section 2.2, the model runoff rates are

$$J_r(P) = 56 \cdot \frac{\mu\text{mol}}{\text{m}^2\text{ yr}} \qquad (20)$$

$$J_r(TA) = 0.20 \frac{\text{mol}}{\text{m}^2\text{ yr}} \qquad (21)$$

Water that travels from the mixed layer to polar surface waters cools from temperature T_m to temperature $T_p = T_b$. Species feed the polar surface box from both water from the mixed layer plus that due to water which bypassed the mixed layer or formed dissolved organic biomass. The flux of phosphate which bypasses the mixed layer and is convected into the polar ocean or bottom water boxes is

$$J(BP) = \int_0^{100\text{ m}} F\rho S_{Phot}\, \partial z \qquad (22)$$

A fraction α of the flux $J(BP)$ appears in the polar surface waters. The remainder $(1 - \alpha)$ appears in the bottom water, presumably transported there by settling of organic particulates formed in the polar surface waters [cf. *Broecker and Peng*, 1989]. In the polar ocean the phosphate concentration is that of the mixed layer, $[TA]_m$, plus that due to $J(BP)$:

$$[P]_p = [P]_m + \alpha\, J(BP)/(\rho w) \qquad (23)$$

The alkalinity of water that bypasses the mixed layer, however, is expected to be greater than that of the mixed layer and is not related just to the Redfield ratio for organic biomass. To give the high observed alkalinity of the polar surface water, a fraction β of the dissolution products of inorganic biomass is added to the polar box to give:

$$[TA]_p = [TA]_m - RN\alpha J(BP)/(\rho w) + \beta \int_{\Omega'} \left\{ 2RC\, v[BI] \frac{\partial A}{\partial z}\, \partial z \right\} \Big/ w \qquad (24)$$

where Ω' is the region in depth z where the carbonate ion concentration is less than the critical value (the region below the lysocline, if only one lysocline exists). Total carbon and oxygen at the ocean surface are near equilibrium with the atmosphere. The polar box is modeled to be thoroughly ventilated. The total carbon concentration in the polar surface water, $[TC]_p$, is that which gives a carbon dioxide partial pressure equal to that of the atmosphere:

$$pCO_2([TC]_p, [TA]_p, T_p) = p_a(CO_2) \qquad (25)$$

Oxygen also is near equilibrium with the atmosphere in the upper polar ocean waters:

$$[O_2]_p = [O_2]_{sat} \qquad (26)$$

at temperature T_p.

Deep in the polar ocean and in the benthic layer, represented by the bottom water box, the remainder of $J(BP)$ appears as decay products and inorganic biomass that arrives below the lysocline dissolves. Both respiration and dissolution release carbon which combines with $[TC]_p$ to produce the bottom water concentration:

$$[TC]_b = [TC]_p + RC(1 - \alpha)$$

$$\cdot J(BP)/\rho + (1 - \beta) \int_{\Omega'} \left\{ v[BI] \frac{\partial A}{\partial z} \partial z \right\} \Big/ w \qquad (27)$$

Bottom ocean water contains less oxygen than polar ocean water due to reduction of oxygen by the flux $RO(1 - \alpha) J(BP)$:

$$[O_2]_b = [O_2]_p + \frac{RO(1 - \alpha) J(BP)}{\rho w} \qquad (28)$$

And, finally, bottom ocean water contains more phosphate than polar surface water due to addition of phosphate by the flux $(1 - \alpha) J(BP)$:

$$[P]_b = [P]_p + \frac{(1 - \alpha) J(BP)}{\rho w} \qquad (29)$$

The polar surface and bottom water boxes are calibrated to present ocean conditions. A value of $\alpha = 0.51$ is found to lead to a bottom water concentration of oxygen, given by (28), that matches present ocean concentration $[O]_b = 210$ μmol/kg listed in (15). A value of $\beta = 0.75$ is found to lead to a polar alkalinity, given by (24), that matches present ocean concentration $[TA]_p = 2390$ μmol/kg [*Broecker and Peng*, 1989]. Finally, the atmospheric concentration of carbon dioxide of 285 μatm, near the preindustrial value listed in (15), is found to lead to a total carbon concentration in the bottom water equal to $[TC]_b$ given in (15).

The modeling equations for the polar surface and bottom water boxes are added to the simultaneous equations that model the ocean system. These added equations allow upwelling concentrations to be calculated from the external variables: (1) atmospheric partial pressure of carbon dioxide, $p_a(CO_2)$, (2) atmospheric partial pressure of oxygen, $p_a(O_2)$; (3) runoff rate of phosphate, $J_r(P)$; and (4) runoff rate of alkalinity, $J_r(TA)$.

The equations that model the ocean system are solved numerically. The 1D equations for the ocean column are reduced to a system of algebraic equations by the Galerkin/finite element method [e.g., *Zienkiewicz*, 1977]. Newton's method is used to solve the system of algebraic equations. Continuation methods are required to provide initial guesses of solutions for Newton iteration. Calculations were carried out on the Cray X-MP/14se computer at the Corporate Research Laboratories of Exxon Research and Engineering Company. Solution required from 3 CPU seconds to as long as 2 min, depending on parameter values.

3.2. *Comparison of Model Carbon Fluxes With Observations*

A central feature of our model is that the vertical fluxes of organic and inorganic carbon, and burial rates, are linked to their vertical compositional structure. Model profiles were calibrated to observed profiles above, thereby determining all model coefficients. Comparison of model biological productivities and burial rates to observed values provides an internal consistency check of the model. The main fluxes of interest are the organic carbon primary and "new" productivity of the surface ocean and the burial rates of organic and inorganic carbon.

In a fairly recent review of the literature, *Sundquist* [1985, Table 9] finds total marine primary productivities estimated by the ^{14}C uptake technique using area-weighted averaging over different ocean ecosystems in the range $\simeq 20$–40×10^{15} g C yr^{-1}. The oft-quoted *Whittaker and Likens* [1975] marine primary productivity of 55×10^{15} g C yr^{-1} [*Ajtay et al.*, 1979; *Valiela*, 1984, Table 1-3] is for "equivalent dry matter" corresponding to 25×10^{15} g C yr^{-1} for a dry weight biomass of 45% carbon [*Sundquist*, 1985]. However, some authors [*Sorokin*, 1978; *Jenkins*, 1982] have suggested that radiocarbon methods may underestimate primary production by as much as a factor of 1.5–2.0, which would put world ocean values in the range 30–80×10^{15} g C yr^{-1}. The high end of these estimates has the marine biosphere about as productive as the land, albeit with a much higher turnover rate. With world ocean primary productivity in the range 20–80×10^{15} g C yr^{-1}, averaging over the world ocean area, 3.61×10^{14} m^2, gives mean primary production flux of the world ocean photic zone, PP, in the range 55–200 g C m^{-2} yr^{-1}.

More pertinent to the present model comparison is the organic carbon flux sinking to deeper levels as fecal pellets or marine snow: the so-called new production flux, NP. In general, NP < PP because heterotrophs feeding on phytoplankton respirate part of this productivity in their own metabolism. *Sundquist* [1985] observes that organic carbon in settling particles is rapidly depleted, causing the flux to the deep sea to decrease significantly within a few hundred meters of the surface. The fraction of primary production which sinks is itself a function of PP, being larger when PP is large (eutrophic waters) than when PP is small (oligotrophic waters), as well as being a function of the depth to which NP is attributed. *Eppley* [1984] has developed the correlation $NP \simeq PP/2$ when $PP > PP_1$, and $NP \simeq (PP_1/2)(PP/PP_1)^2$ when $PP \leq PP_1$, where $PP_1 = 200$ g C m^{-2} yr^{-1}, in which case the "observed" range of world ocean new production (including the uncertainties of the ^{14}C method) is quite broad: $NP \simeq$ 8–110 g C m^{-2} yr^{-1}. This suggests some 15–50% of primary production falls out of the photic zone as organic detritus. In comparison, our standard model calibration predicts a total integrated flux of organic phosphorous leaving the surface layer in the form of BO as 0.007 mol P m^{-2} yr^{-1}. The corresponding "new production" of organic biomass in our model is NP = 9. g C m^{-2} yr^{-1}, which is at the low end of the "observed" range.

Sedimentation rates of phosphorous and total alkalinity predicted by the calibrated model are linked to the runoff rates of these quantities. Since our calibration on concentration profiles did not use actual runoff data, the computed sedimentation rates provide useful consistency tests. Estimates of the global hydrological cycle indicate mean evaporation and precipitation fluxes of water over the world's oceans at 140 and 127 cm yr^{-1}, respectively [*Budyko*, 1982, Table 3.5]. Since the surface ocean water budget is balanced by runoff from the world's rivers, the freshwater flow rate

$$R = \left(1.40 \frac{\text{m}}{\text{yr}} - 1.27 \frac{\text{m}}{\text{yr}} \right) \times 3.61 \times 10^{14} \text{ m}^2 \times 1000 \frac{\text{kg}}{\text{m}^3} \simeq 4.7 \times 10^{16} \frac{\text{kg}}{\text{yr}} \qquad (30)$$

is considered to carry the entire load of dissolved and particulate sediments from the weathering of continental rocks to the oceans.

In particular, the influx of dissolved and particulate P in runoff is balanced by the outflux of P in organic biomass buried in sediments. (Inorganic precipitation of phosphorous as phosphorite is not thought to be a major factor in today's oceans, but may have been a significant sink for marine phosphorous over much of Phanerozoic time [*Holland*, 1978, p. 218].) Global mean elemental P concentrations in riverine runoff are given by *Broecker and Peng* [1982, pp. 26–27] as $[P]_r \simeq 1.3$ μmol P kg^{-1}. The input of P to the sea is thus $[P]_r R \simeq 6.1 \times 10^{10}$ mol yr^{-1} $= 2 \times 10^{12}$ g P yr^{-1}, the value cited by *Holland* [1978, p. 217]. Alternatively, *Froelich et al.* [1982; see also *Broecker and Peng*, 1982, p. 29] estimate a burial rate of phosphorous in organics of 1.4×10^{10} mol P yr^{-1}. This leads to a global mean phosphate burial flux per unit ocean of

$$(1.4\text{–}6.1) \times 10^{10} \frac{\text{mol P}}{\text{yr}} \times \frac{1}{3.61 \times 10^{14}\ \text{m}^2} = (0.4\text{–}1.7) \times 10^{-4} \frac{\text{mol P}}{\text{m}^2\ \text{yr}} \quad (31)$$

For comparison, our standard model calibration predicts a phosphorous burial rate (20), $J_{sed}(P) = 0.6 \times 10^{-4}$ mol P m^{-2} yr^{-1}, which is within the "observed" range and surprisingly close, considering the many approximations involved. Both model and the observations indicate some 0.1–0.4% of the primary production and 1–2% of new production are buried as unoxidized organic carbon in ocean sediments. Since both modeled and observed J(BO) scale with C/P, a different organic carbon to phosphate ratio characteristic of sediments would not affect their relative values, but would influence the J(BO)/P and J(BO)/NP ratios. *Holland* [1978, p. 218], for example, indicates buried organic matter may only contain half as much nitrogen and phosphorous as average phytoplankton, such that $(C/P)_{sed} \simeq 200$.

The observed total alkalinity burial rate is expressible in moles alkalinity per year per unit area of ocean, and is readily approximated from the riverine concentration of calcium ion concentration [*Broecker and Peng*, 1982, p. 26], $[Ca^{2+}]_r \simeq 360$ μmol kg^{-1}:

$$3.6 \times 10^{-4} \frac{\text{mol Ca}^{2+}}{\text{kg}} \frac{4.7 \times 10^{16}\ \text{kg yr}^{-1}}{3.61 \times 10^{14}\ \text{m}^2} 2 \frac{\text{TA}}{\text{Ca}^{2+}} = 0.09 \frac{\text{mol TA}}{\text{m}^2\ \text{yr}} \quad (32)$$

This is roughly half the value obtained in the model calibration (21) of 0.20 mol m^{-2} yr^{-1}. For a ("Broecker") ratio of organic to inorganic particulate carbon leaving the photic zone of Br = 1/RI = 2.3, our model gives a ("Ronov") ratio of organic to inorganic carbon buried in sediments of Ro = 0.06. The small value of Ro is attributed to the higher fraction of BO respired in the water column compared to the smaller fraction of BI dissolved on the ocean floor (below the lysocline). This value of Ro is fairly consistent with that observed in recent sedimentary rock deposits [*Sundquist*, 1985].

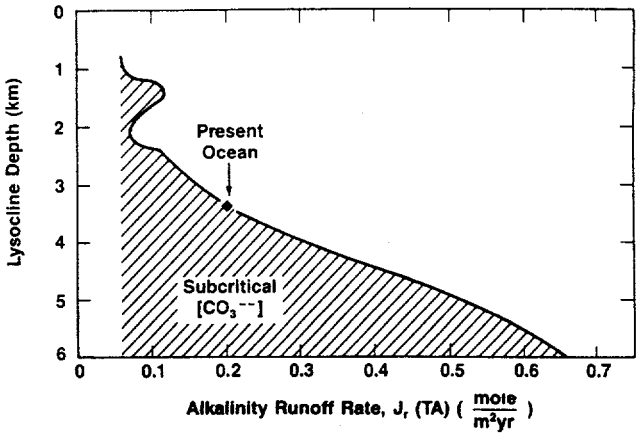


Fig. 4*a*

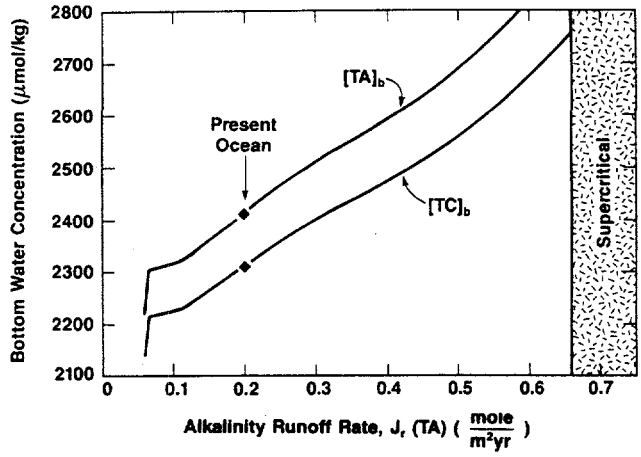


Fig. 4*b*

Fig. 4. Alkalinity runoff rate influences (*a*) lysocline depth and (*b*) bottom water concentrations of total carbon and alkalinity.

3.3. *Response of Ocean System to Changes in External Variables*

Solution of the ocean system model predicts what the response of ocean composition and biological productivity would be to changes in external variables, ocean circulation, and biologically determined parameters.

First, consider the response of the ocean to a change in rate of runoff of alkalinity into the ocean (which, perhaps, originates as the product of weathering of calcium-silicate minerals). The transient response of the ocean to an increase of J_r(TA) would be to have [TA] increase. Biological activity, in the model, does not depend on either TA or TC, and following from this, the $[O_2]$ and [P] in the ocean are not altered. Since the production rate of calcium carbonate particulates is set by the (unaltered) organic biomass productivity, the flux of BI out of the mixed layer is unaltered. To compensate for larger input of alkalinity, in a steady state, the rate of sedimentation of alkaline species must rise. In the model this can only occur by increasing the fraction of the ocean floor over which the carbonate ion concentration is above its critical value.

As J_r(TA) increases from the present ocean condition estimated by the model (see (21)) of 0.20 mol m^{-2} yr^{-1}, the lysocline moves deeper, as shown in Figure 4*a*, where [TA]

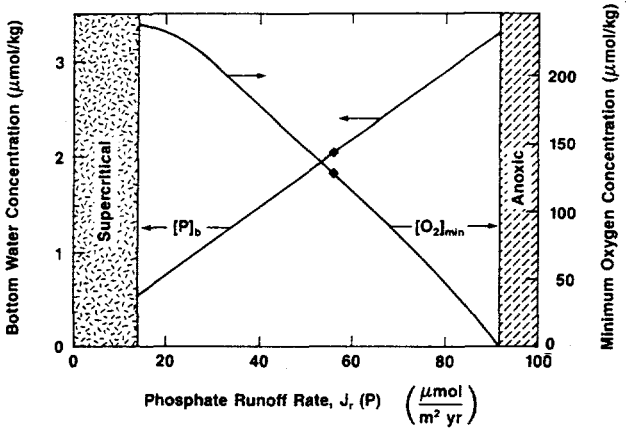


Fig. 5a

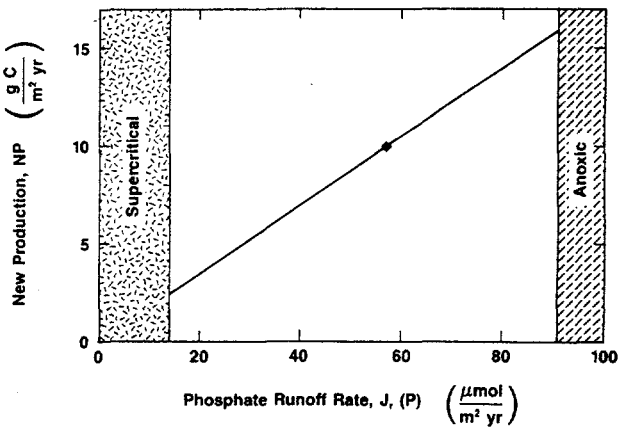


Fig. 5b

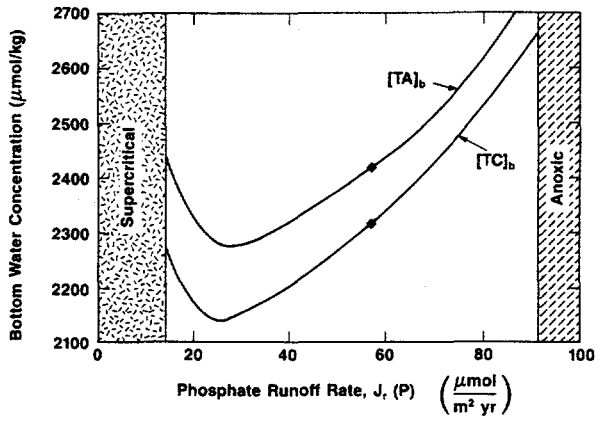


Fig. 5c

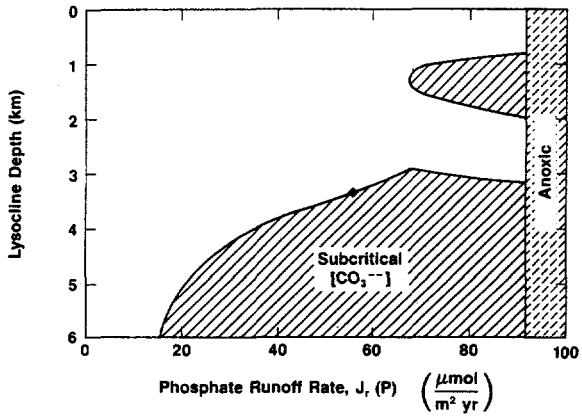


Fig. 5d

Fig. 5. Phosphate runoff rate drives (*a*) phosphate and oxygen concentrations, (*b*) new productivity, (*c*) bottom water concentrations of total carbon and alkalinity, and (*d*) lysocline depth.

and [TC] are nearly depth independent and equal to their bottom water concentrations. A deeper lysocline implies a higher carbonate ion concentration deep in the ocean, while at the surface of the ocean, $p(CO_2) \simeq p_a(CO_2)$. Carbonate chemistry for typical bottom water conditions, $T = 5°C$ and $p(CO_2) = 490$ μatm, gives

$$\frac{\partial[TA]}{\partial[CO_3^{--}]} = 12. \qquad \frac{\partial[TC]}{\partial[CO_3^{--}]} = 11. \tag{33}$$

as is shown graphically by *Baes* [1982], which explains why [TA] and [TC] change in tandem in Figure 4*b* with changes in $[CO_3^{--}]$ which are a consequence of changes in lysocline depth.

For J_r(TA) less than 0.20 mol m^{-2} yr^{-1} there are two depth regions which have a subcritical concentration of $[CO_3^{--}]$. This is a result of multiple intersections of $[CO_3^{--}]$ and the $[CO_3^{--}]_{crit}$ curves in Figure 2*e* which occur if the $[CO_3^{--}]$ curve is moved to the left, due to lower bottom water carbonate concentrations. Multiple lysoclines are observed in regions of the North Pacific [*Broecker and Takahashi*, 1978] and in model predictions of transient responses of the water column [*Sundquist*, 1986]. Evident in Figure 2*e* is the high sensitivity in lysocline depth to small changes in carbonate ion concentration.

As shown in Figure 4*b*, both [TA] and [TC] increase as the runoff rate increases for J_r(TA) near the present ocean value. When the runoff rate reaches 0.66 mol m^{-2} yr^{-1} the lysocline reaches the deepest point of the model ocean, $z_l = B$. At runoff rates higher than this value, the biological production of calcium carbonate cannot keep up with the input of calcium to the ocean. The alkalinity and total carbon of the ocean would then rise until an abiological mechanism for the formation of calcium carbonate became active or, alternatively, the biota could react by increasing the ratio of calcium carbonate to organic biomass formed, RI. In any case, the steady state model defined in section 3.1 does not have a solution beyond this point.

If abiotic calcium carbonate sedimentation occurs, it would occur at the ocean floor where the carbonate concentration is the highest: in shallow portions of the ocean, the continental shelves. The biota could thus influence the ratio of sedimentation on the continental shelves to sedimentation in the deep ocean. The subduction time for the continental shelves is much longer than that of the deep ocean floor, and so the biota can influence the residence time of carbon in the Earth's crust and therefore the amount of carbon in the active carbon cycle.

If phosphate runoff rate J_r(P) is varied from the present ocean conditions, the phosphate flux entering the mixed layer would vary, altering biological activity. Nearly all of

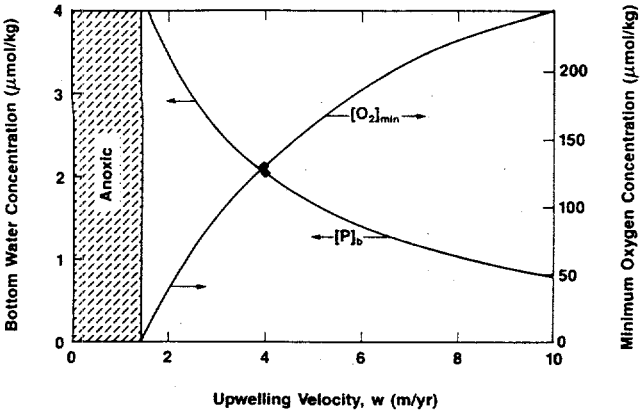


Fig. 6*a*

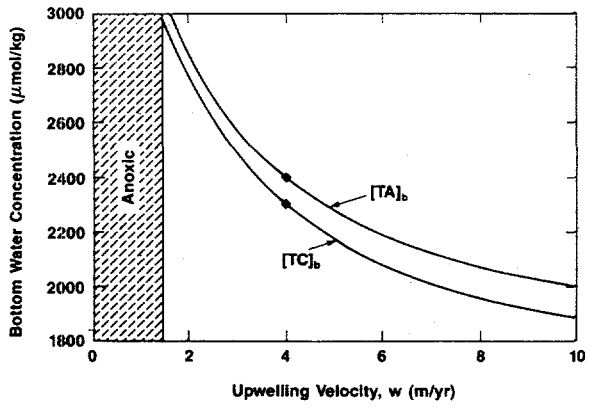


Fig. 6*b*

Fig. 6. Upwelling velocity drives (*a*) phosphate and oxygen concentrations and (*b*) bottom water concentrations of total carbon and alkalinity.

the phosphate that arrives in the mixed layer, and does not escape to the polar oceans, forms particulates which decay before reaching the ocean floor, thereby recycling the phosphate to the mixed layer. At high phosphate runoff rates, the flux of organic particulates down from the mixed layer is greater, and decay depletes oxygen at $z \simeq 1$ km, as shown in Figure 5*a*. At $J_r(\mathrm{P}) = 92.$ μmol m^{-2} yr^{-1} the ocean below the mixed layer becomes anoxic. In an anoxic environment, alternative decay mechanisms, not included in this model, become active, as shown by *Shaffer* [1989]. As the phosphate runoff rate increases, biological activity increases (see Figure 5*b*). Alkalinity and total carbon again move in tandem (see Figure 5*c*). At low phosphate runoff rate, the production rate of $CaCO_3$ is small and the lysocline is located near the deepest part of the ocean to allow nearly all the $CaCO_3$ produced to form sediments (see Figure 5*d*). Below a phosphate runoff rate of 16. μmol m^{-2} yr^{-1} the ocean is entirely above the critical calcium carbonate saturation, and the model does not contain a mechanism to remove the alkalinity runoff into the ocean; no steady solution exists.

As $J_r(\mathrm{P})$ increases from 16. μmol m^{-2} yr^{-1}, the lysocline must move steeply upward (see Figure 5*d*); TC and TA decrease and make the ocean less saturated in $CaCO_3$ (see Figure 5*c*). At even higher phosphate runoff rates, the biological activity is high enough to pump carbon (and alkalinity) out of the mixed layer; TA and TC increase to maintain carbonate ion concentration and $p(CO_2)$ at the air/sea interface. Interestingly, there is a minimum in TA and TC with respect to phosphate runoff rate. Furthermore, multiple lysoclines, apparent in Figures 5*d*, exist over a range of phosphate runoff rates greater than in present oceans.

The upwelling velocity of bottom water determines both the time required for nutrients pumped down in the ocean by particulate biomass settling to return to the mixed layer and also the amount of oxygen that is mixed deep into the ocean; together these mechanisms imply that higher upwelling velocity results in the unaltered biological productivity with higher $[O_2]$ and lower [P] (see Figure 6*a*). At upwelling velocities below 1.5 m yr^{-1} the ocean becomes anoxic. Alkalinity and total carbon concentrations decrease monotonically with increasing upwelling w, as shown in Figure 6*b*, in order to keep the lysocline at an unaltered depth.

The partial pressure of carbon dioxide of the atmosphere affects the alkalinity and carbon concentration of the ocean. The rate of air-sea exchange is rapid enough that $p(CO_2)$ of ocean surface water is near that of the atmosphere. If ocean alkalinity were fixed, a small relative change in $p(CO_2)$ would result in a smaller relative change in TC; this is known as the buffer effect. Over a longer time scale, however, the alkalinity of the ocean would rise until the lysocline regained its position before the change in $p(CO_2)$. Therefore the carbonate ion concentration at the lysocline would regain its old value while $p(CO_2)$ at the surface was changed. This results in a much larger change in both TC and TA (see Figure 7). The sensitivity of TC and TA comes largely from local carbonate chemistry [*Peng et al.*, 1979; *Baes*, 1982], which for typical deep ocean conditions, $T = 5°\mathrm{C}$ and $[CO_3^{--}] = 101$ μmol/kg, gives

$$\frac{\partial[\mathrm{TA}]}{\partial p(CO_2)} = 2.14 \frac{\mathrm{mol}}{\mathrm{kg\ atm}} \qquad \frac{\partial[\mathrm{TC}]}{\partial p(CO_2)} = 2.24 \frac{\mathrm{mol}}{\mathrm{kg\ atm}} \tag{34}$$

The buffer factor, defined as $\{\partial p(CO_2)/\partial[\mathrm{TC}]\}\{[\mathrm{TC}]/p(CO_2)\}$, calculated from results shown in Figure 7 is 19. when [TA] is held fixed and 2.6 when [TA] adjusts to its steady state profile, which results in the present ocean lysocline position and $J_r(\mathrm{TA})$.

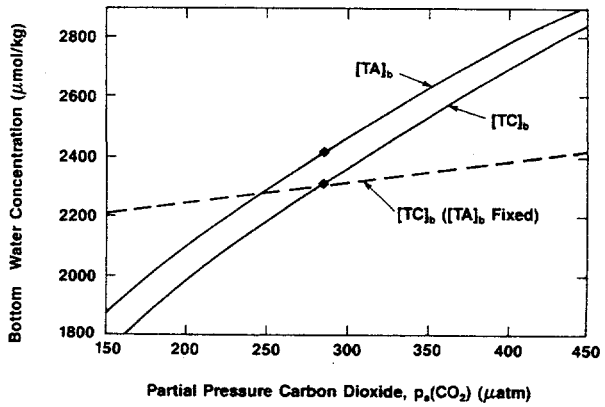


Fig. 7. Carbon dioxide partial pressure in the atmosphere influences bottom water concentrations of total carbon and alkalinity (solid curves) when $J_{sed}(\mathrm{TA})$ is held fixed and therefore lysocline depth is held fixed. If, alternatively, the alkalinity concentration profile were held fixed to present ocean values (dashed curve), then $[\mathrm{TC}]_b$ would be much less sensitive to $p_a(CO_2)$.

Evolution and natural selection of marine biota could alter parameters in the marine biota model; for example, the ratio RI of inorganic to organic biomass produced in the mixed layer, Redfield ratios RC, RN, and RO, or the production of aragonite instead of calcite. Each of the biota traits affects ocean composition.

If the inorganic to organic ratio RI is increased, the increased production of calcite particulates causes the lysocline to rise (Figure 8*a*) and TA and TC to decrease (Figure 8*b*), in order that the sedimentation rate of alkalinity (calcium) remains constant. The organic productivity, along with phosphate, nitrate, and oxygen concentration profiles, remains unaltered. If RI decreases below 0.08, the lysocline reaches the ocean floor; the ocean is supersaturated everywhere. When this happens, $[TC]_b$ and $[TA]_b$ would rise until abiotic sedimentation of $CaCO_3$ commenced.

Another biota trait that has great influence over ocean composition in our model is the Redfield ratio of carbon to phosphorous, RC (and implicitly, $-RO = RC + 2RN$); see Figure 9. As RC increases, the minimum concentration of oxygen decreases, until at RC = 160 the ocean becomes anoxic. As RC decreases, the lysocline sinks in order to maintain a constant burial rate of alkalinity, until at RC = 23 the entire ocean is supercritical with respect to $CaCO_3$ (Figure 9*a*).

Our marine biota/ocean transport model defines mechanisms and quantifies how marine biota influence ocean composition. For example, Table 1 shows the change in external variables or biota traits required for the ocean carbon content to increase by 720 gigatons (about 2%), which is the current amount of carbon in the atmosphere. Most notable is the small changes in RI and RC required to change ocean carbon content. Biota does have great leverage in influencing where carbon resides on Earth and, in turn, in climate variations. We plan a transient model to study the time scales over which marine biota could potentially react and influence the carbon cycle.

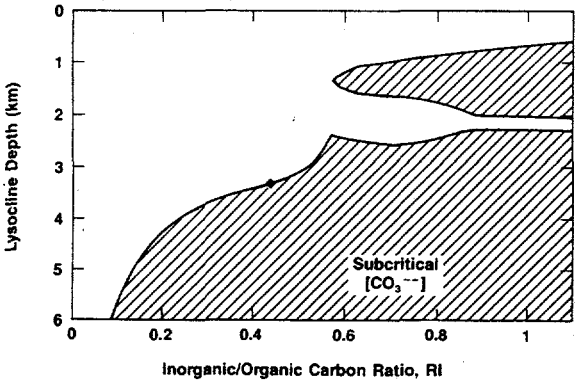


Fig. 8*a*

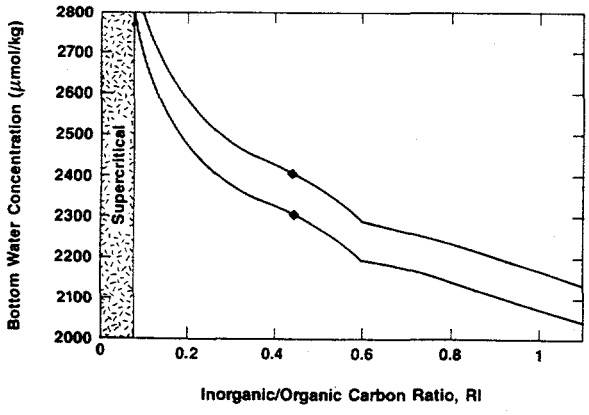


Fig. 8*b*

Fig. 8. Inorganic to organic particulate ratio, RI, influences (*a*) lysocline depth and (*b*) bottom water concentrations of total carbon and alkalinity.

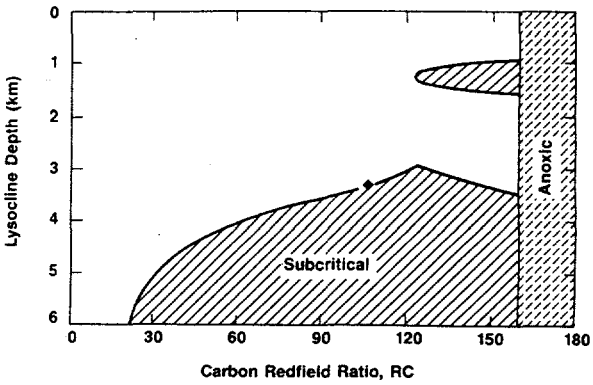


Fig. 9*a*

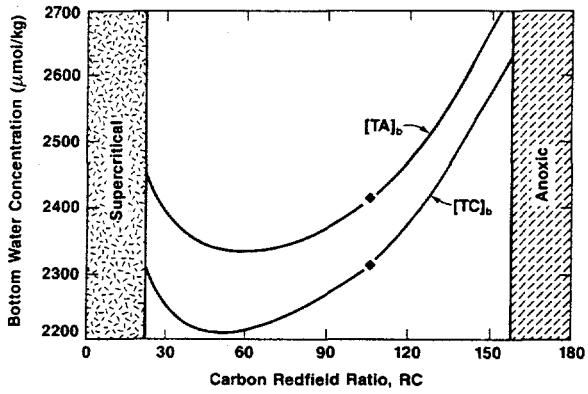


Fig. 9*b*

Fig. 9. Carbon Redfield ratio, RC, influences (*a*) lysocline depth and (*b*) bottom water concentrations of total carbon and alkalinity.

TABLE 1. Extent That an External Variable or Biota Trait Must be Changed in Order for the Model Ocean to Increase Its Carbon Content by 720 Gtons (or 2%)

	Change From Present Ocean Value, %
External variable	
$p_a(CO_2)$	+5
Alkalinity runoff	+21
Upwelling velocity	−10
PO_4 runoff	+10
Biota trait	
Inorganic/organic ratio, RI	−23
Carbon Redfield ratio, RC	+9

4. Concluding Discussion

A steady state model for ocean transport/marine biota defines mechanisms by which the marine biota influences and links organic and inorganic carbon cycles in the world's oceans. Calibrating the model to match current vertical concentration profiles of ocean oxygen and phosphate indicates that roughly half of the nutrients upwelling from the deep ocean are transported to the polar oceans before being incorporated into biomass in the mixed layer; this allows much of the phosphate to escape the settling, rapid decay, phosphate recycle loop in the ocean column. Naturally occurring or anthropogenically caused variations in runoff to the oceans, ocean circulation, atmospheric composition, and climate affect ocean behavior in the global carbon cycle. Moreover, certain "traits" of marine biota, e.g., production of calcareous shells and composition of organic particulates, are shown to have a profound effect on the ability of the oceans to absorb carbon dioxide; and these traits are known to vary significantly between species of phytoplankton.

Conditions of the present ocean appear, in the model, to be coincidental. The present ocean is nested between the Scylla of anoxia and the Charybdis of supersaturation (in $CaCO_3$). These circumstances suggest that biota "traits" such as the ratio of calcite to organic carbon or that of organic carbon to phosphorous fixed by the biota are not constants as they are modeled here, but rather traits of the marine biosphere that react to changes in the marine environment in, as yet, unresolved ways.

Notation

A relative fraction of seafloor area at a given depth
B maximum ocean depth, m
BI inorganic biomass particulates, $CaCO_3$
BO organic biomass particulates
BP components bypassing mixed layer
$c[\ ]$ concentration, mol kg^{-1}
C_p specific heat of seawater, J kg^{-1} $°C^{-1}$
$E(\ )$ exchange coefficient of oxygen, mol m^{-2} yr^{-1}
FI ratio of BP flux to BO produced
H mixed layer thickness, m
J flux, mol m^{-2} yr^{-1}
K vertical eddy diffusivity, m^2 yr^{-1}
k_{Dec} biomass decay rate constant, yr^{-1}
k_{Phot} biomass production rate constant, mol^{-1}
L characteristic length for temperature change, m
N nitrate
NP new productivity flux, g C m^{-2} yr^{-1}
$p(\ )$ partial pressure, μatm
P phosphate
PP primary productivity flux, g C m^{-2} yr^{-1}
R freshwater runoff rate, kg yr^{-1}
RC carbon Redfield ratio
RI inorganic to organic biomass ratio
RN nitrogen Redfield ratio
RO oxygen Redfield ratio
S source rate, mol yr^{-1} kg^{-1}
S_{Dec} decay rate of organic biomass, mol P yr^{-1} kg^{-1}
S_{Phot} decay rate of organic biomass, mol P yr^{-1} kg^{-1}
T temperature, °C
TA titration alkalinity
TC total carbon
v settling velocity m yr^{-1}
w upwelling velocity, m yr^{-1}
z vertical coordinate, m

Greek Symbols

α fraction of BP that appears in upper polar ocean
β fraction of dissolution products that appear in upper polar ocean
γ fraction buried of BO arriving at the ocean floor
Ω fraction of seafloor where neighboring water is supersaturated in $CaCO_3$
Ω' fraction of seafloor where neighboring water is undersaturated in $CaCO_3$
λ heat transfer coefficient, W $°C^{-1}$ m^{-2}
ρ seawater density, kg m^{-3}

Subscripts

a in atmosphere
b bottom water
crit critical saturation value
i species
l lysocline
m in mixed layer
p upper polar water
r input from runoff
sat saturation value
sed sedimentation
0 at ocean surface

References

Ajtay, G. L., P. Ketner, and P. Duvigneaud, Terrestrial primary production and phytomass, in *The Global Carbon Cycle*, edited by B. Bolin et al., pp. 129–218, John Wiley, New York, 1979.

Bacastow, R., and C. D. Keeling, Atmospheric carbon dioxide and radiocarbon in the natural carbon cycle, in *Carbon and the Biosphere*, edited by G. M. Woodwell and E. V. Pecan, pp. 86–135, U.S. Atomic Energy Commission, 1973. (Available as CONF-720510 from Natl. Tech. Inf. Serv., Springfield, Va.)

Bacastow, R., and E. Maier-Reimer, Ocean circulation model of the carbon cycle, *Clim Dyn.*, *4*, 95–125, 1990.

Baes, C. F., Chemistry and biology on atmospheric carbon dioxide, in *Carbon Dioxide Review: 1982*, edited by W. C. Clark, pp. 187–211, Oxford University Press, New York, 1982.

Berger, W. H., K. Fischer, C. Lai, and G. Wu, Ocean productivity and organic carbon flux, I, Overview and maps of primary production and export production, *SIO Ref. 87-30*, Scripps Inst. of Oceanogr., Univ. of Calif., San Diego, La Jolla, 1987.

Bjorkstrom, A., A model of CO_2 interaction between atmosphere, oceans and land biota, in *The Global Carbon Cycle*, *SCOPE 13*, edited by B. Bolin et al., pp. 403–457, John Wiley, New York, 1979.

Bjorkstrom, A., One-dimensional and two-dimensional ocean models for predicting the distribution of CO_2 between the ocean and the atmosphere, in *The Changing Carbon Cycle*, edited by J. R. Trabalka and D. E. Reichle, pp. 258–278, Springer-Verlag, New York, 1986.

Bolin, B., Modeling the oceans and ocean sediments and their response to fossil fuel carbon dioxide emissions, in *The Fate of Fossil Fuel CO_2 in the Oceans*, edited by N. R. Anderson and A. Malahoff, Plenum, New York, 1977.

Broecker, W. S., and T.-H. Peng, *Tracers in the Sea*, Eldigeo Press, Lamont-Doherty Geological Observatory, Palisades, N. Y., 1982.

Broecker, W. S., and T.-H. Peng, Carbon cycle, 1985: Glacial to interglacial changes in the operation of the global carbon cycle, *Radiocarbon*, *28*, 309–327, 1986.

Broecker, W. S., and T.-H. Peng, The cause of the glacial to interglacial atmospheric CO_2 change: A polar alkalinity hypothesis, *Global Biogeochem. Cycles*, *3*, 215–239, 1989.

Broecker, W. S., and T. Takahashi, The relationship between lysocline depth and in situ carbonate ion concentration, *Deep Sea Res.*, *25*, 65–95, 1978.

Budyko, M. I., *The Earth's Climate: Past and Future*, Academic, New York, 1982.

Craig, H., Abyssal radiocarbon in the Pacific, *J. Geophys. Res.*, *74*, 5491–5501, 1969.

Eppley, R. W., Relations between primary productivity and ocean chlorophyll determined by satellite, in *Global Ocean Flux Study*, 85–102, Woods Hole Oceanogr. Inst., Mass., 1984.

Froelich, P. N., M. L. Bender, N. A. Luedthe, G. R. Heath, and T. DeVries, The marine phosphorus cycle, *Am. J. Sci.*, *282*, 474–511, 1982.

Hoffert, M. I., The oceans in one dimension: Upwelling, turbulence, temperature, oxygen, nutrients and carbon, Third International Conference on Analysis and Evaluation of Atmospheric CO_2 Data Present and Past, *WMO Rep. 59*, pp. 195–201, World Meteorol. Organ., Geneva, 1989.

Hoffert, M. I., and B. P. Flannery, Model projections of the time-dependent response to increased carbon dioxide, in *Projecting the Climatic Effects of Increasing Carbon Dioxide*, *Rep. DOE/ER-0237*, edited by M. C. MacCracken and F. M. Luther, pp. 141–190, Washington, D. C., 1985.

Hoffert, M. I., A. J. Callegari, and C. T. Hsieh, The role of deep sea heat storage in the secular response to climatic forcing, *J. Geophys. Res.*, *85*, 6667–6679, 1980.

Hoffert, M. I., A. J. Callegari, and C. T. Hseih, A box-diffusion carbon cycle model with upwelling, polar bottom water formation and a marine biosphere, in *Carbon Cycle Modeling*, *SCOPE 16*, edited by B. Bolin, pp. 287–305, John Wiley, New York, 1981.

Holland, H. D., *The Chemistry of the Atmosphere and Oceans*, Wiley Interscience, New York, 1978.

Jenkins, W. J., Oxygen utilization rates in North Atlantic subtropical gyre and primary production in oligotrophic systems, *Nature*, *300*, 246–248, 1982.

Joos, F., and U. Siegenthaler, Study of the oceanic uptake of anthropogenic CO_2 and ^{14}C using a high-latitude exchange/interior diffusion-advection (HILDA) model, Third International Conference on Analysis and Evaluation of Atmospheric CO_2 Data Present and Past, *WMO Rep 59*, pp. 203–208, World Meteorol. Organ., Geneva, 1989.

Keeling, C. D., The carbon dioxide cycle: Reservoir models to depict the exchange of atmospheric carbon dioxide with the oceans and land plants, in *Chemistry of the Lower Atmosphere*, edited by S. I. Rasool, pp. 251–329, Plenum, New York, 1973.

Lapenis, A. G., and A. I. Kolomeitsev, Effects of ocean circulation on marine biota production (in Russian), *Meteorol. Hydrol.*, no. 1, 77–84, 1987.

Levitus, S., Climatological atlas of the world ocean, *NOAA Prof. Pap. 13*, Geophys. Fluid Dyn. Lab., Princeton, N. J., 1982.

Li, Y.-H., T. Takahashi, and W. S. Broecker, Degree of saturation of $CaCO_3$ in the oceans, *J. Geophys. Res.*, *74*, 5507–5525, 1969.

Lovelock, J. E., *The Ages of Gaia: A Biography of Our Living Earth*, W. W. Norton, New York, 1988.

Maier-Reimer, E., and K. Hasselmann, Transport and storage of CO_2 in the ocean: An inorganic ocean-circulation carbon cycle model, *Clim. Dyn.*, *2*, 63–90, 1987.

Menard, H. W., and S. M. Smith, Hypsometry of ocean basin provinces, *J. Geophys. Res.*, *71*, 4305–4325, 1966.

Moore, B., III, and A. Bjorkstrom, Calibrated ocean models by constrained inverse method, in *The Changing Carbon Cycle*, edited by J. R. Trabalka and D. E. Reichle, pp. 295–328, Springer-Verlag, New York, 1986.

Munk, W. H., Abyssal recipes, *Deep Sea Res.*, *13*, 707–736, 1966.

Oeschger, H., U. Siegenthaler, U. Schotterer, and A. Guglemann, A box-diffusion model to study the carbon dioxide exchange in nature, *Tellus*, *27*, 168–192, 1975.

Peng, T.-H., W. S. Broecker, G. G. Mathieu, Y.-H. Li, and A. E. Bainbridge, Radon evasion rates in the Atlantic and Pacific oceans as determined during the Geosecs program, *J. Geophys. Res.*, *84*, 2471–2486, 1979.

Peng, T.-H., T. Takahashi, W. S. Broecker, and J. Olafsson, Seasonal variability of carbon dioxide, nutrients and oxygen in the northern North Atlantic surface water, *Tellus*, *39B*, 439–458, 1987.

Sarmiento, J. L., Three-dimensional ocean models for predicting the distribution of CO_2 between the ocean and atmosphere, in *The Changing Carbon Cycle*, edited by J. R. Trabalka and D. E. Reichle, pp. 279–294, Springer-Verlag, New York, 1986.

Sarmiento, J. L., and J. R. Toggweiler, A new model for the role of the oceans in determining atmospheric pCO_2, *Nature*, *308*, 621–624, 1984.

Shaffer, G., A model of biogeochemical cycling of phosphorous, nitrogen, oxygen and sulfur in the ocean: One step toward a global climate model, *J. Geophys. Res.*, *94*, 1979–2004, 1989.

Siegenthaler, U., Uptake of excess CO_2 by an outcrop-diffusion model of the ocean, *J. Geophys. Res.*, *88*, 3599–3608, 1983.

Sorokin, Y. I., The primary production of the seas and oceans, in *General Ecology*, *Bioceanology*, *Hydrobiology*, vol. 1 (Biology Series), Translated from Russian, pp. 3–35, G. K. Hall, Boston, Mass., 1978.

Sugimura, Y., and Y. Suzuki, A high-temperature catalytic oxidation method for the determination of non-volatile dissolved organic carbon in seawater by direct injection of a liquid sample, *Mar. Chem.*, *24*, 105–131, 1988.

Sundquist, E. T., Geological perspectives on carbon dioxide and the carbon cycle, in *The Carbon Cycle and Atmospheric CO_2: Natural Variations Archean to Present*, *Geophys. Monogr. Ser. 32*, edited by E. T. Sundquist and W. S. Broecker, pp. 5–59, AGU, Washington, D. C., 1985.

Sundquist, E. T., Geologic analogs: Their value and limitations in carbon dioxide research, in *The Changing Carbon Cycle*, edited by J. R. Trabalka and D. E. Reichle, pp. 371–402, Springer-Verlag, New York, 1986.

Takahashi, T., W. S. Broecker, and A. E. Bainbridge, Supplement to the alkalinity and total carbon dioxide concentration in the world oceans, in *Carbon Cycle Modeling*, *SCOPE 16*, edited by B. Bolin, pp. 159, John Wiley, New York, 1981.

Valiela, I., *Marine Ecological Processes*, Springer-Verlag, New York, 1984.

Volk, T., Multi-property modeling of the marine biosphere in relation to global climate and carbon cycles, Ph.D. thesis, N. Y. Univ., University Microfilms International, Ann Arbor, Mich., 1984.

Volk, T., and M. I. Hoffert, Ocean carbon pumps: Analysis of relative strengths and efficiencies in ocean-driven atmospheric CO_2 changes, in *The Carbon Cycle and Atmospheric CO_2: Natural Variations Archean to Present*, *Geophys. Monogr. Ser. 32*, edited by T. E. Sundquist and W. S. Broecker, pp. 91–110, AGU, Washington, D. C., 1985.

Volk, T., and Z. Liu, Controls of CO_2 sources and sinks in the Earth scale surface ocean: Temperature and nutrients, *Global Biogeochem. Cycles*, *2*, 73–89, 1988.

Weiss, R. F., The solubility of nitrogen, oxygen and argon in water and seawater, *Deep Sea Res.*, *17*, 721–735, 1970.

Whittaker, R. H., and G. E. Likens, The biosphere and man, in *Primary Production of the Biosphere*, edited by H. Leith and R. H. Whittaker, pp. 305–328, New York, 1975.

Wyrtki, K., The oxygen minimum in relation to ocean circulation, *Deep Sea Res.*, *9*, 11–23, 1962.

Zienkiewicz, O. C., *The Finite Element Method*, 3rd ed., McGraw-Hill, New York, 1977.

B. P. Flannery and H. S. Kheshgi, Corporate Research Laboratories, Exxon Research and Engineering Company, Annandale, NJ 08801.

M. I. Hoffert, Department of Applied Science, New York University, New York, NY 10003.

(Received July 13, 1989;
revised July 11, 1990;
accepted August 31, 1990.)

Hoffert, M.I., 1992: Climate sensitivity, climate feedbacks and policy implications. In Mintzer, I.M., ed. *Confronting Climate Change: Risks, Implications and Responses*, Cambridge University Press, Cambridge (Great Britain), New York (USA), Victoria (Australia), pp. 33-54.

CHAPTER 3

Climate Sensitivity, Climate Feedbacks and Policy Implications

Martin I. Hoffert

Editor's Introduction

The study of climatic problems includes much more than the atmosphere. In this chapter, Martin Hoffert looks back over the Earth's history to explore the natural archives of paleoclimate data, using modern techniques of isotopic and oceanographic analysis. He combines these analyses with computer modelling studies, which were designed to estimate the sensitivity of global climate to changes in atmospheric composition, the position of the Earth's orbit, and circulation of the world ocean. In the process, as one of the paper's reviewers (geophysicist Tyler Volk) observes, Hoffert spins a complex web, providing us a rapid and enthralling tour of Ice Ages, the planetary evolution of Earth, Mars, and Venus, atmospheric physics, soil chemistry, and ocean circulation. And he explains the strengths, weaknesses, and fundamental limits of climate models in terms that are readily penetrable by the lay reader.

Because its range is the widest, this chapter is the longest in the book. But it also expands directly on the discussion of linkages in the previous chapter, to show the interconnections among the major components of our planetary system — air, ocean, land, and biota. Hoffert notes that everything on the surface of the planet — from the smallest dust particle to the largest iceberg — absorbs and re-emits the radiation that comes to us from the sun. And he explains how each of these components interacts with Earth's lifeforms to affect the overall radiative balance of the planet with its surroundings.

Woven throughout this chapter, you will find a thread of respect for the fundamental uncertainties in our understanding of natural processes, and the intellectual challenge facing geophysicists, geochemists, and other scientists who try to model future effects of human activities on climate. None the less, this chapter concludes with a strong plea for political action to reduce the risks of rapid climate change while we carefully monitor the state of the planet and continue to pursue basic and applied research in planetary science.

Martin Hoffert is a professor at, and former chairman of, the Department of Applied Science at New York University — one of the pre-eminent interdisciplinary graduate-level science programmes in the United States. He is best known for his work in geochemistry and oceanography and for his path- breaking applications of small, transparent, and understandable models to large complex geophysical processes.

Syukuro Manabe, one of the few true masters of the three-dimensional general circulation models of the atmosphere, caps off Hoffert's presentation, describing a potent new international research strategy for meeting the challenge of uncertainty head-on. (See page 51) Manabe urges us to continue and intensify carefully targeted research on planetary dynamics, linking three current lines of attack: (1) process studies of atmospheric dynamics that can improve scientific knowledge of cloud feedbacks and land surface processes; (2) monitoring of the natural environment with both in situ measurements and observations from satellites; and (3) diagnostic studies to promote model validation, testing, and impact assessment. With this three-pronged approach, Manabe argues, the scientific community can provide additional policy-relevant data on the prospects for rapid climate change. This additional data can help national leaders and business executives to sustain the momentum of continued economic development in a world where future climate is uncertain and, potentially, quite different from that of the recent past.

- I. M. M.

1 Climate Feedbacks and Uncertainty

What limits our capability to project global warming into the future? It is not an inability to compute the direct amount of radiative heating by greenhouse gases. Rather, it is our poor understanding of climate feedbacks. These are processes of the climate system, that amplify or diminish any of the climatic changes from direct global warming (Manowitz, 1990). For example, cloud-radiative feedback (also described by Crutzen and Golitsyn, 1992, this volume) is the process in which greenhouse gas emissions lead not just to rising temperatures, but also to changes in the distribution and characteristics of clouds over the planet. Clouds scatter incident solar radiation back to space (cloud albedo), but they also tend to retain heat radiated from the surface below (cloud greenhouse). As a result, changes in cloud distributions induced by greenhouse gas emissions can promote both warming and cooling, in different ways, simultaneously. It is not yet clear which of these effects prevails when the climate changes.

A feedback is positive if it amplifies a warming or cooling trend, negative if it diminishes the same warming or cooling trend. Positive feedbacks promote large climate changes, negative ones promote climate stability.

Our understanding of the cloud feedback process is incomplete enough that, by itself, it produces a factor-of-three uncertainty in future temperature changes. The statistical term "factor-of-three" refers to a range of prediction so uncertain that the top limit is three times the bottom limit. A typical climate model might, for example, be used to forecast a set of potential global temperature rises. When the feedback effects of cloud cover are built in to the model, they add a "factor-of-three" ambiguity. As a result, we can be no more specific than to predict a rise of between 1.5 and 4.5° Celsius for an atmospheric CO_2 doubling (Houghton et al., 1990). The cloud feedback problem is presently under intense study, partly in hopes of rendering more accurate forecasts; but it is by no means the only "wild card" that can affect global warming.

In this chapter, I will review a number of wide-ranging considerations bearing on feedback effects — first covering the physical and chemical processes underlying greenhouse calculations, then discussing feedbacks in the climatological record of Earth, Mars, and Venus, and focusing ultimately on certain potentially large wild card feedbacks from the oceans and biosphere. Let it be said up front that there are speculative aspects to this review. But there are also very practical reasons to explore these possibilities.

International bodies like the Intergovernmental Panel on Climate Change (IPCC) have projected dramatic temperature rise and other climate changes from anthropogenic greenhouse gases over the next century (Houghton et al., 1990). Some critics emphasize uncertainties of the climate models on which these projections are based, and conclude that global warming is likely to have been overestimated (Marshall Institute, 1989; Lindzen, 1990; Ellsaesser, 1990). That may be. But in fairness, because of the unpredictability of feedback processes, uncertainties in climate models are as likely to *underestimate* future global warming as to overestimate them (Schneider, 1990).

1.1 Lessons from the climatic past

In modelling the future, it is instructive to look to our climatic past for guidance (MacCracken et al., 1990). Even without a fossil-fuel greenhouse, the present world climate is, from the perspective of the past few million years, anomalously warm — an interglacial hiatus from which, in the absence of human intervention, we are destined to descend to another ice age. The climate history of the Earth over the past 1.6 million years, the so-called Pleistocene period, is one of cycles between glacial (cooler) and interglacial (warmer) periods. Ice volume and temperature fluctuations, with periodicities of 100,000, 41,000 and 23,000 years, are seen in the geologic record — and are thought to arise from variations in the Earth's orbit (Crowley and North, 1991).

But one needs to go back millions, even tens of millions, of years to find atmospheric CO_2 levels (and corresponding periods of global warmth) comparable to what humanity can produce in coming centuries. Figure 1, showing a range of prehistoric climate reconstructions, suggests that comparable levels last existed in the mid-Pliocene era, about three million years ago. As a result of various non-equilibrium effects — different feedbacks coming into play over different time scales — simple "paleo-analogues" of global warming may not exist (Crowley, 1990). If current climate and carbon cycle models are even approximately correct, the present era of human history can bequeath on to future generations an anomalous "super-interglacial" period — an unprecedented hot spell persistent enough to overcome the next few glacial cycles. For reasons discussed subsequently, humanity's combustion of the fossil fuel reserve could alter the global climate over time scales comparable to the lifetime of the human species. Those who see *Homo sapiens* as somehow apart from natural processes might well conclude we are at "the end of nature" (McKibben, 1989).

1.2 Implications of uncertainty on climate policy

As we move toward uncharted climatic territory, it seems all the more important to understand relevant feedback mechanisms that may have operated on Earth and on other Earth-like planets. It may become rapidly necessary to make climate policy in the face of substantial uncertainty. The stakes are too high to rule out plausible climate change processes for which there is physical evidence.

Indeed, recent environmental history tells us that surprises can be expected in the precise way that any human-triggered environmental perturbation unfolds. Atmospheric photochemistry models were predicting ozone depletion from chlorofluorocarbon (CFC) emissions well before they were measured. But the specifics of the "ozone hole" were unanticipated. Local ozone depletion have proven much more extreme than the gradual decreases that had been

predicted by models. The Antarctica ozone hole is now understood to depend on chemical reactions catalysed in polar stratospheric clouds, reactions which had not been included in models. Although the implications of ozone depletion are still hotly debated, they were perceived as serious enough in the post-ozone-hole world to rapidly adopt a global CFC regulatory treaty — the Montreal Protocols. Though controversial at first, rollbacks in CFC emissions have proven easier to implement than some had originally thought. Reductions in the emissions of CO_2, the major greenhouse gas, may prove more challenging because at least 85% of global energy consumption is presently from fossil fuels (Hammond, 1990; Holdren, 1992, this volume).

Could some climatic feedback from the oceans or biosphere trigger a response to greenhouse warming for which the world is ill-prepared? A broad perspective, including consideration of climatic long shots, seems only prudent if one seeks to minimize "unpleasant surprises in the greenhouse" (Broecker, 1987).

2 The astrophysics of the Greenhouse Effect

2.1 The validity of greenhouse models

When global climate is constant, conservation of energy dictates that the solar energy absorbed by the atmosphere, clouds and surface must equal the energy radiated back from the Earth's atmosphere, clouds, and surface to cold space. This is the fundamental energy balance on which climate modelling is based. Necessarily — due to the complexities of molecular absorption and emission of radiation — the computer codes from which the solar and infrared components of radiation are computed are too often intimidating to the uninitiated. Atmospheric radiation is a discipline where the modelling art looks mysterious even to scientifically knowledgeable researchers in adjacent fields.

This does not mean radiative heating rates computed from greenhouse gases are a significant source of error or uncertainty. Radiation models developed and intercompared by a dedicated research community over decades, with the help of satellite data, are believed accurate to within a few percent.

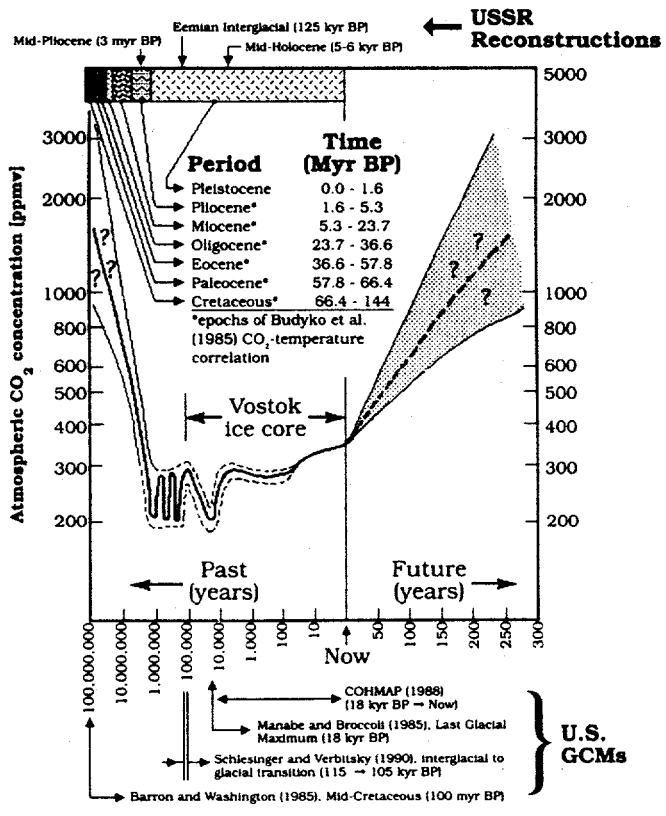


Figure 1. Variations of atmospheric carbon dioxide concentration, shown over time: from 100 million years in the past to 300 years into the future. The upper scale (top left corner) shows only past paleoclimates as reconstructed by Soviet researchers, keyed against a table showing the name and duration of each period. (MacCracken et al., 1990). The lower scale shows periods simulated by general circulation models, both for the past (to the left of "Now") and for the future (shown to the right of "Now"). Note that while the future scale is linear, the past recedes exponentially.

It is believed that the paleoclimate variations, were partly caused by the greenhouse effect, from changes in carbon dioxide levels. Human emissions of fossil fuel carbon dioxide in the coming decades and centuries could produce levels of greenhouse warming comparable to those produced millions of years ago.

2.2 *The Earth as an absorber and emitter of radiation*

Simple theories of the Earth's radiative behaviour compare the Earth to the astrophysical concept of a "black body". Like the black hole of astrophysics, a black body absorbs all radiation which contacts it. Unlike a black hole, it is also a perfect emitter of radiant energy. The amount of electromagnetic radiation which black bodies emit is proportional to the fourth power of their temperature.[1]

The greater the temperature of a black body, the shorter the wavelength of the radiant energy which it emits. For example, stars are very nearly black bodies; which is why hot stars are blue (short wavelength), cold stars red (long wavelength) and our yellow sun is somewhere in between. It is possible to express the Earth's energy loss as equivalent to that of a black body. Most of this radiation leaving the Earth is in the infrared part of the spectrum (between the wavelengths of 8 and 30 microns). These wavelengths of radiation are emitted and absorbed by greenhouse gases, including water vapour, carbon dioxide (CO_2), ozone (O_3), methane (CH_4), and nitrous oxide (N_2O) and the chlorofluorocarbons.

That greenhouse gases warm the surface of the Earth and other planets is a well-established scientific fact. A classic proof that the greenhouse effect is taking place on Earth is that the present global mean surface temperature, about 288 Kelvin or 15° Celsius, is 33°C warmer than the Earth's black-body temperature (255K).[2] The additional thirty-three degrees Celsius of surface warmth is due to the absorption and re-emission of radiation by atmospheric water vapour and (to a lesser extent) by carbon dioxide and ozone. What is, to some extent, still uncertain is the amount and timing of global warming from humanity's greenhouse gas emissions.

2.3 *Greenhouse gas molecules and the shrinking infrared "window"*

An orbiting observer who could "see" in thermal infrared could not see the Earth's surface very well because the view would be blocked by water vapour, carbon dioxide and ozone — trace gases which we can see through, but which absorb infrared radiation. But the atmosphere's infrared opacity has "windows" of transparency at some infrared wavelength bands — that let some wavelengths of radiation through. Increases in the amount of greenhouse gases produce global warming because they further block infrared windows. Some gases, because of the strength and distribution of their absorption spectrum, are more effective blockers than others (Table 1). For example, molecule for molecule an increase in chlorofluorocarbon-12 concentration produces almost sixteen thousand times the greenhouse heating of carbon dioxide.

This comparison does not take into account the lifetime of various greenhouse gas molecules, and their integrated effect over time. For example, CO_2 will almost certainly continue to be the major greenhouse gas, because there are so many molecules of CO_2 being added, and they persist so long in the atmosphere.

2.4 *The link between surface temperature and incoming radiation levels*

It is also critical to the greenhouse effect that in the troposphere — the first 11 kilometres of atmosphere above the surface, where clouds and weather systems exist — the temperature drops off with altitude about 6.5 °C for every kilometre higher. The altitude from which radiation escapes from the Earth to space (called the radiation height), is currently about five kilometers above the surface — roughly halfway up the troposphere. When they escape, the infrared photons carry with them, in their distribution of wavelengths (or "temperature"), a signature of the atmosphere's radiation height. As more greenhouse gases are added, the atmosphere gets optically thicker — it radiates infrared from higher, colder altitudes. This creates an imbalance between the absorbed solar radiation and emitted longwave radiation. The climate system restores the balance by raising the surface temperature to a new value. This process is the greenhouse effect.

Because the rate at which temperature drops off with altitude tends to remain constant, it is possible to find a height at which the absorbed solar energy equals the black body radiation to space. This is called the effective radiation height. The temperature at the radiation height is coupled to the Earth's surface temperature — as one changes, so does the other. The coupling between the two temperatures — that of the surface and that of the radiation height — results from a process known as "convective adjustment," and allows us to make a first estimate of the climate sensitivity.

In analyses of climate change, it is important to distinguish between radiative forcing and surface temperature

Table 1. Radiative heating of anthropogenic greenhouse gases relative to carbon dioxide per molecule added to the atmosphere

Gas	Relative radiative heating
Carbon dioxide (CO_2)	1
Methane (CH_4)	21
Nitrous oxide (N_2O)	206
Chlorofluorocarbon-11 (CFC-11)	12,400
Chlorofluorocarbon-12 (CFC-12)	15,800

(After MacCracken et al., 1990).

[1] Per unit area, the rate of energy loss from a black body at temperature T is σT^4, where σ is the *Stefan-Boltzmann constant* ($\approx 5.67 \times 10^{-8}$ W m^{-2} K^{-4}).

[2] T_{eff} is the effective black body temperature, with S_o the solar energy per unit area falling on a plane perpendicular to the sun's rays at the top of the atmosphere (the solar constant = 1368 watts per square metre (W m^{-2}) and the fraction of solar radiation reflected back to space (the planetary albedo ≈30%). A global energy balance gives: $T_{eff} = [S_o(1-\alpha)/(4\sigma)]$ =255 K (-18°C).

change. Radiative forcing refers to the rate at which heat is added to or removed from the climate system — by changes in incident sunlight, planetary albedo (reflectivity), or greenhouse gas concentrations. Climate *sensitivity* is the ratio of the temperature change to the radiative forcing. Gaining a better understanding of the sensitivity of the Earth's climate is a fundamental goal of climate change research.

If, for example, the atmosphere were heated by an additional four watts per square metre — the heating predicted by radiative models for a CO_2 doubling since the beginning of the industrial period — we would expect the surface to warm eventually by about one degree Celsius.[3] This is because the climate sensitivity based on black-body cooling alone would correspond to a surface warming of about 1° Celsius. However, this warming could be either increased by positive feedback, or decreased by negative feedback. The realized global warming will also be affected by the heat capacity of the oceans, which would tend to postpone the temperature rise somewhat into the future.

A more general discussion will be given presently, linking climate sensitivity, feedbacks and transient climate change on decadal to century time scales. But in keeping with the broad objectives of this survey, it is useful to first discuss how climate feedbacks can operate over much longer time scales — and not only on Earth.

3 The Evolution of Climate

Greenhouse gases seem to have profoundly affected the climates of at least three "terrestrial" planets: Venus, Earth and Mars. These all have surface temperatures that are warmer than they would be without the greenhouse effect (Table 2), and all have very likely experienced evolutionary histories in which the greenhouse gas content of their atmospheres has varied (Kasting et al., 1988).

3.1 The vanished greenhouse of Mars

At present, Mars is so cold (218K, or -55 °C) that, even with its large seasonal and daily temperature fluctuations, liquid water does not normally exist. Such low temperatures mean there can be very little water vapour in the atmosphere, though appreciable amounts of water as permafrost are believed to exist beneath the surface and in permanent ice caps. (Overlying the water ice caps at both poles are "seasonal" ice caps of solid carbon dioxide, which condense from the atmosphere in winter and return as a gas to the atmosphere in summer.) Given this deep freeze, water vapour feedback on Mars today is so weak that even with twice the surface pressure of CO_2 as the Earth, Mars produces less than half the greenhouse warming. This is not too surprising, as most of the 33 degrees of greenhouse warming on Earth comes from water vapour.

What is surprising is that images of the Martian surface, returned by Mariner and Viking spacecraft, show evidence of ancient riverbeds — channels cut into the surface long ago, perhaps billions of years ago, by running water — complete with tributaries, meandering paths, and other properties of rivers on Earth (Figure 2). This suggests Mars was once much warmer, warm enough to sustain liquid water (requiring a surface temperature greater than 273 Kelvin). The most plausible explanation of this prior warmth is a strong greenhouse warming from an earlier, much denser CO_2 atmosphere. Because Mars is smaller than the Earth, it cooled earlier and is tectonically inactive at present. We may surmise that early in its history, like the Earth, it released volatile gases, including carbon dioxide and water. The highest mountain in the solar system is an inactive volcano on Mars: Olympus Mons, taller than Mount Everest.

As Mars cooled, degassing of CO_2 (injection to the atmosphere from fissures in Martian surface rock) slowed. At the same time, carbon dioxide may have been drawn out of the atmosphere and converted to calcium carbonate rocks in the crust — a process that could continue so long as the temperature remained above the freezing point of water.

[3] In this work, we call the energy radiated to space per unit surface area per unit surface temperature change the radiative damping coefficient, λ. The climate sensitivity, λ^{-1}, is inversely proportional to the radiative damping coefficient. A first approximation of λ, including the convective adjustment but neglecting all other atmospheric feedbacks, is the blackbody cooling coefficient, $\lambda' = 4\sigma T_{eff}^3 \approx 4.0\ Wm^{-2}\ K^{-1}$.

Table 2. Variation of effective and surface temperature on three planets showing the influence of greenhouse gases

Planet	Atmosphere pressure [atm]	Greenhouse gases	Orbit [AU]	Solar constant [W m^{-2}]	Albedo [%]	Effective temp [K]	Surface temp [K]	Greenhouse warming [°C]
Venus	90	CO_2	0.723	2620	76	229	750	521
Earth	1	H_2O, CO_2	1.000	1368	30	255	288	33
Mars	0.006	CO_2	1.524	589	25	210	218	8

Figure 2. Images of the Martian surface, returned by the Viking spacecraft, show evidence of ancient river beds, complete with tributaries, meandering paths, and other properties of rivers. This suggests that Mars was once much warmer, warm enough to sustain liquid water. The most plausible explanation is a strong greenhouse warming from a much denser CO_2 atmosphere early in the planet's history. Source: Based on Nasa Viking images assembled by the US Geological Survey - USGS Miscellaneous Investigations Series.

(Liquid water is necessary to precipitate calcium carbonate from carbon dioxide gas.) Presumably, the process stopped when the planet cooled to the freezing point of water. This scenario may explain why Mars' atmospheric pressure is so close to the triple point — the point where solid, liquid and vapour phases of water coexist. A Martian greenhouse more than three-and-a-half billion years ago is the most plausible hypothesis for its apparently wetter climate early on. The critical evidence confirming this hypothesis would be Mars rock samples showing carbonate content consistent with absorption of a prior thick carbon dioxide atmosphere. Such samples will hopefully be obtained and analysed in future missions to Mars.

3.1.1 Earth compared to Mars

Evidently, Earth averted the deep freeze of Mars. This is at least partially because the Earth's crust has remained geologically active. Unlike Mars, the CO_2 in Earth's carbonate rock is recycled to the atmosphere through a combination of chemical reactions and plate tectonics. It takes 100 million years for molten rock to upwell at the boundaries of tectonic plates under the sea, spread across the ocean floor and, finally, plunge back down into the deep mantle. Each of these stages releases CO_2 into the atmosphere or swallows it into the rock shell of the Earth again — a cycle which has taken place many times over since the Earth formed. The carbon rock cycle is a geochemical conveyor belt partly driven by radioactive heat released by the mantle. During the 4.6 billion years of planetary evolution there has been a gradual slowing of this conveyor belt as the Earth cooled, and a reduction in the amount of carbon dioxide in the Earth's atmosphere. Fortunately for life on Earth, the conveyor belt has not altogether stopped, as it has, apparently, on Mars.

3.2 The runaway greenhouse of Venus

An example of a positive feedback gone wild is the runaway greenhouse of Venus, where vaporized greenhouse gases apparently caused progressively more warming until all surface volatiles, including water and carbon dioxide, were driven to the gaseous phase. Venus today is an uninhabitable hell with a dense carbon dioxide atmosphere and a surface temperature of 750 Kelvin (477 °C) — hot enough to melt lead. In the unlikely event they ever existed, Venus' liquid water oceans have vanished without a trace. Even the water vapour that might, under more favourable conditions, have condensed to oceans on Venus is gone — the hydrogen lost to space, the oxygen re-combined in rocks at the surface.

3.2.1 Earth compared to Venus

Fortunately again, a life-threatening runaway greenhouse effect like that of Venus, but caused by humanity's continued fossil fuel carbon combustions, is unlikely on Earth. Burning all the oil, natural gas, coal and shale could not heat the atmosphere enough to volatilize the carbonate rocks to carbon dioxide or boil the oceans.

But we ought not to become too complacent about avoiding a Venus-like runaway greenhouse. Major, long-term changes in the Earth's climate lasting thousands of years and beyond are likely if humanity consumes a significant fraction of its recoverable fossil fuel reserve. The inadvertent introduction of greenhouse gases to the Earth's atmosphere since the worldwide industrial revolution in amounts large enough to affect global climate has been called a "grand geophysical experiment" (Revelle and Suess, 1957). For better or worse, our descendants will experience the climatic impacts of this experiment.

3.3 Evolution of the Earth's greenhouse effect

Stars brighten as they evolve. The sun, for example, has brightened by 30% since the solar system formed. All things being equal, a simple energy balance indicates that the Earth's surface would have been below the freezing point of water at the time that life evolved — a biological impossibility known as the faint-young-sun paradox (Kasting et al., 1988). The greenhouse effect of an early atmosphere much richer in CO_2 can be invoked to resolve this paradox, with geologic evidence to support it. But, in the face of increasing solar luminosity, how was the planetary "thermostat" regulated in the eons after life had evolved? Some stabilizing climatic feedback process is strongly indicated.

Walker et al. (1981) were the first to propose that, on geologic time scales, climate is stabilized by factors affecting the rate at which calcium silicate rocks are geochemically *weathered* (that, is, converted to calcium carbonate by reaction with atmospheric carbon dioxide). Recall that weathering may also have played a significant role in removing Mars' early CO_2 atmosphere. The rate at which CO_2 is removed from the Earth's atmosphere by rock weathering increases as temperature increases for two reasons: (1) rainfall and runoff increase, carrying more carbonate to the ocean; and (2) the respiration of soil organisms increases.

The second mechanism — by releasing more carbon dioxide to the soil, which diffuses to the atmosphere — initially produces a positive feedback: Warmer surface temperatures promote the release of more carbon dioxide, which in turn promotes warmer temperatures, and so on. But in the long run, higher temperatures will increase weathering from rainfall — a negative feedback. The warming of surface temperatures would promote, through weathering, a drain of CO_2 out of the atmosphere, which would drive CO_2 levels down. That, in turn, would create a cooling trend opposed to the initial warming. The cooling trend would eventually slow down the drain of CO_2, until it reached an equilibrium. Thus, the long-term effect of geochemical weathering is to stabilize global temperature — a negative climate feedback.

3.4 The role of life in greenhouse evolution

Although many geochemists believe carbon dioxide variations over geologic time can be explained abiotically, they are increasingly challenged by theories in which life plays a more or less important role. Even the geochemical weather-

ing rate feedback is now understood to be affected by living organisms in many different ways (Volk, 1987, 1989).

The most fiercely controversial idea in this arena may be the Gaia hypothesis of James Lovelock and Lynn Margulis (Lovelock, 1988). This hypothesis — some think "metaphor" is a better term — holds that living organisms on Earth actively regulate atmospheric composition and climate in the face of challenges like the increasing luminosity of the sun. Specific Gaian mechanisms that have been invoked are the emission of greenhouse gases and/or reflective cloud-producing gases (Charlson et al., 1987) by organisms to produce a planetary homeostasis. Though a compelling case has by no means been made for the "strong" — planetary homeostasis — version of Gaia, most earth scientists would accept a "weak" version — that biological process and feedbacks affect global climate. We will examine biological processes and biological feedbacks on climate in more detail presently.

3.5 Continental drift

The gradual reduction in the CO_2 content of the atmosphere, as the Earth cooled, was probably modulated by the detailed dynamics of continental drift. Such modulations are recorded as changes in sea floor spreading rates (measured by magnetic field reversals relative to sea-floor spreading centers like the mid-Atlantic ridge). There is also evidence from variation in carbon isotopes and organic carbon burial rates that over the last 600 million years — the Phanerozoic eon — atmospheric CO_2 rose and fell by large factors (Budyko et al., 1987; Berner, 1990). On these time scales, supercontinents break up and reassemble, and ice caps come and go in a flash.

4 Lessons from Climate Evolution

4.1 Evaluating the case for pessimism: "return to Cretaceous"

One hundred million years ago — the mid-Cretaceous when dinosaurs roamed the Earth — was a time of great warmth. The Earth was about 10 °C hotter and the poles ice-free, very possibly in response to greenhouse warming from an atmosphere four to ten times richer in carbon dioxide (Barron and Washington, 1985). We have only a crude idea of what planetary ecosystems looked like in those days, or of how well a human technological civilization would have fared. Interestingly, the recoverable fossil fuel reserve contains enough carbon to raise atmospheric CO_2 to mid-Cretaceous levels (Figure 1). While paleoclimatic analogies have problems, it is not unreasonable to think of the Cretaceous climate as a "worst case" greenhouse warming scenario, if humanity burns the fossil fuel reserve to depletion.

4.2 Evaluating the case for optimism: "return to Eden"

Given that large variations in atmospheric carbon dioxide and climate have occurred over Earth history — all the while within boundaries habitable to life — it is not obvious that humanity's greenhouse will necessarily be "bad". Some Soviet climatologists have recently used a paleoclimate analogue approach to project a more favourable world climate in the latter half of the 21st century, as the result of continued fossil fuel burning (Budyko, 1991; Budyko and Izrael, 1991). Climate model projections, coupled with the prospect of indefinite climate change, suggest that any regional "winners" in global warming will be transient at best, while other regions would suffer reduced agricultural productivity due to increased drought frequency, with continuous changes. Budyko and his colleagues, on the other hand, see a warmer climate as one with wetter continents everywhere. Combined with increased crop fertilization from higher CO_2 levels, this warmer climate would provide a bonanza of global agricultural productivity — perhaps enough, they suggest, to avert the starvation of billions in the much more populous world of the 21st Century.

pid climate change as well, and about the need for a constructive response.[4] At this point, even the executive branch of the US government (which tends to be quite conservative on global warming) has concluded in its National Energy Strategy that "there is sufficient credible scientific concern to start acting to curb the buildup of so-called greenhouse gases" (White House, 1991).

I will argue that even if the long-term climatic impacts prove to be beneficial — and it is by no means clear that they will — a rationale exists for limiting the rate of global warming based on the rate ecosystems can respond to climatic change. I also want to show that wild card climatic feedbacks can affect these warming rates.

4.3 Evaluating the case for climate engineering

Profound technological developments are sometimes stimulated by accidental discoveries. One outcome of our species' inadvertent effect on planetary climate could be the ability to engineer climatic change. Proposals have already been made to compensate for global warming by geo-engineering — seeding the stratosphere with reflective aerosol particles, deflecting incident solar radiation with space-based mirrors, sequestering atmospheric CO_2 by fertilizing polar zone plankton and other ingenious manipulations (NAS, 1991). These might be considered humanity's intervention in the climate system to provide negative feedbacks. Apart from their technological feasibility, such capabilities raise a host of legal and ethical questions that are only beginning to be addressed.

Opponents of geo-engineering argue that such approaches pre-suppose a much better knowledge and understanding of the climate system than all but the most confident of scientists are prepared to claim. Moreover, these measures would not address the underlying problem at its scource: they won't do anything to reduce co-related problems such as acid rain,

[4] The US National Energy Strategy concludes that "there is sufficient credible scientific concern to start acting to curb the buildup of so-called greenhouse gases" (White House, 1991).

photochemical pollutants, and urban air pollution, which solutions based on reducing fossil fuel use would address. Finally, they are likely to require continuous action which, if interrupted for any reason, could lead to disaster if the greenhouse heating, which these solutions are designed to mask, were allowed to grow indefinitely.

If humanity can make major alterations in the Earth's climate without really trying, what could be done in a more technologically sophisticated age to change the inhospitable climates of neighbouring planets? Terraforming — the creation of Earth-like habitats on other planets — has received serious scientific attention in recent years (McKay et al., 1991). An earlier NASA study, prior to analysis of Viking mission data, assumed enough greenhouse warming potential in the CO_2 Mars polar caps that a modest reduction in their albedo (the amount of radiation they reflect) could raise planetary temperatures above freezing. This could be achieved, the study concluded, by covering the polar caps with low-albedo dust or growing dark plants on them. (Averner and MacElroy, 1976). We now know that the "dry ice" caps at Mars' poles are almost entirely volatilized and re-condensed each Mars year, and that the NASA terraformers overestimated the CO_2 mass of these caps available for greenhouse warming by many orders of magnitude.

More recent ideas include releasing carbon dioxide in surface rocks. McKay (1987) estimates an amount equivalent to one Earth atmosphere of pressure could be released in 200 years by diverting 1% of the incident solar energy to this purpose. But again, it is not yet known how much carbon dioxide exists either absorbed or in carbonate rocks (Fogg, 1989).

Allaby and Lovelock (1984) in their speculative novel, *The Greening of Mars*, explored the creation of a warmer, wetter and generally more habitable Mars by colonists seeding its atmosphere with chlorofluorocarbons. In their novel, CFCs are shipped from Earth as payload in formerly nuclear-tipped missiles left over from the Cold War — though the authors significantly underestimated the mass of CFCs it would take. Even with "free" interplanetary transportation, it might be more cost-effective to make chloroflorcarbons from material available on Mars. Also, though CFCs are certainly efficient greenhouse gases (Table 1), they are not necessarily optimum. The infrared absorption of CFCs is, after all, an incidental property of an artificial molecule developed for refrigeration and aerosol spray can applications. The best approach might be to manufacture on Mars new greenhouse molecules engineered to strongly and broadly absorb infrared radiation at very low concentrations.

Terraforming today is a not yet entirely respectable concept that serves mainly to provide plot opportunities for science fiction. But in the same sense that genetic engineering has the potential to alter the pace and direction of "natural" biological evolution, so geo-engineering has the potential to accelerate and divert natural climate evolution. Given sustainable sources of solar or fusion power, and a long enough term commitment, the feasibility of geo-engineering planetary atmospheres is real enough. Such schemes could profitably exploit climate feedbacks from greenhouse gases liberated from planetary surfaces, particularly the water vapour feedback from permafrost driven from Mars' regolith and ice caps. But before that happens, humanity will have faced the challenge of the "greenhouse century" — the next hundred years of greenhouse-gas-induced climate change on Earth.

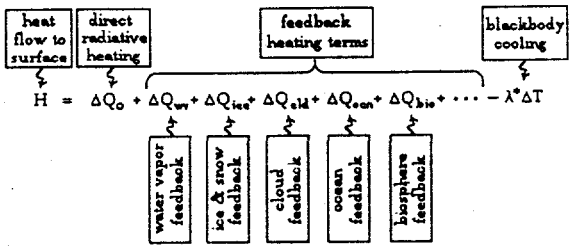
5 Sensitivity, Feedbacks and Climatic Transients

There are important relationships between radiative feedback, the rate of global warming, and climate sensitivity. To get a feeling for these relationships, it is helpful to consider processes that come into play when the atmosphere is heated by some direct radiative forcing.

For our purposes, it is irrelevant whether this forcing comes from a change in the level of solar radiation, a change in planetary albedo (reflectiveness), or from the addition of greenhouse gases — so long as it comes from outside the climate system (Houghton et al., 1990).

5.1 Heat flux to the surface

Climatic changes are due to both external and internal factors. The internal factors can be either autonomous (self-generated) or they can be responses or feedbacks to some externally imposed change. An example of an autonomous internal change is the so-called El Niño — a dramatic change in the wind and temperature system of the equatorial Pacific, that occurs over two- to five-year intervals, and impacts the global climate. This variation is believed to arise from an interaction within the atmosphere-ocean system, rather than a response to radiative forcing imposed from outside. Climate feedbacks are triggered over various time scales, ranging from "fast" meteorological processes to "slow" geologic processes (some of which were discussed in the preceding section). We can calculate the heating of the planetary surface from the following formula that includes both external radiative forcing and internal feedbacks:



Perhaps the best-verified fast feedback is due to water vapour — a greenhouse gas whose concentration is largely controlled by surface temperature. When the surface warms, evaporation and precipitation both increase for the world as a whole; the water vapour content of the atmosphere goes up, and the greenhouse effect from this additional water vapour produces an incremental radiative forcing (ΔQ_{wv}). Satellite studies confirm that this water vapour feedback roughly

doubles the direct radiative heating within a few months (Raval and Ramanathan, 1989). Twenty-five years ago, Manabe and Wetherald (1967) obtained a similar conclusion by the use of a radiative-convective atmospheric model. A warmer surface also normally results in less snow and ice at high latitudes, which means lower planetary albedo and an increased radiative forcing (ΔQ_{ice}). The result is more net radiant heat input to the climate system. Temperature-ice albedo feedback has both fast-reacting (sea ice and land snow) and slowly reacting (glacial ice) components.

A potentially large but presently poorly understood fast radiative feedback is due to clouds (ΔQ_{cld}). This is often broken down into a number of sub-processes. (For more about cloud feedback, see Crutzen and Golitsyn, 1992; this volume.) Other radiative feedbacks are heat fluxes from greenhouse gases or reflective aerosols driven from the oceans (ΔQ_{ocn}) and the biosphere (ΔQ_{bio}).

This leads to the following simple equation for the net heat flux to the Earth's surface:

$$H = \Delta Q_0 - \lambda \Delta T, \text{ where}$$

$$\lambda = \lambda^* - \left(\frac{\Delta Q}{\Delta T}\right)_{wv} - \left(\frac{\Delta Q}{\Delta T}\right)_{ice} - \left(\frac{\Delta Q}{\Delta T}\right)_{cld} - \left(\frac{\Delta Q}{\Delta T}\right)_{ocn} - \left(\frac{\Delta Q}{\Delta T}\right)_{bio} \quad \text{...[2]}$$

is the radiative damping coefficient. Notice that the basic blackbody dampingcoefficient λ^* is reduced if the feedback heating rates, the $(\Delta Q/\Delta T)_x$ terms, are positive. This is called positive feedback. When feedbacks are positive, the radiative damping coefficient decreases, and the climate sensitivity increases. Conversely, a negative feedback tends to decrease the climate sensitivity.

Under hypothetical steady state conditions, after a sufficient length of time subsequent to the imposition of ΔQ_0, the net heat flux H will vanish. The global warming would then depend only on the change in direct radiative forcing (ΔQ_0) and the radiative damping coefficient ($\Delta T = \Delta Q_0/\lambda$).

5.2 *Model predictions of climate feedbacks*

It is possible to use climate models to estimate the radiative damping coefficient and temperature changes for different feedback processes (Schlesinger, 1985; Mitchell, 1989). Typical values derived from current climate models are indicated in Table 3.

Of course, these estimates are only as good as the underlying feedback process models. Major studies are under way to determine the sensitivity of general circulation models (GCMs) to their own internal representations of feedback processes. This work is being coordinated by the Program for Climate Model Diagnosis and Intercomparison (PCMDI) at Lawrence Livermore National Laboratory. This work will help to explain why different models estimate a different climate sensitivity for the same radiative forcing. Although intercomparisons of climate models may seem more preoccupied with the meta-universe of simulations than with the Earth System itself, they are essential if we are to understand how feedbacks work.

5.3 *Models as laboratories of alternatives*

In the preceding section, I discussed a number of useful things that have been learned about feedbacks from the climatic history of the Earth and Earth-like planets. But paleoclimatologists who study ancient climates are at best passive observers. They cannot go back in a time machine to intentionally experiment with, or change, climate history. As humanity conducts its Grand Geophysical Experiment with greenhouse gases, numerical simulations are perhaps the only way to conduct parallel "laboratory experiments."

It is an article of faith that useful results will emerge from these simulations before substantial climate changes in the real world take place. Hopefully the results will arrive soon enough to allow policy to make a difference.

5.4 *Model uncertainties and cloud radiative feedback*

Most of the variability between global climate models can be traced to differences in cloud radiative feedback. This is illustrated in Table 3. When analyzing a CO_2 doubling since the industrial era, the predictions range from 1.5 ° to 4.8 °C. If model-to-model differences in climate sensitivity were entirely due to differences in the way cloud radiative feedback processes are depicted, we would expect the planetary cooling rate to be correlated by the feedback parameter. Figure 3, derived from the intercomparison by Cess et al. (1989) of 14 GCMs, shows this is very nearly the case.[5]

Notice in Figure 3 that different global climate models predicted different cloud radiative feedbacks — ranging from slightly negative to strongly positive. This dramatically illustrates how uncertainties in cloud radiative feedback translate into uncertainties in estimates of global climate sensitivity. Unfortunately we cannot say from an examination of the models alone, which values of cloud radiative feedback are the most realistic. Observational studies using satellites and selected field measurements are underway to resolve this issue. The main point we can make is that radiative feedbacks can strongly impact equilibrium climate sensitivity.

5.5 *Models and the time scale of climate response*

The atmospheric build up of greenhouse gases is occurring on time scales that range from decades to centuries. If the response of global climate to continually increasing greenhouse gas concentrations were immediate, it could be calculated from equilibrium atmospheric general circulation models (GCMs). There are many such models, and as we have seen, they exhibit quite different climate sensitivities. How-

[5] (The symbol λ is defined by these authors as the gain — the reciprocal of the planetary cooling rate defined here. Hopefully there will be no confusion on this point if the differences in symbology are kept in mind.)

Table 3. Effects of "fast" radiative feedbacks on planetary cooling and equilibrium climate response to carbon dioxide doubling.

Feedback	Radiative heating per unit temperature rise [W m^{-2} °C^{-1}]	cumulative radiative damping coefficient, λ [W m^{-2} °C^{-1}]	cumulative change in equilibrium temp., ΔT [°C][a]
Blackbody cooling	-4.0	-4.0	1.1
Water vapour[b]	1.4	2.6	1.7
Sea ice/land snow[c]	0.4	2.2	2.0
Cloud[d]	-0.8 to 1.3	3.0 to 0.9	1.5 to 4.8
Oceans, biosphere	?	?	?

[a] Computed from $\Delta T = \Delta Q_0/\lambda$, where $\Delta Q_0 = 4.4$ W m^{-2} for a CO_2 doubling.
[b] Sum of water vapour greenhouse, lapse rate and baroclinic instability feedbacks.
[c] Omits glacial ice feedback.
[d] Sum of cloud cover, altitude, formation, liquid water and radiative property feedbacks.

ever, the response of the climate system will not be immediate but will be delayed by the time it takes to warm the oceans, an effect called thermal inertia. One can imagine, on entering a cold house, a lag from the moment one turns up the thermostat, until the room actually becomes warm. This is a similar phenomenon.

A complicating factor is that the atmosphere and the oceans come to equilibrium on different time scales — about a month for the atmosphere versus a thousand years for the oceans (Hoffert and Flannery, 1985). This delayed response is called the transient climate change, and to calculate it, it is necessary to employ a coupled atmosphere-ocean climate model. There are fewer such models than atmospheric GCMs, and they contain additional uncertainties, including the very uncertain cloud-radiative feedback.

Because of the long and varied time scales of climate responses, the sheer computational requirements for simulating the coupled atmosphere/ocean system can become enormous. Perhaps future generations of parallel-processing supercomputers will routinely provide detailed transient climate projections for various locales and seasons. Indeed, a comprehensive analysis of the impact of global climate change requires such detailed information.

But at this point in time, it has proven useful in comparative assessments of climate response to radiative forcing, to use simpler ocean climate models, of the type proposed by Hoffert et al. (1980). Such a model, for example, was used in the recent IPCC assessment of changes in potential greenhouse gas concentration, and their effects over the next century (Houghton et al., 1990).

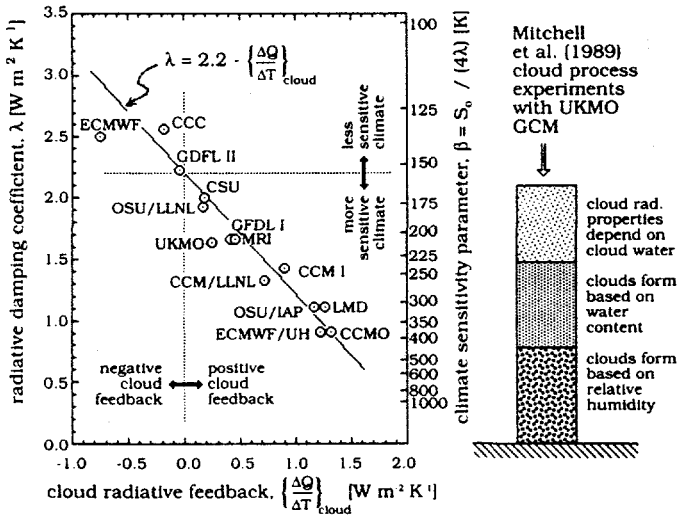


Figure 3. The effect of cloud processes on climate sensitivity. The left panel shows the radiative damping coefficient (left vertical axis) and climate sensitivity parameter β (right vertical axis) as a function of cloud radiative feedback $(\Delta Q/\Delta T)_{cld}$ in 14 general circulation models. GCM sensitivities and acronym labels are those given in the GCM inter-comparison by Cess et al. (1989). Notice that different global climate models predict different cloud radiative feedbacks — ranging from slightly negative to strongly positive (see also Table 3, "cloud"). The bar graph on the right, derived from Mitchell et al. (1989), shows the effect of changing cloud process models (hence cloud feedback) on the radiative damping coefficient and climate sensitivity.

6 Deriving Climate Sensitivity from Paleoclimate Data

Uncertainty about climate sensitivity has had great political impact; it has raised difficult questions about the importance of implementing policies to reduce the risks of rapid climate change. To address some of these issues, Curt Covey of Lawrence Livermore National Laboratory and I undertook a study of whether global climate sensitivity and transient climate change could be derived from paleoclimatic data, as an alternative to general circulation modelling (Hoffert and Covey, 1991). The methodology requires reconstructing both the radiative forcing and the temperature responses of selected ancient climates. Such reconstructions are based on proxy data — including such diverse archives of climate information as tree rings, lake sediments, ice cores, pollen, loess, ocean cores, corals, paleosols, geomorphic features and sedimentary rocks. The systematic reconstruction of climate and atmospheric composition from these proxy records is both an art and a science (see Box).

One useful principle of empirical science is that interpolation is safer than extrapolation, so our first step was to look at the sensitivity implied by climates very different from our own. Presumably, the future climate will fall between the present climate and climatic extremes observed over geologic history.

One hundred million years ago, in the middle of the Cretaceous Era, the Earth was about 10 °C warmer than today. Though the sun was slightly dimmer, this was more than offset by the greater sunlight absorption by a darker, more ocean-covered planet. Substantial greenhouse heating was provided by high concentrations of atmospheric carbon dioxide, six to ten times higher than present levels. At the other extreme, we considered the deep freeze of the Last Glacial Maximum (LGM), 18,000 years ago.

An orbital explanation for interglacial cycling was first proposed by Milutan Milankovitch (1920). He held that cyclic changes in seasonal solar radiation, caused by periodic variation in the Earth's orbit, drives climate fluctuation on glacial-interglacial timescales ranging from 10,000 to 100,000 years. However, the annual mean insolation — the average amount of sunlight incident on the Earth over the year — remains virtually constant. It is now strongly suspected that what causes the Ice Ages is the persistence of snow cover from year to year and changes in concentration of greenhouse gases resulting from the seasonal and latitude changes in sunlight. In a sense, the changes in the Earth's orbit are thought to be the pacemaker of Ice Ages, though they do not themselves cause the large climate changes.

The record of these changes — particularly the changes in CO_2 and CH_4 concentration — can be observed by studying polar ice cores. The best evidence comes from the 160,000-year-long record in the ice drilled at Vostok station on the Antarctic plateau. When combined with data from Greenland ice cores, these records reveal much about the last global-scale glacial-interglacial cycle (Chappellaz et al., 1990). These data suggest that, during the coldest glacial periods, sulphate aerosol particles were more abundant and the atmospheric concentration of CO_2 was 80 ppm below pre-industrial levels.

Most significant, for our present concerns, is the sensitivity of climate to these changed conditions — both at the Last

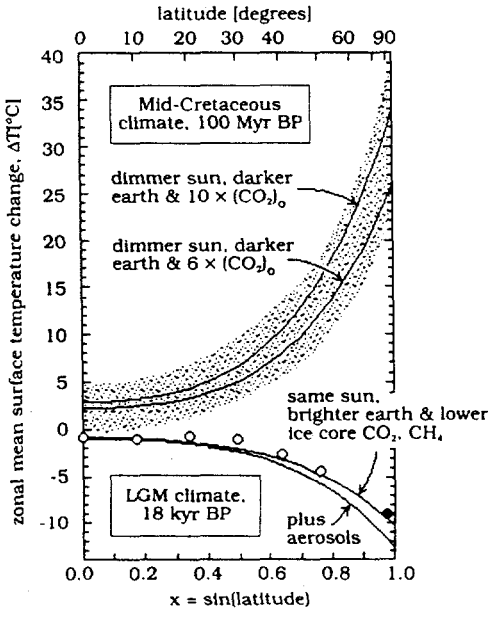


Figure 4. Two historical curves show a similar relationship between temperature change and latitude, when scaled to the radiative forcing of the time — one during the Cretaceous period, 100 million years BP, and the other during the Last Glacial Maximum (LGM), 18,000 years BP.

The stippled region (above) shows the range of Cretaceous climate response as the Earth grew warmer. This was obtained by subtracting the present zonal mean temperature distribution from that of 100 million years BP and averaging the two hemispheres. The LGM curve (below) is synthesized from sea surface temperature changes derived in the CLIMAP (1976) study (small open circles) and from air temperature on the Antarctic plateau, recorded in the Vostok ice core deuterium isotope record (small solid circle).

In both the Cretaceous and LGM reconstructions, the solid curves are derived from a hypothetical "universal" zonal temperature response, scaled to the positive and negative radiative forcing reconstructed for 100 million and 18 thousand years ago. In both cases, the global mean planetary cooling rate was 2.2 W m^{-2} K^{-1}, and the CO_2 doubling sensitivity was 2.0 °C, including the effect of cloud feedback.

Glacial Maximum and during the mid-Cretaceous. As illustrated in Figure 4, both the cooling of 18,000 years ago and the warming of 100 million years ago showed the same pattern of temperature change versus latitude, and very nearly the same climate sensitivity. Though one period was cold and the other warm, the global average temperature change and the distribution of temperature with latitude were quite similar — when scaled to the radiative forcing of those times. Both periods exhibited relatively weak temperature response in the tropics, compared to large high-latitude amplification of warming or cooling. We were impressed that both the very hot and very cold paleoclimates could be recovered with the same assumption about the sensitivity of global climate to a doubling of atmospheric CO_2, i.e. approximately 2 °C. This is also roughly equivalent to the sensitivity of clear sky atmospheric columns today, based on satellite measurements (Hoffert and Covey, 1991). The implication of this is that while the net effect of cloud-related feedback processes is small — only about 10% or less of total sensitivity — the positive feedback effect of water vapour is quite significant. Water vapour increases climate sensitivity to doubled CO_2, from about 1.1 °C to 2 °C.

In light of the possibility of adaptation, the impact of climate change on natural ecosystems may depend more on the rate of climate change than on the ultimate temperature change. To understand the effects of future increases in greenhouse gas concentrations on observed rates of change in the next century, a number of additional factors must be considered. Ocean thermal inertia, for example, could slow the observable rate of temperature increase caused by any greenhouse gas buildup.

To explore the implications of paleocalibration for 21st Century warming rates, we re-ran the IPCC ocean climate model with the Business as Usual scenario, assuming a climate sensitivity to doubled CO_2 of 2.0 °C. Not surprisingly, we found warming rates of about 0.3 °C per decade over most of the 21st Century. This would produce a climatic warming wave, sweeping poleward at about 6 kilometres per year, on average for the Earth as a whole. However, climatic changes would be much smaller in the tropics, and larger at the high latitudes, so this velocity of the warming wave represents a characteristic value for temperature latitudes. This is faster by a factor of three to thirty than the rate tree species could follow moving climate zones by spreading their seeds (Shugart et al., 1986). And this does not account for barriers to poleward migration from urbanization and agriculture. Of course, humanity could intervene to accelerate the migration of desirable tree species, using techniques akin to forest management, but the threat to the survival of natural ecosystems — what some have called "the end of nature" — is a major issue for environmental ethics and economics. If, as Vellinga and Swart (1991) have proposed, global warming rates beyond a "green" limit of 1 °C per century are risky, our work suggests that the IPCC "Business as Usual" scenario would almost certainly pose a severe problem in the 21st Century.

We can now say with some confidence that if climate sensitivity were as sluggish as some critics suggest, it is unlikely that the Earth would have experienced the Cretaceous hothouse and Pleistocene chills. Our paleoclimate-derived sensitivity, while not so large as that of some climate models, is large enough to worry about. Those disinclined to

Paleoclimates and the Future Climate

Despite some excellent large scale studies like CLIMAP and COHMAP, which have attempted to reconstruct world climate since the Last Glacial Maximum, paleoclimatologists have tended to emphasize local as opposed to global climate reconstructions. To project anthropogenic global warming, Budyko and Izrael (1991) proposed the "paleo-analogue" method. They employed reconstructed spatial distributions of temperature and precipitation patterns of "warm" analogue periods. They also presented maps of temperature and precipitation for their periods: the Holocene optimum (6000 years ago), the Eemian interglacial (120,000 years ago) and the Pliocene optimum (three million years ago).

There has been some reluctance by the climate community and the IPCC to accept these results, because the raw data and transfer functions used to reconstruct these periods are not well characterized in the literature and because the radiative forcing of their Holocene, Eemian and Pliocene periods is not well known. Budyko and Izrael derive their global climate sensitivity from other periods. Crowley (1990) has argued that, because of differences in the distribution of radiative forcing, there may be no warm period that is a satisfactory analogue for future climate.

Moreover, Crowley observes that because future temperatures may be increasing at a very high rate — as much as 2-4 °C per century, we will have a very unique combination of warm atmospheres and polar ice sheets. These are conditions very different from the pre-Pleistocene warm periods. General circulation modellers have tended to view paleoclimate reconstructions as sources of surface boundary conditions — sea surface temperatures and ice sheet locations — rather than test cases of climate sensitivity (Street-Perrott, 1991). All of these factors suggest that the potential of paleoclimate reconstructions to reduce the uncertainty in climate sensitivity has been insufficiently exploited, although a number of legitimate objections must still be addressed.

consider global warming a problem on the grounds that global climate model results are inherently uncertain might want to re-examine their positions in light of these results. Independent of complex computer models, there is good evidence of risk from continued greenhouse gas emissions. If certain wild card feedbacks not yet included in our paleo-calibration are activated, then warming rates would increase even more, and the risk to planetary ecosytems would have to be faced even earlier.

7 Biosphere and Ocean Feedbacks

What are some of these additional "wild card" feedbacks that have not yet found their way into global climate models? The Earth's biosphere has the potential to induce climatic feedbacks during global warming by changing the reflectivity of the Earth's surface or cloud cover (Charlson et al., 1987) or by releasing additional quantities of greenhouse gases (Lashof, 1989). Ocean circulation affects the distribution of carbon between the atmosphere and oceans, as well as the rate at which heat is transferred from the relatively warm surface layer to the cold depth. Both these processes can have unpredictable feedback effects on the rate of climate change.

7.1 Potential feedbacks from the terrestrial biosphere

So long as the "standing crop" of biomass remains constant, virtually all the carbon dioxide removed from the atmosphere through photosynthesis is regenerated by respiration, mostly as a result of the oxidation by bacteria of dead organic matter.[6] Of every 10,000 carbon atoms cycled from atmosphere to life forms and back this way, a few escape oxidation. They are buried in anoxic environments and transformed into an organic rock called kerogen. For every ten thousand carbon atoms in kerogen, a few are transformed into recoverable fossil fuels. In other words, the rate of fossil fuel formation is only a very small fraction of the rate of the primary productivity of organic carbon. But, because the process has been going on for hundreds of millions of years, the fossil fuel carbon reservoir is now large compared with the amount currently in the atmosphere. None the less, if human energy consumption continues to grow at its present rate, the global stock of fossil fuels would be burned to depletion in a few hundred years.

Table 4 presents current estimates of the fossil fuel reserve, given in units of both energy and carbon. The table shows that conventional fossil fuel resources contain about 2500 gigatonnes (Gt) of C. This estimate could easily double to 5000 Gt if shale oil were included.[7] If all this carbon were to remain in the atmosphere as CO_2, it would equal about 2400 parts per million by volume (ppmv), or six times present atmospheric levels. Table 5 indicates the amount of carbon in other reservoirs of the Earth System.

One well-documented feedback mechanism affecting atmospheric carbon dioxide is deforestation — either the direct cutting back of the standing crop of trees for agriculture, or inadvertently, through a forest ecosystems' inability to survive climate change. The problem of determining the effect of deforestation is complicated by regrowth of trees and vegetation, which takes some of the CO_2 out of the atmosphere that deforestation puts in. Despite the well-documented loss of tropical rainforests, it has been difficult to ascertain from limited sampling whether the world's forests are a net source or sink of carbon dioxide today. This underscores the difficulty of projecting whether terrestrial ecosystems will be sources or sinks of carbon in the future.

[6] Photosynthesis is the fundamental energy-gathering process of life: sunlight + carbon dioxide + water → organic carbon + oxygen. This occurs mainly in the leaves of terrestrial plants and in microscopic blue green algae in the ocean. The rate of organic carbon fixation by photosynthesis is called the primary productivity.

[7] The unit 1 Gt C equals 1 gigatonne carbon or 10^{15} grams of carbon, independent of the chemical compound it is incorporated in. If the carbon is in the form of atmospheric carbon dioxide, then 1 Gt C is equivalent to 0.473 parts per million by volume (ppmv) of CO_2.

Table 4. Energy and carbon content of proven fossil fuel and U-235 reserves.

Energy source[a]	Energy content [10^{21} J]	Carbon content [Gt C]
Hard coal	77.3	1900
Soft coal	17.7	440
Oil	5.2	104
Natural gas	4.2	57
Uranium (recoverable at < $130 kg)	1.4	0
Totals	**105.8**	**2501**

[a] Hammond, 1990, p.320, excluding shale derivatives

Table 5. Carbon in the earth system (after MacCracken et al., 1990).

Carbon reservoir	Carbon content [Gt C = 10^{15} g C]
Atmosphere (present value)	740
Terrestrial biota	560
Detritus and labile humus	1,400
Stable humus and peat	700
Marine biota	2
Dead marine organic matter	1,800
Inorganic dissolved carbon in oceans	35,000
Organic carbon rocks (see Table 4 for fossil fuels)	12,000,000
Carbonate rocks	94,000,000

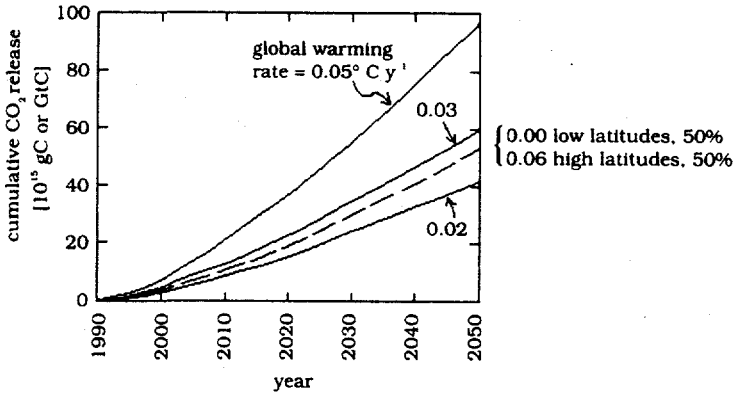


Figure 5. A model developed by Jenkinson et al. (1991) predicts that, if the global warming rate estimated by paleo-calibration occurs (3 °C per century), about 60 gigatonnes of carbon would be released from the soil during the next 60 years. This figure shows the model's estimates of carbon released as CO_2 from soil organic matter, from now to the year 2050.

Solid curves represent net carbon loss from the global stock of soil organic carbon, for 17 life zones excluding wetlands, at three different latitude warming rates: 0.02 °C y^{-1} (bottom), 0.03 (middle), and 0.05 (top). The dashed curve is net carbon loss, as a result of a 0.03 °C y^{-1} average warming rate, divided unequally between the zones, with a rise of 0.06 °C y^{-1} in high- latitude zones (tundra, boreal desert, cool desert, cool temperature steppe, moist boreal forest, wet boreal forest, and cool temperate forest, which together contain 50% of the total non-wetlands soil organic carbon), and no temperature rise in the ten remaining life zones, again excluding wetlands.

7.2 Soil feedbacks

Although most of the public's concern about the biota has focused on the release of carbon through deforestation, the first meter of soil contains even more carbon than do all the world's forests. Because of the size of the soil carbon reservoir, changes in the organic carbon content of soil could be even more important than changes in the standing crop of trees. Table 5 shows that 1400 Gt C is tied up as organic detritus and labile humus, typically in the uppermost meter of soil. This is twice as much carbon as is now contained in the atmosphere. One effect of global warming could be to accelerate respiration and decomposition of soil organic matter, thereby releasing more CO_2 to the atmosphere and further enhancing the warming trend. As Figure 5 illustrates, a model developed by Jenkinson et al. (1991) predicts that, if the global warming rate estimated by paleo-calibration occurs (3 °C per century), about 60 gigatonnes of carbon would be released from the soil during the next 60 years.

On the 100-year time scale, increased soil respiration would act as a modest positive climate feedback, reinforcing the rate of global warming. But on the longest geologic time scales, an increase in soil respiration could become a negative feedback, as the rate of atmospheric CO_2 removal to sedimentary carbonates by calcium-silicate weathering increases.

7.3 Potential feedbacks from ocean life

The amount of carbon in the marine biota is less than 1% that of the land biosphere, though its primary productivity is comparable. The marine biosphere "runs faster" than the land biota, with organic carbon lifetimes of plankton measured in days as opposed to decades for trees. Despite the small biomass of the marine organisms, changes in their productivity can be invoked as a possible explanation for variations of atmospheric CO_2 of 170 Gt C (80 ppmv) recorded between glacial to interglacial periods in polar cores at 170 Gt C (Chappellaz et al.,1990).

Although the entire living marine biosphere contains only 2 Gt C of biomass, plankton can exert considerable leverage on the distribution of the surrounding 35,000 Gt of inorganic carbon dissolved in the ocean (Table 5). The combination of photosynthesis near the surface and the respiration of fecal pellets falling through the water column drives a "biological pump" that moves carbon from the warm surface layer into the deep ocean. During periods of stable climate, this biological pump maintains a vertical gradient in dissolved inorganic carbon (Volk and Hoffert, 1985). A change in ocean productivity affects the distribution of this dissolved carbon, as well as the concentration of carbon dioxide at the surface, and hence the atmospheric CO_2 concentration that is in equilibrium with the surface ocean.[8]

[8] Mix (1989) used data on planktonic foraminifera species in modern and ice-age Atlantic sediments to assess spatial patterns of changes in marine productivity. These changes, if extrapolated to the global ocean, support models in which a significant portion of CO_2 changes are driven by variations in productivity. This is an interesting finding, but incomplete for purposes of predicting future CO_2 feedbacks on climate. Some other mechanism, like changes in sea level or ocean circulation, is needed to explain how the intensity of the ocean's biological pump changes when the climate changes.

One of the factors controlling the biological pump is the supply of essential nutrients to marine organisms. Over most of the ocean surface, phosphate and nitrate are the "biolimiting" nutrients, virtually depleted in most surface waters where phytoplankton use all there is. Because they are rich in nutrients, high-latitude waters surrounding Antarctica continent are relatively high in productivity. But the productivity would be even higher if dissolved phosphate nutrients near the surface were fully utilized by phytoplankton. In the Southern Ocean, Martin et al. (1990) suggest that another element, iron, is biolimiting. They hypothesize that significant amounts of atmospheric carbon might be sequestered by fertilizing the Southern Ocean with supertanker loads of iron filings. However, Peng and Broecker (1991) conclude, on the basis of model calculations, that even if iron fertilization worked perfectly it could not significantly reduce atmospheric CO_2 concentrations. Joos et al. (1991) did find, on the basis of their model calculations, that atmospheric CO_2 concentrations could be lowered by up to 190-227 Gt C (90-107 ppmv) by the year 2100, for the IPCC Business as Usual scenario. These authors believe, however, that such large biotic uptake is unlikely in practice — again, in part, because of feedbacks. For example, organic carbon produced in the surface waters of Antarctica would be oxidized at more equatorward latitudes, returning CO_2 to the water column and eventually to the atmosphere.

A possible link between ozone depletion and climate change is a CFC-generated ozone hole around Antarctica — increasing ultraviolet radiation and killing plankton and other marine organisms in the food chain in the Southern Ocean. Without these organisms, the global level of net primary productivity would fall, CO_2 would build up more rapidly in the atmosphere, and the warming effect would increase. Another link arises from the fact that a CO_2 increase leads to a cooling of the stratosphere at the same time that it warms the surface and the lower atmosphere. The Antarctic ozone hole is indirectly caused by the occurrence of stratospheric water-nitric acid clouds, which form only at temperatures below about -80 °C (Solomon, 1990). By increasing the region over which these clouds can form, the CO_2 could very well exacerbate stratospheric ozone depletion. These are only two of a host of potential feedbacks between biogeochemical cycles and climate change.

7.4 Ocean circulation

One of the most intriguing wild cards is the possibility of changes in ocean circulation. The hemispheric asymmetry of the present continents produces a dramatic difference in the way deep water forms in the two hemispheres. In contrast to the "open" Southern Ocean around Antarctica, the main site of seasonal sea ice formation in the Northern Hemisphere is the Arctic ocean, a landlocked water body virtually inacces-

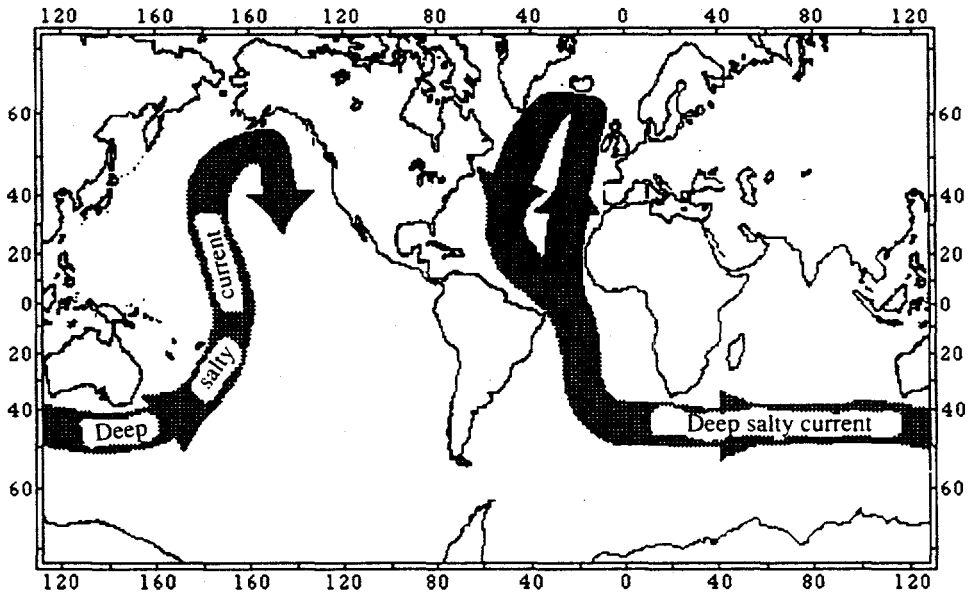


Figure 6. A mechanism operating in today's ocean, which may create rapid climate shifts. Salt-laden North Atlantic Deep Water (NADW) flows down the length of the Atlantic, around Africa through the Southern Indian Ocean, and finally northward in the deep Pacific. This acts as a large-scale salt transport system, compensating for the transport of water vapour through the atmosphere from the Atlantic to the Pacific. This system may be self-stabilizing. Records from ice and sediment cores suggest that it was disrupted in glacial times and replaced by an alternate mode of operation. Source: Adapted from Broecker, 1987. Reprinted by permission from Nature, vol. 328, pp. 123-126. Copyright (c) 1987, MacMillan Magazines Ltd.

sible to the North Pacific and communicating weakly (if at all) with the North Atlantic. It is presently believed that the source of North Atlantic Deep Water (NADW) is the permanent density-stratified layer of sea water of the Iceland and Greenland Seas. This is a region of high salinity, created because evaporation removes more moisture there than is replaced by local precipitation.

This salty North Atlantic water enters the Atlantic through the Norwegian sea; as it moves southward down the Atlantic it flows beneath the high-salinity (but warmer) Mediterranean outflow, and never does attain the density needed to sink to the bottom of the Atlantic. But on reaching the Southern Ocean, the NADW mixes to the surface and is cooled in the Antarctic Circumpolar Current, where it becomes part of the source water for the densest, coldest and deepest bottom water that feeds all the world's oceans: Antarctic Bottom Water (AABW).

Figure 6 illustrates the paths of deep, dense salty currents in the oceans. Broecker et al. (1985) speculated that these currents have two alternate modes of operation, one in which NADW formation is strong and another in which it is weak or absent.[9] A temporary shut-off of NADW may have occurred during the so-called Younger Dryas period, 10,800 to 10,000 years ago, when fresh water from the Mississippi may have been diverted to the North Atlantic, punctuating the gradual warming since the Last Glacial Maximum with a short-lived global cooling. This was followed by a very rapid rewarming and return to the slow emergence from the Last Glacial Maximum. It has also been suggested that the North Atlantic was flooded with fresh water, via the St. Lawrence River's outlet, by an abrupt drainage of Lake Agassiz, a giant mid-continental lake of meltwater. What is known, from Greenland ice core records, is that there was a local temperature rise as large as 7 °C over 50 years, and a 50% increase in rainfall over 20 years, during this Younger Dryas Termination (Dansgaard et al., 1989). This has prompted speculation that a future shut-down of NADW could occur during the next century's global warming by greenhouse gases, possibly threatening Europe with localized cooling in the midst of a more general global warming (Calvin, 1991). In addition to such direct climatic effects, rapid changes in ocean circulation could trigger changes in ocean chemistry that feedback on global climate.

The implication of Broecker's hypothesis is that rapid changes in ocean circulation can trigger changes in ocean chemistry that feed back on climate (Broecker and Peng, 1986, 1989; Boyle, 1988). Such circulation-driven changes in ocean carbon chemistry (or in the biological productivity of the ocean) may explain the rapid increases of atmospheric carbon dioxide and methane that are observed in ice cores and seem to occur during glacial terminations.

We do not understand the dynamics of glacial-interglacial transitions well enough yet to make predictions about deep water circulation feedbacks. We need to know much more about the natural variability of the oceans. The most sophisticated current atmospheric-ocean climate model, run for 100 simulated years with realistic build ups of CO_2, does not suddenly flip its NADW circulation (Manabe et al., 1991). But this does not prove that the real ocean will not experience such a change. There are indications that climate variability may be partly due to inherently unpredictable chaotic-dynamic interactions between the atmosphere and the oceans (Gaffin et al., 1986). The best we can say at this point is that an abrupt change in ocean circulation during the next century is an important wild card that cannot be ruled out.

Carbon dioxide is not the only greenhouse gas that can be driven to the atmosphere by global warming. No one has definitively explained the large rate of methane increase currently in progress — but some feedback processes may be involved. Gas clathrates are crystalline inclusions of water (hydrates) bound with light hydrocarbons — primarily methane. The crystals are found in abundance beneath Arctic permafrost and along the continental shelf below the ocean floor. Many authors have hypothesized a climatic feedback that would involve methane being degassed from high-latitude methane hydrate reservoirs as a consequence of global warming (Bell, 1982; Khalil and Rasmussen, 1988; Lashof, 1989; MacDonald, 1990). No observational evidence yet supports this hypothesis. But, depending on the actual rate of methane release if it occurred, this could be a significant positive climate feedback.

On the other hand, data from Antarctic ice cores indicate that atmospheric methane concentration varied by a factor of two in remarkably close association with glacial-interglacial climatic changes during the past 160,000 years (Chappellaz et al., 1990). These changes could have involved methane hydrates (Nisbet, 1990), or other methane sources such as natural wetlands (Fung et al., 1991). In any case, the geological record strongly suggests that feedbacks between climate and methane sources or sinks will affect methane concentration as the climate warms. Observational studies currently in progress will hopefully help to determine the potential importance of climate-methane feedbacks, which are likely to involve methane sources both from hydrates and from wetlands.

8 The Dilemma for Policy Makers

8.1 Lessons of the IPCC scenarios

In 1990, the Intergovernmental Panel on Climate Change (IPCC) sought to assess the effect of various greenhouse emission policies on climate change. Because of publication time constraints, Houghton et al. (1990) developed a set of

[9] These speculations were supported by Manabe and Stouffer (1988), who found two stable equilibria in time integrations of a coupled atmosphere-ocean climate model. These had identical boundary conditions but different initial conditions. In one of these solutions, a fairly realistic interhemispheric thermohaline circulation was maintained, in the other it was absent. The solution is quite sensitive to the salt water versus fresh water balance in the North Atlantic.

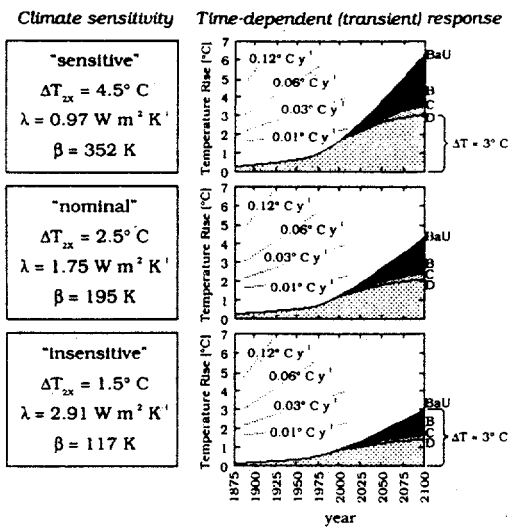


Figure 7. Twelve predictions of the global mean surface temperature response, projected to the year 2100. The left column shows three equilibrium climate sensitivities. For each sensitivity, the right column shows temperature changes under each of four assumptions regarding future greenhouse gas concentrations in the atmosphere. The four assumptions are the IPCC BaU (Business as Usual), B, C, and D scenarios (adapted from Houghton et al., 1990).This figure shows the importance of climate sensitivity to policy decisions in the face of uncertainty. If sensitivity is high (upper panel), then draconian rollbacks are needed to keep global warming below 3 °C by the year 2100. If climate sensitivity is low (lower panel), then the same three- degree warming will occur for Business as Usual.

scenarios which projected the evolution of greenhouse gas concentrations in the atmosphere over the 21st Century, and the corresponding radiative forcing.

A more policy-relevant approach would have been to develop scenarios based on the chain of causality from economic development to emissions to concentrations and finally to global warming. This is possible in principle, because there is a causal link between the rate of greenhouse gas emissions over time, and the concentration history of greenhouse gases in the atmosphere that results from these emissions. For example, the CO_2 concentrations versus time can be computed from global carbon-cycle models, which account for the distribution of carbon in the atmosphere, in the ocean, and in terrestrial and marine biota as CO_2 is emitted from fossil fuel burning and possibly deforestation. Hopefully, as part of the ongoing IPCC process, comprehensive systems models will emerge which permit analysis of all the links in the causal chain. This would make it easier to separate the factors which are well known from the factors which are uncertain.

None the less, there is much to be learned about the impact of climatic feedbacks and climate sensitivity on global warming policy options from the responses of an atmosphere-ocean-climate model to the four IPCC "scenarios" (Figure 7).

In an approximate sense, the first scenario, "Business as Usual (BaU)," portrays a 21st Century world in which greenhouse gas emissions increase as they will if fossil fuels remain the dominant world energy source, and population growth continues as in current projections, particularly in the third world. "BaU" is meant to represent an average projection, though it could significantly underestimate emissions from developing countries if they fuel the economic growth they aspire to with relatively abundant (and cheap) local coal reserves (Table 4). Given the energy and technology transfer policies of the developed countries, there would be a strong economic incentive for the developing countries to exploit relatively inexpensive coal reserves.

Scenarios "B" through "D" represent progressively stronger greenhouse gas emissions controls, with "B" corresponding approximately to a freeze in greenhouse gas emissions at 1990 levels, and "D" a draconian rollback in which fossil fuels are cut back by 60-80% during the next 50 years (Krause et al., 1989; Houghton et al., 1990, Annex).

Figure 7 shows the projected temperature rise of all four scenarios, for each of three different climate sensitivities. It is not surprising that the greatest warming takes place under Business as Usual with the highest sensitivity; and the smallest warming (although still more than 1 °C in the next century) is for the draconian rollback ("D") scenario with the least sensitive climate. But economic considerations make intermediate combinations of these results more interesting. For example, if climate sensitivity is high (upper panel), then draconian rollbacks are needed to keep global warming *below* three degrees Celsius by the year 2100. If climate sensitivity is near its lower uncertainty boundary (lower panel), then the same three degree warming will occur for Business as Usual. Though the global warming is approximately the same, the costs to implement a mitigation policy are vastly different.

8.2 Feedback uncertainties in policy making

The central dilemma of global warming policy is decision making in the face of large uncertainties. Some critics of global warming policy have emphasized the uncertainty in GCMs, and further argued that global warming will be near the lower end of these projections, implying thereby that corrective actions are premature. On the other hand, these projections are as likely to be underestimates as overestimates, and may be too conservative if the wild card feedbacks discussed in this chapter become important in the next century.

The problem is somewhat analogous to that faced by a patient whose physician has informed him that his condition requires surgery. Though the studies of the impact of this surgery are somewhat equivocal, to not make a decision is a decision in itself of major significance to his survival.

The central question for the policy maker is: What are allowable rates of greenhouse gas emissions that would maintain climate change below some "threshold of acceptability"? We cannot answer this question yet, but we can specify the ingredients of an answer. We need to know the climate sensitivity (including cloud feedbacks so poorly characterized by GCMs); how ocean heat storage links climate sensitivity to global warming rates; and how feedbacks from the biosphere and oceans, along with possibly compensatory cooling from sulphate aerosols, would affect the ultimate climate response. We also need better criteria specifying what is meant by "unacceptable" warming and rates of warming. This brings into play a panoply of philosophical, ethical, political, environmental, and ecological perspectives.

The decision making process is also complicated by fundamental cultural differences between scientists and policy makers that can render communication on science policy nonproductive. For professional reasons, the scientist is guarded, sometimes hedging his or her conclusions in with so many caveats that they are virtually useless to the policy maker. Politics, the art of the possible, requires decision making in the face of uncertainty. This volume attempts to bridge the gap between science and policy. And in this chapter, I have focused attention on how climatic feedbacks could impact global warming policy decisions. It is clear from this survey that the connections are many, and the ramifications virtually endless. Some possible "wild card" feedbacks or other significant factors have almost certainly been left out because no one has thought of them yet.

8.3 Charting the route of acceptable action

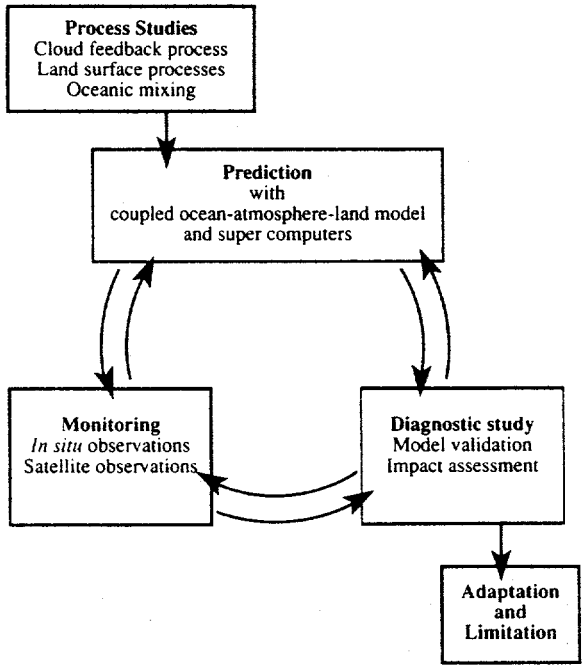
None the less, there is light at the end of the tunnel. We have found that a semi-empirical approach to estimating climate sensitivity can be quite useful. For example, the paleoclimatic

Triad Strategy for Improving Climate Prediction

Syukuro Manabe

Unfortunately, the projections of future climate change made by the Intergovernmental Panel on Climate Change (IPCC; Houghton et al., 1990) are subject to large uncertainties. These uncertainties reflect our inability to model the various processes that control any future climate change. It is therefore necessary to improve various components of climate models, such as the representation of cloud feedback and land surface processes.

In addition, it is essential to carefully assess the model's predictions of future climate change in light of the results obtained by monitoring actual climate changes and the factors causing these changes. Comparing the predicted and observed climate will eventually enhance our confidence in model prediction.



A comprehensive triad strategy for continuously improving the art and science of climate prediction is illustrated by the flow diagram in Figure 8. This strategy involves coordinating three concurrent activities. The first involves the monitoring of the coupled ocean-atmosphere-land surface system by *in situ* and remote sensing. The second applies state-of-the-art models to the prediction of future climate change. The third activity incorporates in-depth analysis of predicted changes in conjunction with the observations of the climate system.

The insights gained from this coordinated strategy would be indispensable — not only for enhancing our confidence in model prediction of future climate change, but also for adapting effectively to altered climate conditions and for reducing the rate of buildup in atmospheric concentration of greenhouse gases (see also Wigley, 1992; this volume).

record may provide an alternative path to general circulation models around the uncertainties of cloud-radiative feedback. Most experts in this field would probably agree that a resolution of the cloud feedback issue by more realistic cloud models — tested against observations and ultimately incorporated as parameterizations in general circulation climate models — is the most scientifically acceptable approach to resolving this issue. But that may be decades away. Satellite observations, too, could be very helpful, but the Earth Observing System is also decades away from operational status. In the meantime, the paleoclimatic record awaits analysis. We believe that its implications may be reliable enough to use as a basis for political decisions. Preliminary results with this method are encouraging enough to pursue it much more vigorously (Oeschger and Mintzer, 1992; this volume).

A common near-term goal for an international climate convention is a global emission rate freeze (Krause et al., 1989). A committee of the US National Academy of Sciences (NAS, 1991), studying this problem advocated a "no regrets" strategy — a response focused on zero- or negative-cost options like improved energy efficiency and fuel switching (Pachauri, 1992; this volume). Despite lingering criticisms on grounds of climate model uncertainties (Michaels, 1991), "No regrets" is generally perceived as an economically painless, if not particularly farsighted, policy. Indeed in 1990, even without a carbon agreement, annual global CO_2 emissions, at 5.803 gigatonnes of carbon (Gt C), were down slightly from 5.813 Gt C in 1989 (Wald, 1991) — a *de facto* emissions freeze.

Over the next few years, an emissions freeze based on energy conservation, efficiency and fuel switching is technologically feasible (Holdren, 1992; this volume). Given the pressures of population growth and economic development, it will become increasingly difficult to implement in the coming decades without new non-fossil energy technology, particularly as the already committed population growth of the next century makes itself felt. An important implication of the causal link between emissions and concentrations is that an emission rate freeze results in continually increasing atmospheric carbon dioxide concentrations, radiative heating and global warming. All of these continue to rise to some extent throughout the next century in all four IPCC scenarios (Figure 7), though the amount of warming clearly decreases as more stringent emissions controls are applied. If, for example, stabilization of greenhouse gas levels were needed, a worst-case scenario something close to a fossil-fuel phaseout in the next 50 years could become necessary. Very little contingency planning has gone into the transition from a fossil fueled to a solar/fusion powered world implied by scenario D. This is a further example of how climate modelling and systems analysis can identify key problems for policy makers.

In the immediate future, first-order policy questions will continue to be impacted by the uncertainties of climate feedbacks. Their resolution may be an acid test of the fledgling science of Global Change. I have argued in this chapter that semi-empirical methods, based on the history of the Earth's climate imprinted in the geological record, can provide a useful approach to this problem. Perhaps those policy makers unconvinced of the reality of global warming by GCM model predictions might reassess their positions based on the paleoclimate record.

Acknowledgments

It is a pleasure to thank C. Covey, L.D.D. Harvey, M.C. MacCracken, S. Manabe, K. Taylor and T. Volk for insightful contributions to this paper. The help of H. Jones in transmitting edited manuscripts via electronic mail was also invaluable. Research on the climate sensitivity, cloud feedback, and ocean modelling at New York University, some of which is reported here, was supported by the US Department of Energy, Carbon Dioxide Research Program, Office of Health and Environmental Research, and the National Institute for Global Environmental Change.

References

Allaby, M. and J. Lovelock, 1984: *The Greening of Mars*. New York, St. Martin's Press.

Averner, M.M. and R.D. MacElroy, eds., 1976: *On the Habitability of Mars: An Approach to Planetary Ecosynthesis*. NASA **SP-414**, Washington, DC, National Aeronautics and Space Administration.

Barron, E.J. and W.M. Washington, 1985: Warm Cretaceous Climates: High atmospheric CO_2 as a plausible mechanism. In *The Carbon Cycle and Atmospheric CO_2: Natural Variations Archean to Present*, E.T. Sundquist and W.S. Broecker, eds, *Geophys. Mong* **32**, 546-553, Washington, DC, American Geophysical Union.

Bell, P.R., 1982: Methane hydrate and the carbon dioxide question. In *Carbon Dioxide Review*: 1982. W.C. Clark, ed., 401-406, New York, Oxford University Press.

Berner, R.A., 1990: Atmospheric carbon dioxide levels over Phanerozoic time. *Science*, **249**, 1382-1386.

Boyle, E.A., 1988: The role of vertical chemical fractionation in controlling late Quaternary atmospheric carbon dioxide. *J. Geophys. Res.*, **93C**, 15,701-15,714.

Broecker, W.S., 1987: Unpleasant surprises in the greenhouse? *Nature*, **328**, 123-126.

Broecker, W.S. and T.-H. Peng, 1986: Carbon cycle, 1985: glacial to interglacial changes in the operation of the global carbon cycle. *Radiocarbon*, **28**, 309-327.

Broecker, W.S. and T.-H. Peng, 1989: The cause of the glacial to interglacial atmospheric CO_2 change. *Global Biogeochemical Cycles*, 3, 215-239.

Broecker, W.S., D.M. Peteet and D. Rind, 1985: Does the ocean-atmosphere system have more than one stable mode of operation? *Nature*, **328**, 123-126.

Budyko, M.I.,ed., 1991: Climate of the Future. Manuscript submitted by USSR delegation of Working Group VIII of US-USSR Agreement on Protection of the Environment for review as a possible supplement to MacCracken et al. (1990). US

National Climate Program Office, Suite 518, 1825 Connecticut Ave, N.W., Washington, DC.

Budyko, M.I. and Y.A. Izrael, 1991: *Anthropogenic Climatic Change*, pp. 277-318, Tucson, University of Arizona Press.

Budyko, M.I., A.B. Ronov and A.L. Yanshin, 1987: *History of the Earth's Atmosphere*. New York, Springer-Verlag.

Calvin, W.H., 1991: Greenhouse and Icehouse. In *Whole Earth Review*, **73**, Winter 1991, 106-111.

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. Morchette, D.A. Randell, M.R. Riches, E. Roeckner, U. Schlese, A. Slingo, K.E. Taylor, W.M. Washington, T.R. Wetherald and I. Yagai, 1989: Interpretation of cloud-climate feedback as produced by 14 atmospheric general circulation models. *Science*, **245**, 513-516.

Chappellaz, J., J.M. Barnola, D. Raynaud, Y.S. Korotkevich and C. Lorius, 1990: Ice-core record of atmospheric methane over the past 160,000 years. *Nature*, **345**, 127-131.

Charlson, R.J., J.E. Lovelock, M.O.Andreae and S.G. Warren, 1987: Ocean phytoplankton, atmospheric sulfur, cloud albedo and climate. *Nature*, **326**, 655-661.

CLIMAP Project Members, 1976: The surface of the ice age earth. *Science*, **191**, 1131-1137.

Crowley, T.J., 1990: Are there any satisfactory geologic analogs for a future greenhouse warming? *Journal of Climate*, **3**, 1282-1292.

Crowley, T.J. and G. North, 1991: *Paleoclimatology*. New York, Oxford University Press.

Dansgaard, W., J.W.C. White and S.J. Johnsen, 1989: The Abrupt Termination of the Younger Dryas Climate Event. *Nature*, **339**, 532-534.

Ellsaesser, H.W., 1990: A different view of the climatic effect of CO_2 — updated. *Atmósfera*, **3**, 3-29.

Fogg, M.J., 1989: The creation of an artificial dense Martian atmosphere: a major obstacle to the terraforming of Mars. *J. Brit. Interplanet. Soc.*, **42**, 577-582.

Fung, I., J. John, J. Lerner, E. Matthews, M. Prather, L.P. Steele and P.J. Fraser, 1991: Three-dimensional model synthesis of the global methane cycle. *JGR*, **96**, 13, 13033-13065.

Gaffin, S.R., M. I. Hoffert and T. Volk, 1986: Nonlinear coupling between surface temperature and ocean upwelling as an agent in historical climate variations. *J. Geophys. Res.*, **91**, 3944-3950.

Hammond, A.L., ed., 1990: *World Resources 1990-91: A Guide to the Global Environment*, New York, Oxford University Press.

Hoffert, M.I. and C. Covey, 1991: Projecting 21st Century greenhouse warming from paleoclimate data and ocean models. Submitted to *Nature*.

Hoffert, M.I. and B.F. Flannery, 1985: Model projections of the time-dependent response to increasing carbon dioxide. In *Projecting the Climatic effects of Increasing Carbon Dioxide*. M.C. MacCracken and F.M. Luther, eds., DOE/ER-0237, 149-190, Washington, DC, US Department of Energy, Atmospheric and Climate Research Division.

Hoffert, M.I., A.J. Callegari and C.-T. Hsieh, 1980: The role of deep sea heat storage in the secular response to climatic forcing. *J. Geophys. Res.*, **85** (C11), 6667-6679.

Houghton, J.T., G. Jenkins and J.J. Ephraums, 1990: *Climate Change: The IPCC Scientific Assessment*. New York, Cambridge University Press.

Jenkinson, D.S., D.E. Adams and A. Wild, 1991: Model estimates of CO_2 emissions from soil in response to global warming. *Nature*, **351**, 304-306.

Joos, F., J.L. Sarmiento and U. Siegenthaler, 1991: Estimates of the effect of Southern Ocean iron fertilization on atmospheric CO concentrations. *Nature*, **349**, 772-775.

Kasting, J.F., O.B. Toon and J.B.Pollack, 1988: How climate evolved on the terrestrial planets. *Scientific American*, February 1988, 90-97.

Khalil, M.A.K. and R.A. Rasmussen, 1988: Climate-induced feedbacks for the global cycles of methane and nitrous oxide. *Tellus*, **41B**, 554-559.

Krause, F., W. Bach and J. Koomey, 1989: How much fossil fuel can still be burned? In *Energy Policy in the Greenhouse*, Vol. I, Chapter I.4, International Project for Sustainable Energy Paths (IPSEP), El Cerrito, California.

Lashof, D.A., 1989: The dynamic greenhouse: Feedback processes that may influence future concentrations of atmospheric trace gases and climatic change. *Climatic Change*, **14**, 213-242.

Lindzen, R.S., 1990: Some coolness concerning global warming. *Bull. Am. Meter. Soc.*, **71**, 288-299.

Lovelock, J., 1988: *The Ages of Gaia: A Biography of Our Living Earth*. New York, W.W. Norton & Company.

MacCracken, M.C., A.D. Hecht, M.I. Budyko and Y.A. Izrael, eds., 1990: *Prospects for Future Climate: A Special US/USSR Report on Climate and Climate Change*. Chelsea, Michigan, Lewis Publishers.

MacDonald, G.J., 1990: Role of methane clathrates in past and future climates. *Climatic Change*, **16**, 247-281.

Manabe, S. and R.J. Stouffer, 1988: Two stable equilibria of a coupled ocean-atmosphere model. *J. Climate*, **1**, 841-866.

Manabe, S. and R.T. Wetherald, 1967: Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *J. Atmos. Sci.*, **24**, 241-259.

Manabe, S., R.J. Stouffer, M.J. Spelman and K. Bryan, 1991: Transient responses of a coupled ocean-atmosphere model to gradual changes of atmospheric CO_2. Part I: Annual mean response. *J. Climate*, **4**, 785-818.

Manowitz, B., ed., 1990: Global Climate Feedbacks, *Proceedings of the Brookhaven National Laboratory Workshop* June 3-6, 1990, Washington, DC, US Department of Energy, Atmospheric and Climate Research Division.

Marshall Institute, 1989: *Scientific Perspectives on the Greenhouse Problem*. Washington DC, George C. Marshall Institute.

Martin, J.H., S.E. Fitzwater and R.M. Gordon, 1990: Iron deficiency limits phytoplankton growth in Antarctic waters. *Global Biogeochemical Cycles*, **4**, 5-12.

McKay, C.P., 1987: Terraforming: Making an Earth of Mars. *The Planetary Report*, **VII**(6), 26-27.

McKay, C.P., O. B. Toon and J. S. Kasting, 1991: Making Mars habitable. *Nature*, **352**, pp. 489-496.

McKibben, W., 1989: *The End of Nature*. New York, Random House.

Michaels, P., 1991: Global warming: The new National Academy of Sciences report, *Cato Review of Business and Government*, Summer 1991, 20-23.

Milankovitch, M., 1920: *Theorie Mathematique des Phenomenes Thermiques Produits par la Radiation Solaire*, Paris, Gauthier-Villars, Paris.

Mitchell, J.F.B., 1989: The "greenhouse" effect and climatic change. *Rev. Geophys.*, **27**, 115-139.

Mitchell, J.F.B., C.A. Senior and W.J.Ingram, 1989: CO_2 and climate: a missing feedback? *Nature*, **341**, 132-134.

Mix, A.C., 1989: Influence of productivity variations on long-term atmospheric CO_2. *Nature*, **337**, 541-544.

NAS (National Academy of Sciences), 1991: *Policy Implications of Greenhouse Warming*. Washington, DC, National Academy Press.

Nisbet, E.G., 1990: The end of the ice age. *Can. J. Earth Sci.*, **27**, 148-157.

Peng, T.H. and W.S. Broecker, 1991: Dynamical limitations on the Antarctic iron fertilization strategy. *Nature*, **349**, 227-229.

Ramanathan, V. and W. Collins, 1991: Thermodynamic regulation of ocean warming by cirrus clouds deduced from observations of the 1987 El Niño. *Nature*, **351**, 27-32.

Raval and V. Ramanathan, 1989: Observational determination of the greenhouse effect. *Nature*, **342**, 758-761.

Revelle, R.R. and H.E. Suess, 1957: Carbon dioxide exchange between atmosphere and ocean and the question of an increase of atmospheric CO_2 during past and present decades. *Tellus*, **9**, 18-27.

Schlesinger, M.E., 1985: Analysis of results from energy balance and radiative-convective models. In *Projecting the Climatic Effects of Increasing Carbon Dioxide*, DOE/ER-0237, 281-319, M.C. MacCracken and F.M. Luther, eds., Washington, DC., US Department of Energy, Atmospheric and Climate Research Division.

Schneider, S.H., 1990: The global warming debate heats up: an analysis and perspective. *Bull. Am. Meter. Soc.*, **71**, 1292-1304.

Shugart, H.H., M. Ya. Antonovsky, P.G. Jarvis and A.P. Stanford, 1986: CO_2, climatic change and forest ecosystems. In *The Greenhouse Effect, Climatic Change and Ecosystems*, 475-521, B. Bolin, B.R. Döös, J. Jager, and R.A. Warrick, eds., SCOPE 29, 475-521, New York, Wiley.

Solomon, S., 1990: Progress towards a quantitative understanding of Antarctic ozone depletion. *Nature*, **347**, 347-354.

Street-Perrott, F.A., 1991: General circulation (GCM) modelling of paleoclimates: a critique. *The Holocene*, **1**, 74-80.

Vellinga, P. and R. Swart 1991: The greenhouse marathon: A proposal for a global strategy. *Climatic Change*, **18**, viii-xii.

Volk, T. and M.I. Hoffert, 1985: Ocean carbon pumps: analysis of relative strengths and efficiencies of in ocean-driven atmospheric CO_2 changes. In *The Carbon Cycle and Atmospheric CO_2: Natural Variations Archean to Present*, E.T. Sundquist and W.S. Broecker, eds., *Geophys. Monogr. Ser.*, **32**, 91-110, Washington, DC, American Geophysical Union.

Volk, T., 1987: Feedbacks between weathering and atmospheric CO_2 over the last 100 million years. *Am. J. Sci.*, **287**, 763-779.

Volk, T., 1989: Rise of angiosperms as a factor in long-term climatic cooling. *Geology*, **17**, 107-110.

Wald, M.W., 1991: Carbon dioxide emissions dropped in 1990, ecologists say. *New York Times International*, Sunday, Dec. 8.

Walker, J.C.G., P.B. Hays and J.F. Kasting, 1981: A negative feedback mechanism for the long-term stabilization of Earth's surface temperature. *J. Geophys. Res.*, **86**, 9776-9782.

White House, 1991: *National Energy Strategy: Executive Summary*, First Edition 1991/1992, February 1991, Washington, DC.

infrared opacity of the troposphere as more water evaporates into air columns over warmer surfaces. This reduces radiative damping and increases climate sensitivity relative to black-body cooling. GCMs show positive water vapour feedback, and agree to within ~10% on clear-sky damping and climate sensitivity[1]; $\lambda_{clear} = 2.2 \pm 0.2\ W\ m^{-2}\ K^{-1}$, $\Delta T_{2\times} = \Delta Q_{2\times}/\lambda_{clear} = 2.0 \pm 0.2\ °C$.

Lindzen[13] holds that GCMs err on this point and that water vapour feedback is in fact negative ($\Delta Q_{clear}/\Delta T < 0$, because tropical cumulus towers enhanced by global warming dry the upper troposphere enough to reduce infrared opacity worldwide). In that case, $\lambda_{clear} > \lambda_b$; perhaps[8] $\lambda_{clear} \approx 8.8\ W\ m^{-2}\ K^{-1}$ ($\Delta T_{2\times} \sim 0.5\ °C$).

In GCMs, the main uncertainty in $\Delta T_{2\times}$ is associated with radiative fluxes, ΔQ_{cloud}, induced by cloud formation and cloud radiative processes when the climate changes. The energy balance $\Delta\bar{Q} + \Delta Q_{clear} + \Delta Q_{cloud} = \lambda_b \Delta\bar{T}$ gives $\lambda \equiv \Delta\bar{Q}/\Delta\bar{T} = \lambda_b - \Delta Q_{clear}/\Delta T - \Delta Q_{cloud}/\Delta T$, or

$$\lambda = \lambda_{clear} - (\Delta Q_{cloud}/\Delta T) \quad (2)$$

Figure 1 shows that λ and $\Delta T_{2\times}$ of GCMs are uncertain by a factor of three because of uncertainties in cloud radiative feedback. Even the sign of $\Delta Q_{cloud}/\Delta T$ is uncertain.

The Intergovernmental Panel on Climate Change (IPCC) in 1990 projected global warming from a transient ocean-climate model[14] for a range of GCM-derived sensitivities[2], $\Delta T_{2\times} = 1.5$, 2.5 and 4.5 °C. Four future radiative forcing scenarios, A, B, C and D, were examined, representing progressively more constrained greenhouse gas emissions. All the scenarios produced some global warming, but unconstrained emissions (Business as Usual, A; "SA90" in IPCC updates) in combination with $\Delta T_{2\times} = 1.5$ °C produced about the same warming as the most highly constrained emissions (D) with $\Delta T_{2\times} = 4.5$ °C. As D would be much more costly to implement than A, $\Delta T_{2\times}$ uncertainties can severely affect the cost projections of emission controls.

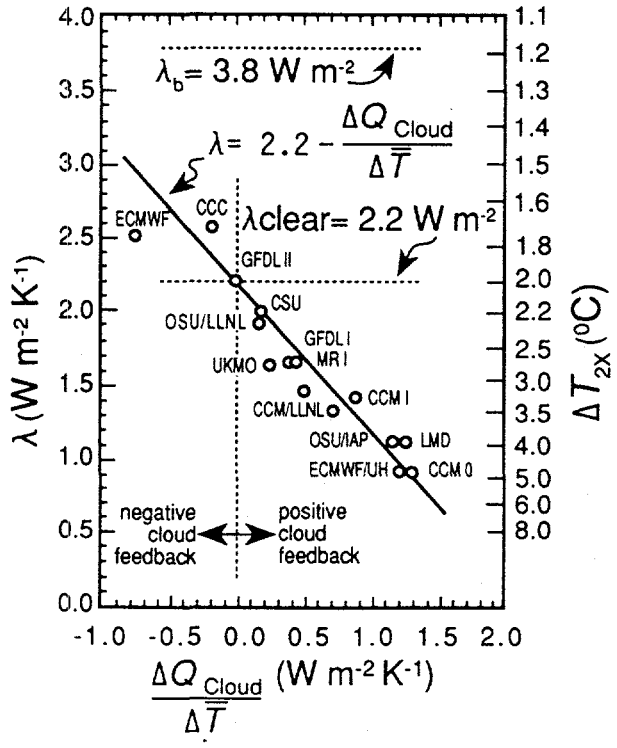


FIG. 1 Radiative damping coefficient, λ, and CO_2 doubling climate sensitivity $\Delta T_{2\times}$ plotted against the cloud radiative feedback parameter of 14 GCMs. Data points are derived from Cess *et al.*[1] where the GCM acronyms are defined (but note the different definition of the symbol λ in ref. 1). The linear correlation (solid line) implies that the differences in GCM-derived climate sensitivity arise mainly from differences in the way that cloud formation and cloud radiation physics are represented in simplified form (parameterized) in different GCMs.

TABLE 1 Sensitivity estimates of two palaeoclimates

Last Glacial Maximum (LGM) cooling ~21.5 kyr BP

Radiative forcing ($W\ m^{-2}$)

	Component mean	Component r.m.s.	Cumulative mean	Cumulative r.m.s.
Component	$\langle \Delta Q_i \rangle$	σ_{Q_i}	$\langle \Delta Q \rangle$	$\sigma_Q = \sqrt{\sum \sigma_{Q_i}^2}$
Sun	0.0	0.2	0.0	0.2
Albedo	−3.0	0.5	−3.0	0.5
Greenhouse	−2.8	0.3	−5.8	0.6
Aerosol	−0.9	0.7	−6.7	0.9
Temperature response (°C)			$\langle \Delta T \rangle$	σ_T
			−3.0	0.6
Climate sensitivity (°C)			$\langle \Delta T_{2\times} \rangle$	$\sigma_{T_{2\times}}$
			2.0	0.5

Middle Cretaceous Maximum (MCM) warming ~100 Myr BP

Radiative forcing ($W\ m^{-2}$)

	Component mean	Component r.m.s.	Cumulative mean	Cumulative r.m.s.
Component	$\langle \Delta Q_i \rangle$	σ_{Q_i}	$\langle \Delta Q \rangle$	$\sigma_Q = \sqrt{\sum \sigma_{Q_i}^2}$
Sun	−1.2	0.2	−1.2	0.2
Albedo	5.8	0.9	4.6	0.9
Greenhouse	11.1	6.7	15.7	6.8
Temperature response (°C)			$\langle \Delta T \rangle$	σ_T
			9.0	2.0
Climate sensitivity (°C)			$\langle \Delta T_{2\times} \rangle$	$\sigma_{T_{2\times}}$
			2.5	1.2

Are more accurate $\Delta T_{2\times}$ estimates possible that are independent of GCMs? One approach is to vary $\Delta T_{2\times}$ in an ocean-climate model[14] to fit the historical temperature record[15]. The radiative forcing history from greenhouse gas buildup (~2 $W\ m^{-2}$ by 1990 relative to pre-industrial) yields a relatively low sensitivity by this technique[7,8], $\Delta T_{2\times} \approx 1.5$ °C. But the calculations omit effects of increasing sulphate aerosols from fossil-fuel burning which could have produced partially compensating negative forcing (~−1 $W\ m^{-2}$ by 1990 from direct scattering of sunlight and/or cloud condensation nuclei increasing marine stratus albedo[16]). Wigley and Raper[17] found the historical temperature record could be fitted when anthropogenic sulphate aerosol forcing was included with $\Delta T_{2\times} \approx 3.4$ °C, whereas Schlesinger *et al.*[18] estimate $\Delta T_{2\times} = 2.2 \pm 0.8$ °C with aerosol forcing. Apart from uncertainties in transient forcing and ocean mixing, natural (unforced) variability during the historical temperature period can affect climate sensitivity derived by this method. This is minimized by working with the long-term average palaeoclimates whose forcings and responses are large.

We sought to derive $\Delta T_{2\times}$ for the Last Glacial Maximum (LGM) about twenty thousand years before present (21.5 kyr BP) and the Mid-Cretaceous Maximum (MGM), about 100 Myr BP, periods of large negative and positive climatic perturbation relative to today[19]. Besides the fast feedbacks included in $\Delta T_{2\times}$ (timescales < 1 yr), 'slow' feedbacks (timescales of decade or more) can come into play involving greenhouse gas concentrations, ocean circulation changes[5] and icecap–bedrock dynamics[20]. Following equilibrium climate models[21,22], we treat fast feedbacks as part of the response, slow feedbacks as direct radiative forcing.

For the LGM, we estimated a mean solar radiative forcing of zero with some residual uncertainty based on 'solar-cycle' variations of Sun-like stars[23]. Surface albedo forcing was estimated from changes in glacial ice, vegetation and topography[3,22]. Greenhouse forcing was from IPCC radiative parameterizations for pre-industrial concentrations[9], $(CO_2)_0 = 279$ p.p.m.v.,

$(CH_4)_0 = 790$ p.p.b.v., and LGM concentrations from the Vostok ice core[3,24], $CO_2 = 195$ p.p.m.v., $CH_4 = 350$ p.p.b.v. Aerosol forcing was based on LGM sulphate particulate loading in ice cores[25] with numerical values from Harvey's[26] LGM base case reduced by a factor of 2 to 10 (ref. 27). For the MCM, a slightly negative solar forcing was estimated based on the Sun brightening by 0.5% per 100 Myr (ref. 28; ±10% uncertainty). Surface albedo forcing was based on calculations of the present Earth relative to the darker (ice-free and more ocean-covered) Earth 100 Myr BP with r.m.s. deviations from uncertainties in surface vegetation and continental boundaries[29]. MCM greenhouse forcing was computed from the parameterization of Kiehl and Dickenson[30] for atmospheric palaeo-CO_2 concentrations 2–11 times present levels estimated from several independent reconstructions[31–35]. Values are given in Table 1. Component uncertainties are treated as r.m.s. deviations σ_{Q_i} about the means $\langle \Delta Q_i \rangle$ of independent random variables ΔQ_i. Also in Table 1 are the cumulative forcing mean and r.m.s. uncertainty computed from[36] $\langle \Delta Q \rangle = \sum \langle \Delta Q_i \rangle$ and $\sigma_Q = (\sum \sigma_{Q_i}^2)^{1/2}$. Our net forcing estimates are -6.7 ± 0.9 W m^{-2} (LGM) and 15.7 ± 6.8 W m^{-2} (MCM).

We estimated the statistical mean and r.m.s. global surface temperature changes, $\langle \Delta T \rangle$ and σ_T, from LGM and MCM palaeotemperature reconstructions as follows. Figure 2*a* shows distributions of zonal mean temperature change reconstructed for the Mid-Holocene (~5–6 kyr BP), Eemian (~125 kyr BP) and Mid-Pliocene (~3.3–4.3 Myr BP) periods normalized to their global mean temperature changes[5]. Budyko and Izrael[6] hypothesized that such normalized surface temperature changes are 'universal'. The least-squares curve fit

$$t(x) = \Delta T(x)/\Delta \bar{T} = 0.36 + 3.21x^4 \qquad (3)$$

where $x = \sin$ (latitude) fits the data of Fig. 2*a* well (correlation coefficient $r^2 = 0.89$), although it is not obvious that it describes temperature distributions of all periods[37]. A counterexample (not shown) may have occurred in the Eocene when sea surface temperatures were apparently cooler in the tropics and warmer at high latitudes[38].

Figure 2*b* shows that Equator-to-pole temperature changes relative to present of both the LGM and MCM are described well by equation (3). The LGM temperature change distribution was synthesized from CLIMAP sea surface temperature (SST) changes equatorward of sea ice[21,22,39] and high-latitude air temperature changes derived from ice cores. The surface air temperature change at Vostok station on the Antarctic plateau is from deuterium isotope fractionation (δD)[24]. Surface cooling relative to present at Dye 3 and Camp Century, Greenland, was derived from oxygen isotope fractionation (δ^{18}O)[40] assuming a change of ~1.0‰ δ^{18}O per degree centigrade[41]. Equator-to-pole temperature distributions in Fig. 2*b* are meant to represent long-term equilibrium climates. The ~9 °C cooling predicted by

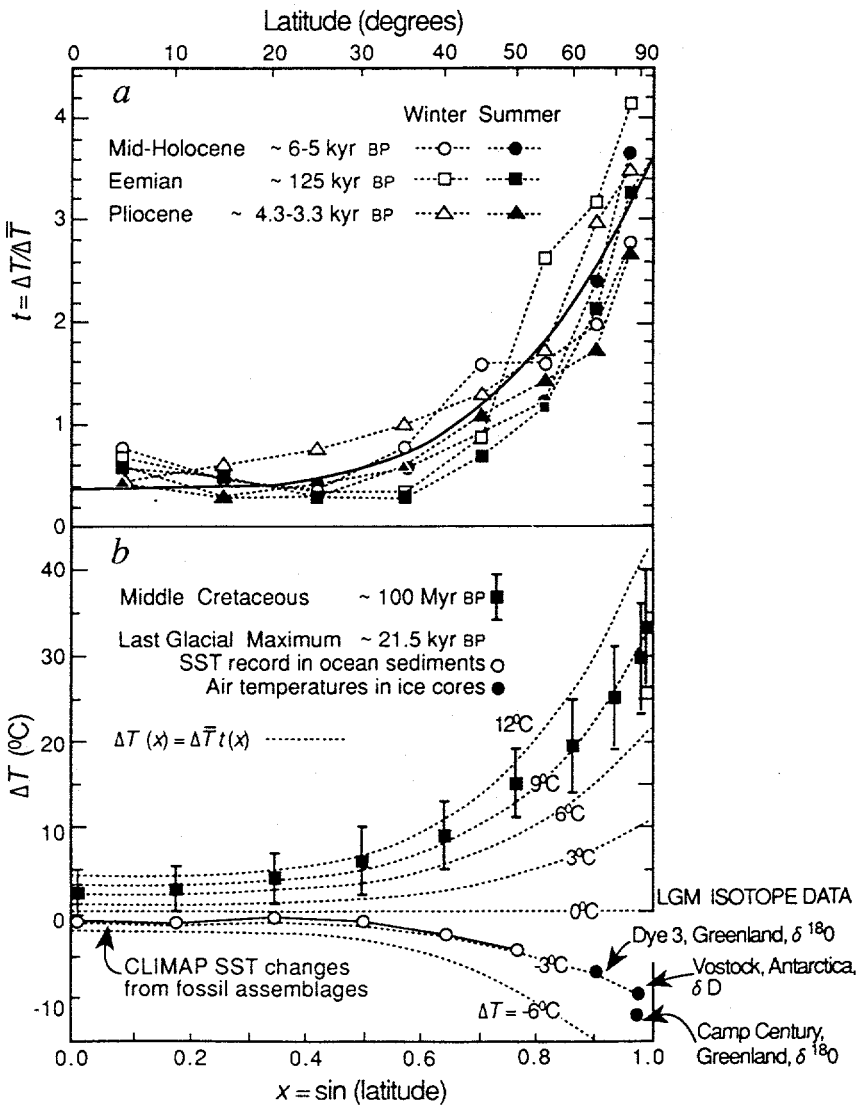


FIG. 2 *a*, Longitudinally averaged surface temperature changes from the present-day climate, normalized by dividing by globally averaged temperature changes. Points connected by dotted lines are based on palaeodata from the mid-Holocene, Eemian and Pliocene warm periods. The solid line is a least-squares fit to this data for a fourth-order polynomial in the sine of latitude (equation (3)). *b*, Longitudinally averaged surface temperature changes from the present-day climate (average of the northern and southern hemispheres). Points are derived from palaeodata for the mid-Cretaceous maximum warming and Last Glacial Maximum cooling (see text). Lines show the best-fit normalized temperature curve of *a* multiplied by the indicated selection of positive and negative global mean temperature changes. Comparison of the lines with a set of points gives the range of global mean temperature change that best fits the data.

the $\Delta\bar{T} = -3$ °C curve at the latitude of Summit, Greenland (72° N) is a reasonable time-average of alternating mild (−7 °C) and severe (−12 °C) LGM glacial climates observed at this station[42]. MCM data in Fig. 2*b* are surface temperature reconstructions of Barron and Washington[43] minus the present surface air temperature distributions. The points shown are an average response of the two hemispheres, with the error bars showing upper and lower bound (not r.m.s.) uncertainties. Comparison of this data with the $\Delta T(x) = \Delta\bar{T}\,t(x)$ curves, also shown in Fig. 2b for different global mean temperature changes, yielded our $\Delta\bar{T}$ estimates, -3.0 ± 0.6 °C (LGM) and 9.0 ± 2.0 °C (MCM).

The climate sensitivity mean and standard deviation in Table 1 were estimated from

$$\langle \Delta T_{2\times} \rangle \approx \frac{\Delta Q_{2\times} \langle \Delta T \rangle}{\langle \Delta Q \rangle}$$

$$\sigma T_{2\times} \approx \langle \Delta T_{2\times} \rangle \sqrt{\left(\frac{\sigma_Q}{\langle \Delta Q \rangle}\right)^2 + \left(\frac{\sigma_T}{\langle \Delta T \rangle}\right)^2} \qquad (4)$$

The approximation assumes statistically independent radiative forcing and temperature estimates and $(\sigma_Q/\langle \Delta Q \rangle)^2$, $(\sigma_T/\langle \Delta T \rangle)^2 \ll 1$. (Small errors associated with nonlinearities and uncertainties in $\Delta Q_{2\times}$ are neglected.) Our $\Delta T_{2\times}$ climate sensitivities are then 2.0 ± 0.5 °C (LGM) and 2.5 ± 1.2 °C (MCM).

Although our method retrieves a Pleistocene $\Delta\bar{T} \approx -3$ °C, Lorius *et al.*[3] cite LGM global cooling of ~−4 to −5 °C. If the polar cooling estimated from ice cores is roughly correct, a global mean cooling of −4 to −5 °C implies that the weak (<1 °C) CLIMAP sea surface cooling from the Equator to 30 degrees latitude is a considerable underestimate (see Fig. 2*b*). Preliminary data from palaeothermometers in groundwater[44] and isotopic and micropalaeontological records in ocean sediment cores[45] admit the possibility of LGM tropical and mid-latitude cooling >3 °C. A recalibrated, colder CLIMAP data set along with consistent land data could increase our LGM sensitivity estimate—perhaps to the mid-point of the IPCC range[2], $\langle \Delta T_{2\times} \rangle \approx 3.0$ °C.

Still, it is remarkable that independently estimated sensitivities of the LGM and MCM are so close to each other and to the $\Delta T_{2\times} \approx 2.2$ °C that Schlesinger *et al.*[18] obtained by fitting the historical temperature record to the output of an ocean-climate model forced by anthropogenic greenhouse gases and aerosols. The 'best guess' $\Delta T_{2\times} \approx 3.4$ °C of Wigley and Raper[17], obtained with slightly different assumptions about the net radiative forcing history, is closer to the 'colder' LGM. In general, $\Delta T_{2\times}$ can change as new palaeoclimate estimates are added to the statistical data base. More focused analysis of palaeodata along with palaeo-reconstructions of additional periods could narrow the uncertainties further.

The numbers are already good enough to rule out a very low sensitivity implied by Lindzen[13]. For example, $\Delta T_{2\times} \approx 0.5$ °C requires relatively large, and nonphysical, forcings (~−26 W m^{-2} to chill the LGM, ~+80 W m^{-2} to heat the MCM). The climate fluctuations imprinted in the geological record could not have occurred if sensitivity was so sluggish. Interpolation of the IPCC's Business-as-Usual scenario curves[2] to our empirical sensitivity estimates for a variety of choices of the ocean model's adjustable parameters, and our own calculations with the Hoffert *et al.* ocean model[14], indicate that global warming by the end of the next century will be 3-4 °C. Such a warming is unprecedented in the past million years, and represents a secular climatic change much faster than previously experienced by natural ecosystems during glacial-interglacial transitions. □

Received 23 June; accepted 27 October 1992.

1. Cess, R. D. *et al. Science* **245,** 513-516 (1989).
2. Houghton, J. T. *et al.* (eds) *Climate Change: The IPCC Scientific Assessment* (Cambridge Univ. Press, 1990).
3. Lorius, C. *et al. Nature* **347,** 139-145 (1990).
4. Hansen, J. E. & Lacis, A. A. *Nature* **346,** 713-719 (1990).
5. MacCracken, M. C. *et al.* (eds) *Prospects for Future Climate: A Special US/USSR Report on Climate and Climate Change* (Lewis, Chelsea, Michigan, 1990).
6. Budyko, M. I. & Izrael, Y. A. *Anthropogenic Climate Change*, 277-318 (Univ. of Arizona Press, Tucson, 1991).
7. Wigley, T. M. L. & Raper, S. C. B. in *Climate Change: Science, Impacts and Policy* (eds Jäger, J. & Ferguson, H. L.) 231-242 (Cambridge Univ. Press, 1992).
8. Schlesinger, M. E. & Jiang, X. *Nature* **350,** 219-221 (1991).
9. Shine, K. P. *et al.* in *Climate Change: The IPCC Scientific Assessment* (eds Houghton, J. T., Jenkins, G. & Ephraums, J. J.) 41-68 (Cambridge Univ. Press, 1990).
10. Mitchell, J. F. B. *Rev. Geophys.* **27,** 115-139 (1989).
11. Manabe, S. & Weatherland, R. T. *J. atmos. Sci.*, **24,** 241-259 (1967).
12. Raval, A. & Ramanathan, V. *Nature* **342,** 758-761 (1989).
13. Lindzen, R. S. *Bull. Am. met. Soc.* **71,** 288-299 (1990).
14. Hoffert, M. I. *et al. J. geophys. Res.* **85** (C11) 6667-6679 (1980).
15. Wigley, T. M. L. & Raper, S. C. B. in *Climate and Sea Level Changes: Observations, Projections and Implications* (eds Warrick, R. A., Barrow, E. M. & Wigley, T. M. L.) (Cambridge Univ. Press, in the press).
16. Charlson, R. J. *et al. Science* **255,** 423-430 (1992).
17. Wigley, T. M. L. & Raper, S. C. B. *Nature* **357,** 293-300 (1992).
18. Schlesinger, M. E., Jiang, X. & Charlson, R. J. in *Climate Change and Energy Policy* (eds Rosen, L. & Glasser, R.) 75-108 (American Institute of Physics, New York, 1992).
19. Crowley, T. J. & North, G. *Paleoclimatology* (Oxford Univ. Press, 1991).
20. Deblonde, G. & Peltier, W. R. *Clim. Dynam.* **5,** 103-110 (1990).
21. Manabe, S. & Broccoli, A. J. *J. atmos. Sci.*, **42,** 2643-2651 (1985).
22. Harvey, L. D. D. *J. geophys. Res.* **94** (D10), 12873-12884 (1989).
23. Baliunas, S. & Jastrow, R. *Nature* **348,** 520-523 (1990).
24. Chappellaz, J. *et al. Nature* **345,** 127-131 (1990).
25. Delmas, R. J. & Legrand, M. in *The Environmental Record in Glaciers and Ice Sheets* (eds Oeschger, H. & Langway, C. C.) 319-341 (Wiley, New York, 1989).
26. Harvey, L. D. D. *Nature* **334,** 333-335 (1988).
27. Anderson, T. L. & Charlson, R. J. *Nature* **345,** 393 (1990).
28. Newman, M. J. & Rood, R. T. *Science* **198,** 1035-1037 (1977).
29. Thompson, S. L. & Barron, E. J. *J. Geol.* **89,** 143-167 (1981).
30. Kiehl, J. T. & Dickinson, R. E. *J. geophys. Res.* **92** (D3), 2991-2998 (1987).
31. Budyko, M. I., Ronov, A. B. & Yanshin, A. L. *History of the Earth's Atmosphere* (Springer, New York, 1987).
32. Berger, W. H. & Spitzy, A. *Paleoceanography* **3,** 401-411 (1988).
33. Popp, B. N. *et al. Am. J. Sci.* **289,** 436-454 (1989).
34. Berner, R. A. *Science* **249,** 1382-1386 (1990).
35. Berner, R. A. *Nature* **358,** 114 (1992).
36. Minkoff, J. *Signals, Noise & Active Sensors* (Wiley, New York, 1992).
37. Crowley, T. J. *J. Clim.* **3,** 1282-1292 (1990).
38. Barron, E. J. *Paleoceanography* **2,** 729-739 (1987).
39. CLIMAP Project Members *Science* **191,** 1131-1137 (1976).
40. Dansgaard, W. *et al.* in *Climate Processes and Climate Sensitivity* (eds Hansen, J. E. & Takahashi, T.) 288-298 (American Geophysical Union, Washington DC, 1984).
41. Bradley, R. S. *Quaternary Paleoclimatology*, 132 (Unwin Hyman, Boston, 1985).
42. Johnsen, S. J. *et al. Nature* **359,** 311-313 (1992).
43. Barron, E. J. & Washington, W. M. in *The Carbon Cycle and Atmospheric CO_2: Natural Variations Archean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 546-553 (American Geophysical Union, Washington DC, 1985).
44. Stute, M. *et al. Science* **256,** 1000-1003 (1992).
45. Emiliani, C. *Science* **257,** 1462 (1992).

ACKNOWLEDGEMENTS. We thank T. Wigley for suggestions, M. Budyko, K. Caldeira, A. Lapenis and D. Rind for discussions, and D. Lischner and J. M. Schantz for programming assistance. This work was supported at NYU by the U.S. Department of Energy (DOE), Atmospheric and Climate Research Division, and the National Institute for Global Environmental Change; and performed in part under the auspices of the DOE Environmental Sciences Division by the Lawrence Livermore National Laboratory.

GEOPHYSICAL RESEARCH LETTERS, VOL. 20, NO. 22, PAGES 2455-2458, NOVEMBER 19, 1993

SEASONAL, LATITUDINAL, AND SECULAR VARIATIONS IN TEMPERATURE TREND: EVIDENCE FOR INFLUENCE OF ANTHROPOGENIC SULFATE

David E. Hunter[1], Stephen E. Schwartz, Richard Wagener, and Carmen M. Benkovitz

Environmental Chemistry Division, Brookhaven National Laboratory

Abstract. Tropospheric aerosols increase the shortwave reflectivity of the Earth-atmosphere system both by scattering light directly, in the absence of clouds, and by enhancing cloud reflectivity. The radiative forcing of climate exerted by anthropogenic sulfate aerosols, derived mainly from SO_2 emitted from fossil fuel combustion, is opposite that due to anthropogenic greenhouse gases and is estimated to be of comparable average magnitude in Northern Hemisphere midlatitudes. However, persuasive evidence of climate response to this forcing has thus far been lacking. Here we examine patterns of seasonal and latitudinal variations in temperature anomaly trend for evidence of such a response. Pronounced minima in the rate of temperature increase in summer months in Northern Hemisphere midlatitudes are consistent with the latitudinal distribution of anthropogenic sulfate and changes in the rate of SO_2 emissions over the industrial era.

Introduction

Tropospheric aerosol sulfate is derived mainly from biogenic sulfur-containing gases (mostly dimethylsulfide) and SO_2 emitted principally from fossil fuel combustion. SO_2 emissions have increased substantially over the past 150 years and now considerably exceed estimated natural emissions of gaseous precursors of sulfate aerosol (Figure 1). The resultant increase in sulfate aerosol concentration is thought to have exerted a change in Earth's shortwave radiation budget that is substantial in the context of longwave forcing by the increase in greenhouse gases over the same period [Charlson et al., 1992; Wigley and Raper, 1992].

More than 90% of anthropogenic SO_2 introduced into the atmosphere is emitted in the Northern Hemisphere (NH), and approximately 70% between 30°N and 60°N [Spiro et al., 1992; Dignon, 1992]. Model calculations indicate that both the resulting sulfate aerosol [Charlson et al., 1991; Langner and Rodhe, 1991] and radiative forcing [Charlson et al., 1991; Kiehl and Briegleb, 1993] are confined primarily to 30-60°N, in agreement with observed sulfate aerosol concentrations [Schwartz, 1988] and satellite observations of clear-sky [Durkee et al., 1991] and cloud [Han et al., 1993] albedo. However there is scant evidence of the influence of this radiative forcing in interhemispheric differences of annual temperature anomaly trend over the industrial era [Schwartz, 1988; Wigley, 1989]. On the other hand, comparison of trends of mean daily maximum (mostly daytime) and minimum (mostly nighttime) temperatures in industrial regions [Karl et al., 1991] indicates a greater rate of warming at night than during daytime, perhaps indicative of a shortwave forcing by anthropogenic sulfate aerosol in conjunction with overall global warming over this period.

Because of greater insolation in summer than winter, the shortwave radiative influence of anthropogenic sulfate aerosol is expected to be greater in summer, and this might lead to a lesser rate of warming in NH midlatitudes (30° to 60°) in summer than in winter that would serve as a sensitive indicator of climate response to sulfate aerosol forcing. To examine the seasonal dependence of the forcing we carried out a radiative transfer model calculation for a spatially and temporally uniform incremental aerosol optical depth of 0.04, corresponding to the difference between a background aerosol optical depth of 0.05 [Forgan, 1987] and that characteristic of industrial latitudes, 0.09. The forcing was computed for a horizontally homogeneous, cloud-free Earth at 0.55 μm using the doubling and adding code of Hansen and Travis [1974]. For this calculation the model atmosphere was divided into three layers: top layer, gaseous absorption and Rayleigh

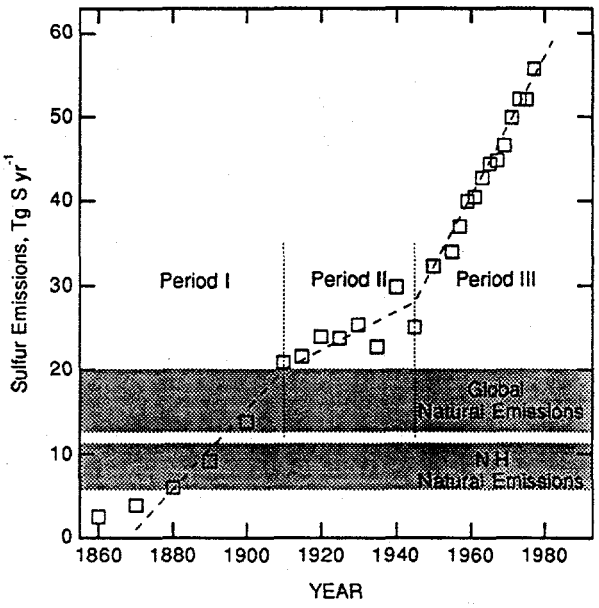


Fig. 1. Time series of estimated annual anthropogenic SO_2 emissions in midlatitudes (30-60°) of the Northern Hemisphere. Midlatitude NH emissions are calculated from estimated global emissions [Möller, 1984] as 98% of global emissions until 1910, 90% in 1945, and 70% in 1986 [Dignon and Hameed, 1989; Spiro et al., 1992] with linear interpolation in periods II and III. Other estimates, e.g., Robinson and Robbins [1971], are similar. Dashed lines represent rate of increase of 0.50, 0.23, and 0.84 Tg S yr^{-1} for periods I, II, and III, respectively, although such a rate of increase for period III was probably not sustained in the 1980s [Spiro et al., 1992]. Also shown are estimates of total global and NH natural emissions of gaseous reduced sulfur compounds and sulfur dioxide [Bates et al., 1992]; the widths of the bands represent ranges of estimates.

[1]Present Address: Scripps Institution of Oceanography, La Jolla CA 92093.

Copyright 1993 by the American Geophysical Union.

Paper number 93GL02808
0094-8534/93/93GL-02808$03.00

scattering (τ_{abs} = 0.03, τ_{Ray} = 0.1); aerosol layer, having log-normal particle size distribution with geometric mean radius 0.0285 μm, geometric standard deviation 2.239, and refractive index 1.43; Lambertian surface layer with albedo 0.15. The

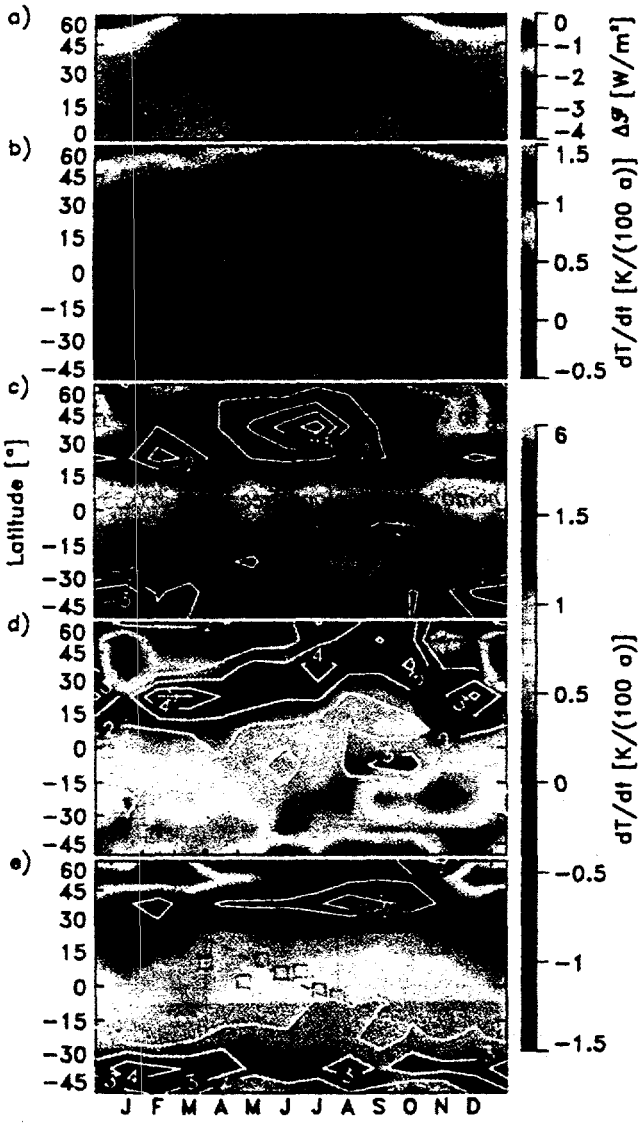


Fig. 2. Latitudinal and seasonal dependence of aerosol forcing and temperature anomaly trend. (a) Diurnal-average shortwave radiative forcing due to a uniform incremental aerosol optical depth of 0.04 (i.e., optical depth 0.09 versus optical depth 0.05) as a function of month of year and latitude for the NH only. (b) Dependence of temperature anomaly trend on month of year and latitude for the period 1854-1989 evaluated from the combined land and sea-surface temperature anomaly data set of Jones et al. [1986a] as obtained from NCAR [1990]; the figure presents a bilinear interpolation of the slopes calculated for 15° zonal area-weighted means of the original 5° by 5° gridded monthly-mean temperature anomaly data set. (c, d, e) Same as (b) but for the periods 1854-1910, 1910-1945, and 1945-1989, respectively. Note compressed color scale at high NH latitudes in winter. Labeled contours represent ratio of the difference between the temperature trend during the indicated period and that for the overall period 1854-1989 to the estimated resultant 1-σ standard error of that difference. Latitude is displayed on a sine scale in all panels.

calculated forcing in midlatitudes is considerably greater (negative) in summer than winter (Figure 2a). Kiehl and Briegleb [1993] similarly note a greater forcing in summer than winter in industrial regions, despite similar modeled sulfate column burdens, which they likewise attribute to greater summertime insolation.

The summer-winter difference in forcing would be enhanced by higher summertime sulfate concentrations that might result from a greater summertime rate of conversion of SO_2 to sulfate [NAPAP, 1990]. Higher surface concentrations of aerosol sulfate in summer than winter (by factors of 2 to 5) are noted in eastern North America [Husain and Dutkiewicz, 1990], although not in Europe [Schaug et al., 1989]. The summer-winter difference in forcing might be enhanced also by a greater rate of new particle production in summertime and resultant increased cloud reflectivity, as might be expected from the dominance of gas-phase oxidation of SO_2 by OH in summer, versus dominance of aqueous-phase oxidation in winter [Langner et al., 1992].

It has been noted previously that the summertime rate of warming in the NH over the industrial period as a whole (1854-1986) has been considerably less than the annual mean [Gordon, 1992], with the lesser rate of summertime warming confined mainly to midlatitudes [Schönwiese and Stähler, 1991]. In contrast the SH exhibits little seasonal variation in warming rate over this period. These observations may be indicative of an increase in seasonal variation in forcing resulting from an increase in anthropogenic sulfate aerosol in NH midlatitudes. Wigley et al. [1992] have noted a summertime cooling trend subsequent to 1950 in regions influenced by anthropogenic aerosols relative to other regions. Recently Engardt and Rodhe [1993] have examined temperature trends between the periods 1926-45 and 1971-90 for locations which exhibit high sulfate column burdens in the Langner et al. model calculations, finding indication of summertime cooling in regions of high burden, albeit not statistically significant.

Results and Interpretation

Here we examine the seasonal and latitudinal dependence of temperature anomaly trend (hereinafter temperature trend) for the period 1854 to 1989, evaluated for 15° latitude zones from the monthly combined land and sea-surface temperature anomaly data set of the University of East Anglia Climatic Research Unit [Jones et al., 1986 a-c] as obtained from the National Center for Atmospheric Research [NCAR, 1990]. Temperature trends were evaluated as the linear regression slope of area-weighted zonal means of the monthly-mean temperature anomalies versus time, with explicit treatment of temporal autocorrelation by the Prais-Winsten generalized least-squares approach [Johnston, 1984]. The resulting temperature trends are shown as a function of latitude and season in Figure 2b. Latitudes above 75°N and below 60°S are excluded because of the short temporal record and restricted spatial coverage. Because the spatial coverage is sparse before 1880, especially for the SH, the slopes for the SH must be viewed as fairly uncertain [Jones et al., 1986 b,c]. Despite these uncertainties a strong seasonal variation in warming rate is exhibited in mid (30-60°) and high (60-75°) latitudes of the NH over the time period as a whole that contrasts markedly with the much more uniform pattern of warming exhibited in the SH. The six monthly-mean warming rates for the 30-45° and 45-60° zones of the NH in June, July, and August are the lowest globally and are in fact near zero (range 0.12 ± 0.09 K/century to -0.07 ± 0.10 K/century, i.e.,

slight cooling; the uncertainties are the 1-σ standard errors). The coincidence of this minimum warming rate with the latitude of maximum anthropogenic SO_2 emissions and sulfate concentrations and with the calculated seasonal maximum in shortwave aerosol forcing is suggestive of a seasonal component of the climate response to shortwave radiative forcing by anthropogenic sulfate aerosol.

Insight into possible alternative explanations of the seasonal and latitudinal variation in NH temperature trend may be gained from a study with a coupled ocean-atmosphere model [Manabe et al., 1992], which examined temperature trend as a function of latitude and season in response to a hemispherically symmetric forcing, corresponding to a 1% yr^{-1} increase in CO_2, roughly equal to the current rate of increase in forcing by greenhouse gases. Except for NH midlatitude summer the model results agree closely with the observed trends at all latitudes and seasons, specifically including the marked enhancement of the wintertime warming rate above 60°N. The enhanced high-latitude wintertime warming is attributed to latent heat trapped during sea ice melting in summer and released in winter through a thus thinner ice layer [Manabe et al., 1992]. However little seasonal variation was exhibited in the modeled temperature trends in the latitude range 30°N to 60°N, and in particular no lesser rate of summertime warming at these latitudes, in contrast to the observations. The observed warming rate of the 60-75°N band, while greatly elevated in wintertime, does not drop in summer more than 0.2 K/century below the hemispheric mean of 0.4 K/century, whereas that for the midlatitude bands is not greatly elevated in wintertime but is well below the hemispheric mean in summertime. The seasonal variation in the 60-75°N band thus seems largely unrelated to the observed variation in the 30-45°N and 45-60°N bands.

To further examine for evidence of influence of sulfate aerosol on temperature trends, we note that the rate of increase in anthropogenic SO_2 emissions (Figure 1) has not been constant throughout the industrial era. Prior to 1910 (period I) and between 1945 and 1989 (period III) emissions increased rapidly, initially in Europe and North America, and later in Asia. In contrast, between 1910 and 1945 (period II) the rate of increase was much lower, primarily because of large-scale conversion from coal to oil. Since sulfate aerosol is short lived in the atmosphere, a change in forcing is brought about only by a change in emissions rate. Therefore the effect of SO_2 emissions on temperature trend would be expected to be greatest in periods I and III, but minimal in period II. Consequently, we examined the latitudinal and seasonal dependence of temperature trend separately for the three time periods (Figure 2c-e). Periods I and III exhibit temperature trends in NH midlatitudes that have seasonal and latitudinal distributions similar to that of the decreased warming rate exhibited by the period as a whole and in fact show significant cooling in summer months. The labeled contours in these figures indicate the statistical quality of the difference in temperature trend between the individual time periods and the 1854-1989 period as a whole. For periods I and III the temperature trend in NH midlatitude summer is less than that for the time period as a whole by more than three times the standard error of the difference. In contrast, during Period II the NH midlatitudes warmed at a rate comparable to that in other seasons and much greater than that for the time period as a whole, again by several times the standard error of the difference. Moreover this warming pattern agrees closely with the model results of Manabe et al. [1992]. To be sure, there is not a one-to-one correspondence between temperature trend and rate of increase in sulfur emissions; this could be due to nonlinear response to aerosol forcing or to differences in geographical distribution of emissions, and/or to influences unrelated to aerosol forcing. Nonetheless, the markedly different NH midlatitude summertime temperature trend associated with the differing rates of increase in SO_2 emissions characterizing the three time periods supports attribution of the lower warming rate in NH midlatitudes in summer to the influence of anthropogenic sulfate.

Conclusions

The seasonal variation of shortwave radiative forcing by aerosols indicated in Figure 2a suggests that examination of the spatial and temporal patterns of temperature trend should provide a powerful means of testing climate response to increased sulfate aerosol over the industrial era. The seasonal dependence of observed temperature trend in NH midlatitudes over the industrial period as a whole and in the two periods of greatly increasing SO_2 emission rates is qualitatively consistent with the historical pattern of this forcing. The seasonal variation in forcing and observed temperature trend thus suggests that response to this variation in forcing be explicitly examined in climate model studies.

Three final points should be noted. First, if the minimum in warming rate noted for midlatitude summers is due to shortwave radiative forcing by anthropogenic sulfate aerosol, then the observed seasonal and latitudinal differences in warming rate must be viewed as a substantial underestimate of the magnitude of the contribution of this forcing to global cooling, as the forcing will be active in all seasons (Figure 2a) and because a forcing in a given season and location may be expected to influence temperature in other seasons and locations. Second, to the extent longwave forcing by anthropogenic greenhouse gases has been offset by shortwave forcing by sulfate aerosols, the warming that would have occurred from the greenhouse forcing alone may be substantially greater than that which has actually occurred thus far over the industrial era. Sulfate forcing may thus account for the apparently too great response to greenhouse forcing, relative to observed trends, of climate models that do not include this forcing [Houghton et al., 1992; Wigley and Raper, 1992]. Third, improvements in description of the spatial and temporal pattern of the sulfate aerosol forcing, in conjunction with observed spatial and temporal patterns of temperature trend, may lead to improved empirical estimates of the sensitivity of temperature to perturbations in radiative forcing.

Acknowledgments. D.E.H. gratefully acknowledges support as a Semester Program participant from the U.S. Department of Energy's Division of University and Industry Programs, Office of Energy Research. We thank S. Nemesure for assistance with the calculations. This research was performed under the auspices of the United States Department of Energy, under Contract No. DE-AC02-76CH00016.

References

Bates, T. S., B. K. Lamb, A. Guenther, J. Dignon, and R. E. Stoiber, Sulfur emissions to the atmosphere from natural sources, *J. Atmos. Chem.*, 14, 315-337, 1992.

Charlson, R. J., J. Langner, H. Rodhe, C. B. Leovy, and S. G. Warren, Perturbation of the Northern Hemisphere

radiative balance by backscattering from anthropogenic aerosols, *Tellus*, 43AB, 152-163, 1991.

Charlson, R. J., S. E. Schwartz, J. M. Hales, R. D. Cess, J. A. Coakley, Jr., J. E. Hansen, and D. J. Hofmann, Climate forcing by anthropogenic aerosols, *Science*, 255, 423-430, 1992.

Dignon, J., NO_x and SO_x emissions from fossil fuels: A global distribution, *Atmos. Environ.*, 26A, 1157-1163, 1992.

Dignon, J. and S. Hameed, Global emissions of nitrogen and sulfur oxides from 1860 to 1980, *J. Air Pollut. Contr. Assoc.*, 39, 180-186, 1989.

Durkee, P. A., F. Pfeil, E. Frost, and R. Shema, Global analysis of aerosol particle characteristics, *Atmos. Environ.*, 25A, 2457-2471, 1991.

Engardt, M. and H. Rodhe, A comparison between patterns of temperature trends and sulfate aerosol pollution, *Geophys. Res. Lett.*, 20, 117-120, 1993.

Forgan, B. W., Aerosol optical depth, in *Baseline Atmospheric Program (Australia) 1985*, edited by B. W. Forgan and P. J. Fraser, p. 56, Department of Science/Bureau of Meteorology and CSIRO/Division of Atmospheric Research, Australia, 1987.

Gordon, A. H., Interhemispheric contrasts of mean global temperature anomalies, *Internatl. J. Climatology*, 12, 1-9, 1992.

Han, Q., W. B. Rossow, and A. A. Lacis, Near-global survey of effective droplet radii in liquid water clouds using ISCCP data, *J. Climate*, in press, 1993.

Hansen, J. E. and L. D. Travis, Light scattering in planetary atmospheres, *Space Sci. Rev.*, 16, 527-610, 1974.

Houghton, J. T., B. A. Callander, and S. K. Varney, *Climate Change 1992: The Supplementary Report to the IPCC Scientific Assessment*, Cambridge Univ. Press, 1992.

Husain, L. and V. A. Dutkiewicz, A long-term (1975-1988) study of atmospheric SO_4^{2-}: Regional contributions and concentration trends, *Atmos. Environ.*, 24A, 1175-1187, 1990.

Johnston, J., *Econometric Methods*, McGraw Hill, New York, 1984.

Jones, P. D., T. M. L. Wigley, and P. B. Wright, Global temperature variations between 1861 and 1984, *Nature*, 322, 430-434, 1986a.

Jones, P. D., S. C. B. Raper, R. S. Bradley, H. F. Diaz, P. M. Kelly, and T. M. L. Wigley, Northern Hemisphere air temperature variations 1851-1984, *J. Clim. Appl. Meterol.*, 25, 161-179, 1986b.

Jones, P. D., S. C. B. Raper, and T. M. L. Wigley, Southern Hemisphere surface air temperature variations: 1851-1984, *J. Clim. Appl. Meterol.*, 25, 1213-1230, 1986c.

Karl, T. R., G. Kukla, V. N. Razuvayev, M. J. Changery, R. G. Quayle, R. R. Heim, D. R. Easterling, and C. B. Fu, Global Warming: Evidence for asymmetric diurnal temperature change, *Geophys. Res. Lett.*, 18, 2253-2256, 1991.

Kiehl, J. T. and B. P. Briegleb, The relative roles of sulfate aerosols and greenhouse gases in climate forcing, *Science*, 260, 311-314, 1993.

Langner, J. and H. Rodhe, A global three-dimensional model of the tropospheric sulfur cycle, *J. Atmos. Chem.*, 13, 225-263, 1991.

Langner, J., H. Rodhe, P. J. Crutzen, and P. Zimmermann, Anthropogenic influence on the distribution of tropospheric sulphate aerosol, *Nature*, 359, 713-716, 1992.

Manabe, S., M. J. Spelman, and R. J. Stouffer, Transient responses of a coupled ocean-atmosphere model to gradual changes in atmospheric CO_2. Part II: Seasonal response, *J. Climate*, 5, 105-126, 1992.

Möller, D., Estimation of the global man-made sulphur emission, *Atmos. Environ.*, 18, 19-27, 1984.

NAPAP, National Acid Precipitation Assessment Program (U.S.), Acidic deposition: State of science and technology report, Vol. 1, Report 2, Atmospheric processes and deposition, U.S. Gov't Printing Office, Washington, DC, 1990.

NCAR, National Center for Atmospheric Research, Boulder, CO; file date December, 1990.

Robinson, E. and R. C. Robbins, Sources, abundance and fate of gaseous atmospheric pollutants--Supplement, *Rep. 4015*, American Petroleum Institute, Washington, DC, 1971.

Schaug, J., J. E. Skjelmoen, S. E. Walker, U. Pedersen, and A. Harstad, Data report 1987. Part 1: Annual summaries, *Rep. EMEP-CCC 1/89*, Norwegian Institute for Air Research (NILU), Lillestrom, Norway, 1989.

Schönwiese, C. D. and U. Stähler, Multiforced statistical assessments of greenhouse-gas-induced surface air temperature change, *Climate Dynamics*, 6, 23-33, 1991.

Schwartz, S. E., Are global cloud albedo and climate controlled by marine phytoplankton?, *Nature*, 336, 441-445, 1988.

Spiro, P. A., D. J. Jacob, and J. A. Logan, Global inventory of sulfur emissions with $1° \times 1°$ resolution, *J. Geophys. Res.*, 97, 6023-6036, 1992.

Wigley, T. M. L., Possible climate change due to SO_2-derived cloud condensation nuclei, *Nature*, 339, 365-367, 1989.

Wigley, T. M. L., P. D. Jones, P. M. Kelly, and M. Hulme, Recent global temperature changes: ozone and aerosol influences, *Proc. Sixteenth Annual Climate Diagnostics Workshop*, edited by L. Mannello, 194-202, NOAA, U.S. Dept. of Commerce, 1992.

Wigley, T. M. L. and S. C. B. Raper, Implications for climate and sea level of revised IPCC emissions scenarios, *Nature*, 357, 293-300, 1992.

Carmen M. Benkovitz, David E. Hunter, Stephen E. Schwartz, and Richard Wagener, Environmental Chemistry Division, Brookhaven National Laboratory, P. O. Box 5000, Upton, NY 11973-5000.

Received: September 3, 1993
Accepted: October 6, 1993

GLOBAL CLIMATE CHANGES AND MOISTURE CONDITIONS IN THE INTRACONTINENTAL ARID ZONES

A. G. LAPENIS

Department of Climatic Change, State Hydrological Institute, St. Petersburg, Russia. Now at Department of Earth System Science, New York University, New York, N.Y. U.S.A.

and

M. V. SHABALOVA

Department of Climatic Change, State Hydrological Institute, St. Petersburg, Russia. Now at Royal Netherlands Meteorological Institute, De Bilt, Netherlands

Abstract. In order to estimate a transient response of the local hydrological cycle and vegetation cover in the African monsoon area to global climate changes, a simple two-dimensional water vapor transport model coupled with a carbon cycle model for the soil was used. The key difference from other models is that we take into account a positive feedback between the precipitation and development of the vegetation root system in the underlying surface. As our calculation shows, this feedback is responsible for a long-term transient response of local hydrological cycles to the global temperature changes. In the case of a four component vegetation system – tropical forests, savannah, semi desert and desert, (and 2 °C ocean surface water warming), a new steady-state is reached in about 1500 years.

In previous works of other authors, the increase of summer precipitations during Holocene or Last Interglacial could be explained only as a result of the surface temperature increase in the intracontinental parts of Africa. However, from paleodata indicates, the temperature in the intracontinental regions of Africa rather decreased during warm epochs of geological past: Holocene optimum, Last Interglacial and middle Pliocene climatic optimum. Our simple model simulations agree with both paleoprecipitation and paleotemperature data.

1. Introduction

During the Holocene optimum (5–6 Kyears (thousands of years) ago) and Eemian time (Last Interglacial, about 125 Kyears ago), the water content of the air masses transported from over the ocean to continents, exceeded that of the present period (due to an increase in sea surface temperature (Barash *et al.*, 1980; CLIMAP, 1984)). The total water vapor flux at the ocean-land boundary should also be greater to provide increased precipitation in the coastal regions of Africa and Australia as known from paleodata (Murzayeva, 1984; Borzenkova, 1992; Borzenkova *et al.*, 1992).

Analyses of African lake levels (Williams and Faure, 1980) and the discovery of fresh-water diatom algae transported by wind from these lakes, in the deep-sea sediments of the Last Interglacial (Sarthein *et al.*, 1982), indicate pronouncedly better moisture conditions and decreases of Sahara desert area in the intracontinental areas of Africa during the Last Interglacial and Holocene optimum.

Climatic Change **27**: 283–297, 1994.
© 1994 *Kluwer Academic Publishers. Printed in the Netherlands.*

Recently, the same features (increasing the precipitations in the intracontinental areas and decreasing the surface temperatures in the desert and semidesert areas) were found for the warm climate of middle Pliocene (Pliocene climatic optimum, about 3.3 Myears ago) (Borzenkova *et al.*, 1992; Poore *et al.*, 1993).

Increased advective transport in the coastal regions and increased precipitation in the interior of Northern Africa and elsewhere in the Northern Hemisphere during Holocene optimum and Eemian time are in good agreement with results of African monsoon simulations obtained by Kutzbach *et al.* with the use of steady-state low resolution general circulation model (Kutzbach and Otto-Bliesner, 1982; Kutzbach and Guetter, 1984; Kutzbach and Street-Perrott, 1985; Prell and Kutzbach, 1987).

In Kutzbach model experiments, the forcing is increased summer insolation by changes in Milankovich cycles. Due to higher summer insolation and greater thermal inertia of ocean masses compared to continents, summer temperatures in the middle parts of continents in Kutzbach's model increased more than surface ocean temperature. This increased ocean-continent temperature gradient intensifies monsoon transport in Kutzbach's model. For Australia, Kutzbach's model fails and did not show increased precipitation. As mentioned above, in the intracontinental parts of Northern Africa Kutzbach's model shows increased surface temperature. And this increase of the temperature in the intracontinental part of Africa is very important for Kutzbach's mechanism. However, as paleodata evidence indicate, the temperature in the intracontinental parts of Africa (as well as Australia) did not increase (and might even decrease) during the Holocene optimum and Eemian time (Borzenkova, 1987; Borzenkova, 1992; Gasse, 1980; Glenn *et al.*, 1986). Even if we would suggest that the separation of the precipitation and temperature signal in the above mentioned works was made incorrectly, there is still not any clear evidence of significant increase of summer temperature in the middle part of Africa during the Holocene optimum. This circumstance allowed us to have some freedom in the interpretation of the paleoclimate data and propose a new hypothesis which will not contradict with the data.

We should note also that it would be difficult to apply Kutzbach's model results to the Middle Pliocene climatic optimum where forcing was different from Holocene or Eemian. The main reason for warmer climate at Pliocene climatic optimum was increased atmospheric CO_2 (Crowley, 1991). However, the precipitation in the intracontinental areas during Pliocene warming increased, as well as during Holocene and Eemian time (Borzenkova *et al.*, 1992; Poore *et al.*, 1993).

In this paper we hypothesize that the primary mechanism responsible for the decreased temperature in the intracontinental regions mentioned above, is the intensification of the local moisture circulation, shift of vegetation and (with some time lag) soil zones and reduction of the deserts. Moreover we suggest that this mechanism, which provides the possibility over long-time scales of stimulating the intensification of the local moisture circulation in midcontinental arid zones, is a common feature of any warm climate.

2. Statement of the Problem

In this paper, to simulate precipitation changes in the midcontinental regions of Africa as a result of warmer ocean, we put forward a hypothesis of a long-term positive feedback between soil parameters and local precipitation. For calculations we used 2-D transient water transport model coupled with simple soil carbon cycle model. The forcing in our model is changes in ocean surface temperature (+2 °C) (summer insolation as well as atmospheric CO_2 and velocity of water vapor advection from over the ocean to continents are constant and equal to their present values).

3. Description of the Supposed Mechanism of Feedback

It is supposed that this mechanism is a positive feedback between precipitation and soil parameters of the underlying surface, which affect soil moisture parameters of the underlying surface, which affect soil moisture content on a long (thousands of years) time scale. According to our understanding, the cause-and-effect relationship in this mechanism is as follows:

(i) The initial precipitation increase due to increasing above the ocean the water vapor content (as a result of 2 °C warming of the ocean surface) should provide the possibility for vegetation zones to shift because in the tropics their positions are strongly determined by annual amount of precipitation (Bukyko, 1974; Kobak, 1988).

(ii) The shift of vegetation zones is accompanied with the forming (in the marginal regions) of new soils (that takes about 1500 years, as will be shown below). According to Sellers (1965), Strahler (1969), and others, every stable soil type can be associated with a certain field capacity value (available water at saturation expressed in cm of water column, which any given type of soil can hold in its 1 m surface layer). Zubenok (1975) showed that the relationship between field capacity and types of soils depends mainly on the development of the vegetation root system.

(iii) Higher field capacity values result in increased absolute values of moisture in the soil and, consequently, in intensified evapotranspiration because in arid regions it is limited by absolute values of soil water content (Zubenok, 1975; Hansen *et al.*, 1983; Kellog and Zhao, 1988). This, in turn, would increase water vapor advection (even with the constant advection velocity) to the adjacent more arid regions, forming new marginal areas, i.e. precipitation (and therefore vegetation) would move to the intracontinental regions.

It should be mentioned that if the vegetation zones change, there is the change of surface albedo that potentially affects climate as well as soil type. As it was shown by Charney (1975), albedo feedback in arid areas can only amplify the proposed mech-

anism, therefore, in this paper we investigate soil-precipitation feedback regardless of albedo change.

In case of cooling, the proposed mechanism should work in the opposite direction. Lower sea surface temperature should decrease water vapor coming to continents from over the sea air masses and decrease the precipitation (if the advective transport were the same or lower). In some time, after the mineralization of old vegetation root systems, local precipitation in these areas should decrease and the process should develop until the new steady state condition.

The impact of soil moisture and evapotranspiration on the Earth's climate has been studied by various modelers (Shukla and Mintz, 1982; Hansen *et al.*, 1983; Rind, 1984; and others). Since the precipitation and soil moisture are secondary features of the general circulation, and depend on a complex sequence of processes, the results of GCM simulations of the hydrological cycle are hardly interpretable and partly contradicting (The IPCC Scientific Assessment, 1990). We suppose that a simplified hydrological model can demonstrate large-scale interactions between hydrological variables more clearly than GCMs.

The key point of this work is that in our study we take into account a slowly changing field capacity due to the formation of new soil types (which takes about 1500 years) in regions where precipitation increases.

4. Model Used

To obtain quantitative estimates of the feedback between precipitation and soil parameters we have used the simple two-dimensional model of water exchange between ocean and land developed by Byutner and Shabalova (1989, 1990). The model calculates the vertical distributions of temperature and specific humidity due to the air masses transported from over the ocean to the land.

4.1. *Model Equations*

Our 2-D model is a parameterized climate model where the horizontal coordinate corresponds to the distance from the coast. The governing equations are the conservation of energy and water vapor. These are:

$$u\frac{\partial \vartheta}{\partial x} - \frac{\partial}{\partial z}k_{\mathrm{T}}(z)\frac{\partial \vartheta}{\partial z} - \nu\frac{L}{c_{\mathrm{p}}} = 0, \tag{1}$$

$$u\frac{\partial q}{\partial x} - \frac{\partial}{\partial z}k_{\mathrm{q}}(z)\frac{\partial q}{\partial z} + \nu = 0, \tag{2}$$

where z is the vertical coordinate, ϑ the potential temperature, q the specific humidity, c_{p} the specific heat at constant pressure, L the latent heat of condensation, ν the condensation heating rate, u the rate of advection in the predominant direction, p the pressure, k_{T} and k_{q} the eddy thermal and water vapor diffusion coefficients.

Following most papers (see Eagleson, 1982), we set $k_T = k_q = k$ in the model, where k is a function of z.

Equations (1) and (2) describe the transformation of the profiles of potential temperature and specific humidity when air masses move from over the ocean to the land. We also used the equation of conservation of water vapor mass above the ocean surface, in order to estimate the value of the temperature at the top of the troposphere (9 km). The boundary conditions at the top of the troposphere are that the specific humidity is equal to zero and $\vartheta(H) = \vartheta(H)\,|_{x=0}$. The last value ($\vartheta(H)$ above the ocean) is calculated from (1) – (2) and the equation of water vapor conservation, with prescribed relative humidity (80%) and sea surface temperature under the condition that precipitation is equal to evaporation.

4.2. *Parameterization of Precipitation*

To calculate the precipitation, the convective adjustment routine is used. The procedure of adjustment, starting at the bottom of the model and progressing toward the top of the troposphere, examines a vertical column of grid points. For a saturated layer, the excess water vapor is rained out as precipitation and, at the same time, latent heat is added to the layer. The process is controlled by the constraint that

$$\delta\left(\vartheta + \frac{L}{c_p}q\right) = 0. \tag{3}$$

The total precipitation is then the sum of all individual layer precipitation.

4.3. *Parameterization of the Land-Atmosphere Interactions*

The heat balance equation in the form:

$$Q(1-\alpha) - I - LE - P - G = 0 \tag{4}$$

serves as a boundary condition for temperature at the bottom of the model. Here α is the surface albedo, Q is the downward solar radiative flux, I is the effective terrestrial radiative flux, P is the eddy heat flux, LE is the latent heat flux and G is the flux of the sensible heat in the soil.

The effective terrestrial radiative flux parameterization in terms of surface air temperature, ϑ, and surface water vapor pressure, e, comes from papers of Budyko (1974):

$$I = s\sigma\vartheta^4(1 - 0.7n)(0.254 - 0.0053e) + 4s\sigma\vartheta^3(\vartheta_s - \vartheta), \tag{5}$$

where s is the surface long-wave emissivity, σ is the Stefan-Boltzmann constant, n is the cloud cover, and ϑ_s is the ground surface temperature.

For the heat flux in the soil, a simple parameterization is generally used, G is expressed as a simple function of net downward radiative flux R at the surface:

$$G = 0.1R = 0.1(Q(1-\alpha) - I). \tag{6}$$

Sensible and latent heat fluxes P and LE are expressed as

$$P = \rho c_p C_H u(\vartheta_s - \vartheta), \tag{7}$$

$$LE = LC_E \rho u(q_s - q)\frac{w}{w_0}, \tag{8}$$

where ρ is the air density, C_H and C_E are the drag coefficients (the parameterization of Arakawa (1972) is used), w is the soil moisture content, w_0 is the soil field capacity, q_s is the saturated specific humidity at the surface temperature, ϑ_s, q is the surface air specific humidity.

So, the partitioning of energy among the components of surface heat balance (4) is controlled by soil wetness and low level atmospheric variables.

4.4. *Soil Water Budget*

We use a simple 'bucket' model of ground moisture changes. The soil is assumed to have a water-holding capacity w_k, which, as well as field capacity, depends upon the soil type.

The change of soil moisture content w is altered by rainfall r, evaporation E, and runoff Y according to a water mass balance equation of the form:

$$\frac{dw}{dt} = r - E - Y. \tag{9}$$

Here, r – precipitation; E – evaporation; Y – runoff. The parameterization of Y is given in the paper of Budyko (1974):

$$Y = \begin{cases} r\frac{w}{w_k}\sqrt{0.04\left\{1-\left\{1-\frac{E_0}{r}\right\}^2\right\}+\left\{1-\frac{E_0}{r}\right\}^2} & \text{at } r > E_0 \\ 0.2\frac{w_k}{w_k} & \text{at } r <= E_0 \end{cases} \tag{10}$$

where E_0 is the potential transpiration, and w_k the water holding capacity. The physical processes taken into account in the model are advection, vertical turbulent exchange, and condensation when relative humidity exceeds 100%. In our calculations we used constant advection velocity and albedo of the surface. This was done to separate the proposed feedback from the mechanism described by Kutzbach (1985, 1987) and Charney (1975) (see above) and to make the 'lower bound' estimation of the affect of the proposed mechanism.

This model allows calculation of the mean-annual (and seasonal) precipitation, air surface temperature, soil moisture, evaporation and some other characteristics of the hydrological cycle for the different distances from the ocean coast in the direction of prevailing wind. The calculation results are compared quantitatively and qualitatively with empirical climatic data from the Atlas (1974) where the model efficiency was confirmed (Byutner and Shabalova, 1989).

TABLE I: Soil characteristics (After Zubenok, 1975, and Kobak, 1988)

	P kg/m^2/year	C kg/m^2	β years	w_0 cm
Desert	6	5.6	2000	5
Semi desert	24	10	800	8
Savannah	72	20	500	12–15
Tropical forest	93–250	14–21	~200	18–20

Using this model in combination with the model of evolution of the soil field capacity we can project the evolution of the system for a period of the order of several thousand years with reasonable expenditures of computer resources.

5. Model of Soil Parameters Evolution

Our calculation uses the classification of natural zones given by Kobak (1988). According to this classification, tropical areas with precipitation lower than 400 mm/year are deserts; those with precipitation 400–800 mm/year: semideserts; the areas with precipitation of 800–1200 mm/year: savannas; and the areas with precipitation higher than 1200 mm/year: tropical forests. (The values of organic carbon concentration in different types of soil as well as values of soil parameters are given in Table I.)

As mentioned above, every stable soil type can be associated with a certain field capacity value which depends mainly on the development of root systems (Zubenok, 1975). Quantitative data characterizing the development of root systems in different soils are not available. However, we make the reasonable assumption that the soils in which well-developed root systems exist would be characterized by high organic carbon concentrations. This is evident from the data in Table I. According to these data, the highest soil organic carbon concentration (in Africa) is characteristic of tropical forests and savannas, while the lowest one – of deserts. Based on these data, it can be supposed that the boundaries of certain soil organic carbon concentration values mark the changes in field capacity which influence the value of evaporation from the underlying surface. This supposition is the formalization of the relationship between soil field capacity and root systems.

Suppose the accumulation of carbon in the soils of marginal regions can be described by the following balance equation:

$$\frac{dC}{dt} = \alpha_1 P - \alpha_2 C, \tag{11}$$

where C – the quantity of the organic carbon; P – the productivity of the plants on the given soil type; α_1 – the portion of soil organic carbon provided by litter fall

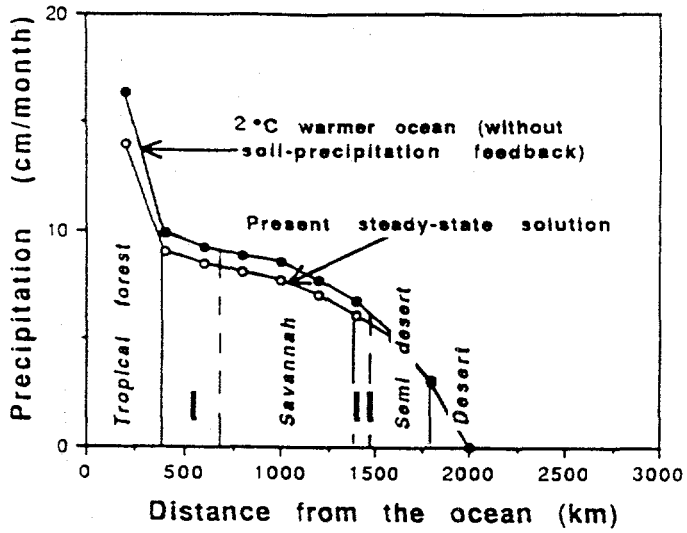


Fig. 1. The distributions of the mean-annual precipitation for present steady-state climate and for 2 °C warmer ocean (without the proposed feedback). Air advection velocity is constant. Changes in precipitation take place due to changes in water vapor content of the air masses coming from over the ocean to the continent. Vertical solid lines for the present and dashed lines for 2 °C warmer ocean, indicate positions of vegetation zone boundaries (the vegetation zone positions are determined by amount of mean-annual precipitation, see Table I.). Transient area I – transition from savannas to tropical forest (about 500 km from the ocean); transient area II – transition from semidesert to savannah (about 1500 km from the ocean (transient area III (see the text) is about 1800 km from the ocean and it appears later as a result of the influence of the proposed feedback).

($\alpha_1 = 0.06$ for all soils (Kobak, 1988)); $\beta = \alpha_2^{-1}$ – characteristic time necessary for the organic carbon of the given soil type to mineralize. The plant's productivity and the mineralization of organic carbon are the main source and sink of organic carbon in soils.

Presented in Figure 1 are the distributions of mean annual precipitation with the distance from the ocean in the direction of monsoon transport for the modern period (curve 1) and for an increase of 2 °C in sea water temperature (curve 2). This calculation was carried out without taking into account the changes in soil parameters and characterizes the changes in precipitation over the continents after an instantaneous sea water temperature change. In our calculation this change serves as a forcing for the studied mechanism (at the same time, we can consider curve 2 as a reaction of the system without the proposed feedback). The comparison of curve 1 and curve 2 (Figure 1) shows two marginal regions where the situation is favorable for the change of soil type. Region I corresponds to savannas-tropical forest transition, and region II to semidesert-savannas transition. As it will be shown later, in a few hundred years of calculations by our model with the proposed feedback, a new, third transient area will appear (transition from desert to semidesert).

Based on Equation (11) which describes changes in organic carbon concentration in different soil types and with the above-mentioned supposition taken into account, it is possible to calculate changes in field capacity in the marginal regions (for example, marginal region II):

$$\begin{cases} \frac{dC}{dt} = \alpha_1 P^{(s)} - \alpha_2^{(s)} C \\ C|_{t=0} = C^{(s/d)} \end{cases} \tag{12}$$

$$w_0 = w_0^{(s/d)} + l(t)\left\{w_0^{(s)} - w_0^{(s/d)}\right\} \tag{13}$$

$$l(t) = \frac{C(t) - C^{(s/d)}}{C^{(s)} - C^{(s/d)}}, \tag{14}$$

where parameters with index (s) refer to savannas, with index (s/d) – semidesert, and are taken from Table I. The general scheme of the calculation which takes into account the feedback mechanism between soil parameters and precipitation is as follows:

1. The value of coastal sea water temperature increase is prescribed.
2. Seasonal precipitation distribution in the direction of predominant transport depending on the distance from the coast is calculated for initial steady-state conditions and warmer ocean (greater water vapor flux to the land). Advection velocity is constant, while humidity increases. Simultaneously, the heat balance and surface air temperature are calculated.
3. The mean annual characteristics of the hydrological cycle and temperature are calculated.
4. In the marginal regions where soil organic carbon starts to accumulate, field capacity increases and local moisture circulation changes.

The last item could be explained by following simple logic analysis of the simulations results:

1. For a fixed advective input of water vapor into a region, any increase in advective output must be balanced by a decrease in runoff. This is because for the atmosphere:

   Advective input – Advective output = Precip. – Evap. (15)

   and for the surface we have Equation (9).
2. Combining water vapor balance in the atmosphere (9) with water balance on the land surface (15) we get:

   Advective input – Advective output = Runoff + Drainage (16)

   since the difference between precipitation and evaporation on the right-hand side of (15) is duplicated by the same difference in (9). The factor dw/dt vanishes for steady state conditions.

3. An increase in field capacity will tend to produce a decrease of runoff, according to equation (10). This leads to increase in soil moisture and advective output. Since in arid regions evapotranspiration is limited by absolute values of soil water content (Zubenok, 1975; Hansen *et al.*, 1983; Kellog and Zhao, 1988), evapotranspiration intensifies and leads to an increase of local precipitation.
4. After absolute soil moisture increases evapotranspiration intensifies and enhanced evaporation leads to an increase of local and downwind precipitation and a decrease in surface temprature.

The resulting steady state annual distributions of hydrological variables are characterized by higher values of precipitation, evaporation, soil moisture and total water vapor content in the troposphere, but by lower surface temprature values.

6. Results and Discussion

Calculations show that first the semidesert area contracts due to the savannas extension, and then (after 400 years) the desert starts to contract slowly. The rate of the savannas extension can be estimated based on the data from Figure 2. The mean rate of the savannas advance on semidesert is equal to 100 km per 250 years. The potential increase of the tropical forests area due to the considered feedback is quite small (within the limits of the model resolution).

According to our calculations, the field capacity increase in semidesert during its transition to savannas occurs with a mean rate of 1 cm per 200 years (transition zone II). During desert-semidesert transition field capacity increases with the rate of 1 cm per 300–400 years (transition zone III) (Figure 3a).

After a period of the order of 1–2 Kyears, moisture conditions and temperature in the intracontinental regions become stationary (curve 4 in Figure 2). At the condition when soil is closed to the saturation – on the boundary of the forest-savannah transition, the feedback mechanism, which we proposed, does not have a significant influence on the hydrological cycle. At the same time, the total shift of the semidesert on the desert is about 400 km, with the sea surface water temperature increase of 2 °C. This shift is much larger than for the case without this feedback. Hence, the primary effect of feedback is felt at a distance of about 1800 km inland, for desert-semidesert transition. The calculated expansion of savannah on Desert (~400 km) is close to that observed, for example, for Holocene optimum (Petit-Maire and Riser, 1983; Van Zinderen Bakker and Goetzee, 1988).

The transient temperature response with distance from the ocean is quite different (Figure 3b). In transient zone II it is almost monotonically decreasing, but does not cross the preforcing value. The rise of the temperature after initial forcing happens in both areas for a few seasons, so it is difficult to show such changes in the scale of Figure 3 (we put an arrow on the abscissa axes, which point to the preforcing values of the surface temperatures).

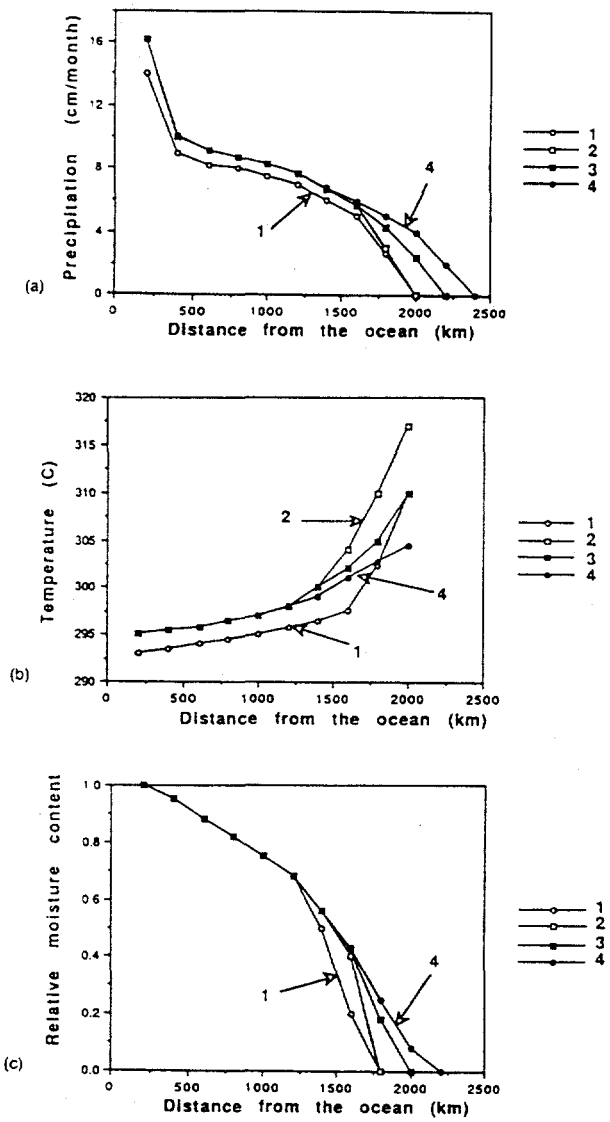


Fig. 2. (a) Evolution of the mean-annual precipitation (r); (b) Evolution of the air surface temperature (T). Curve 1a, b presents present day steady state conditions; curve 2a, b – steady-state distributions without long term soil – precipitation feedback (prescribed field capacity); curve 3a, b – 400 years after forcing with feedback influence; curve 4a, b – new steady-state distributions with feedback influence (about 1500 years after forcing); (c) Evolution of the relative soil moisture (w/w_0). Curve 1c – presents present day steady-state conditions; curve 2c – 400 years after forcing; curve 3c – 800 years after forcing; curve 4c – new steady-state distributions (about 1500 years after forcing). The distribution of the relative soil moisture for 2 °C warmer ocean and without soil-precipitation feedback is not shown because it practically coincides with the present day steady state distribution.

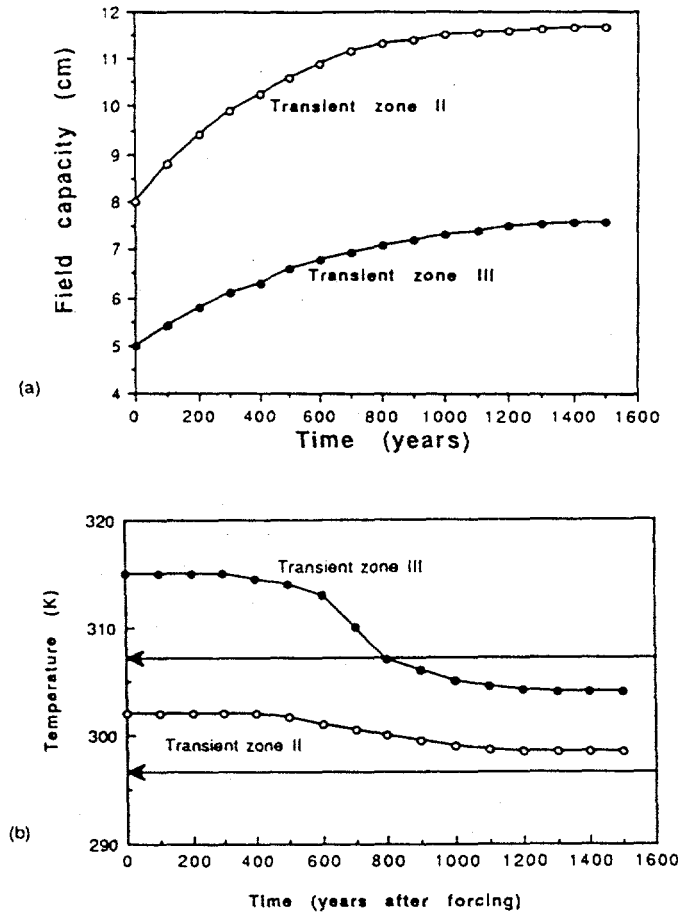


Fig. 3. (a) Evolution of the soil field capacity w_0 for transition from semi desert to savannah (transient area II, about 1500 km from the ocean) and from desert to semi desert (transient area III, about 1800 km from the ocean). At time equal to zero, the field capacities are equal to their present values; (b) Evolution of the surface underlying air temperature for transition from semi desert to savannah (transient area II), and from desert to semi desert (transient area III). The temperature values before the forcing are shown by arrows. At time equal to zero, the temperatures are equal to their corresponding values for the steady-state distribution without feedback (see curve 2b on Figure 2). The model needs only a few seasons to reach steady state conditions without the long term soil-precipitation feedback.

In transient area III (transition from desert to semi-desert, about 1800 km from the ocean), temperature is almost not changing for the first 400 years, and 400 years after forcing (due to a significant increase of the evapotranspiration), is decreasing fast, even crossing the present (preforcing) level. Such a decrease corresponds to the paleotemperature reconstructions based on the paleobotanic data (Borzenkova, 1987; Borzenkova, 1992; Gasse, 1980; Glenn *et al.*, 1986).

As we can see, the above obtained result agrees both with the observed increased precipitations and decreased temperature in continental interiors.

7. Conclusion

Simulation of the advance of vegetation into the arid area when long-term evolution of field capacity is taken into account inspires us to suppose that introduction into climatic models of this kind of mechanism would provide improvement of the results of calculations for steady state climates and better agreement with paleoclimatic data.

Acknowledgements

We thank Dr. T. Volk, and Dr. I. I. Borzenkova for discussions. The notes that were made by the reviewers, especially 'B' and 'D', were very useful. The comments and suggestions of the Re-re-reviewer were extremely helpful and we express our gratitude to this person. We thank D. Collins for help with editing this paper. This work was supported at NYU by the U.S. Department of Energy (DE-FG02-90ER 61014).

References

Arakawa, A.: 'Design of the UCLA General Circulation Model', Numerical simulation of weather and climate technical report No. 7, Dept. of Met., UCLA.

Atlas of the World Water Balance: 1974, Moscow, Gidrometeoizdat.

Barash, M. S., Oskina, N. S., Ivanova, Ye. I.: 1980, 'The Upwellings of African Coastal Waters in the Late Pleistocene Studied by the Data on Planktonic Foraminifers', *Oceanology* **20**, in Russ.

Borzenkova, I. I., 1987, 'The Features of Moisture Conditions on the Land in the Northern Hemisphere in the Geological Past', *Meteorol. Gidrol.* **10**, in Russ.

Borzenkova, I. I., 1990, 'Moisture Conditions in the Arid Regions of the Northern Hemisphere during the Late Glacial-Holocene', *Izv. Acad. Sci. USSR* **2**, in Russ.

Borzenkova, I. I., Zubakov, V. A., Lapenis, A. G., 1992, 'Global Climate Changes during the Warm Epochs of the Past', *Meteorol. Gidrol.* **8**, 25–37, in Russ.

Budyko, M. I.: 1974, *Climate and Life*, English translation edited by Miller, D. H., Academic Press, N.Y.

Budyko, M. I. and Izrael, Ya. A.: 1991, *Anthropogenic Climatic Change*, University of Arizona Press, Tucson, pp. 277–318.

Byutner, E. K. and Shabalova, M. V.: 1989, 'Modeling the Seasonal Course of Hydrologic Cycle Components', *Meteorol. Gidrol.* **12**, in Russ.

Byutner, E. K. and Shabalova, M. V., 1990, 'Modeling of Hydrological Cycle in Ocean-Atmosphere-Active Soil Layer system', *Meteorol. Gidrol.* **10**, in Russ.

Charney, J. C.: 1975, 'Dynamics of Desert and Drought in the Saheli', *Quart. J. R. Met. Soc.* **101**, 193. CLIMAP Project Members: 1984, 'The Last Interglacial Ocean', *Quat. Res.* **21**, 123–224.

Climate Change, The IPCC Scientific Assessment: 1990, Houghton, J. T., Jenkins, G. J., and Ephraums, J. J. (eds.), Cambridge University Press.

Climate Change in Australia Would Bring Heat and Extreme Rainstorms, Climate Alert **4**(4), 1991.

Crowley, T. C.: 1990, 'Are there Any Satisfactory Geologic Analogs for a Future Greenhouse Warming?' *J. Clim.* **3**, 1282–1292.

Crowley, T. C.: 1991, 'Modeling Pliocene Warmth', *Quat. Sci. Rev.* **10**, 275–282.

Eagleson, P. S.: 1982, 'Land Surface Processes in Atmospheric General Circulation Models', Cambridge Univ. Press, Cambridge, U.S.A.

Gasse, F.: 1980, 'Quaternary Changes in Lake-Levels and Diatom Assemblage on the South-Eastern Margin of the Sahara', *Paleoecol. Africa* **12**, 333–350.

Glenn, A., Goodfriend, M., Magaritz, M., Carmi, I.: 1986, 'A High Stand of the Dead Sea at the End of the Neolithic Period: Paleoclimatic and Archaeological Implications', *Clim. Change* **9**(3), 349–356.

Hansen, J. *et al.*: 1983, 'Efficient Three-Dimensional Global Models for Climate Studies: Model I and II', *Mon. Wea. Rev.* **111**(4), 609–662.

Hansen, J. *et al.*: 1988, 'Global Climate Changes as Forecast by Goddard Institute for Space Studies Three-Dimensional Model', *J. Geophys. Res.* **93**(D8).

Hoffert, M. I. and Covey, C.: 1992, 'Deriving Global Climate Sensitivity from Paleoclimate Reconstructions', *Nature* **360**, 573–576.

Houghton, J. T. *et al.* (eds.): 1990, *Climate Change: The IPCC Scientific Assessment*, Cambridge Univ. Press.

Kellog, W. W. and Zhao, Z.: 1988, 'Sensitivity of Soil Moisture to Doubling of Carbon Dioxide in Climate Model Experiments', *J. Clim.* **1**, 348–378.

Kobak, K. I.: 1988, 'Biotical Compounds of Carbon Cycle', Leningrad, Gidrometeoizdat, in Russ.

Kutzbach, J. E. and Otto-Bliesner, B. L.: 1982, 'The Sensitivity of the African-Asian Monsoonal Climate to Orbital Parameter Changes for 9000 Years B.P. in a Low-Resolution General Circulation Model', *J. Atmos. Sci.* **39**, 1177–1188.

Kutzbach, J. E. and Guetter, P. J.: 1984, 'The Sensitivity of Monsoon Climates to Orbital Parameter Changes for 9000 Years B.P.: Experiments with the NCAR General Circulation Model', in Berger, A., Imbrie, J., Hays, J. D., Kukla, G. J., and Saltzmann, B. (eds.), *Milankovich and Climate*, D. Reidel, Dordrecht, Netherlands, pp. 801–820.

Kutzbach, J. E. and Street-Perrott, F. A.: 1985, 'Milankovich Forcing of Fluctuations in the Level of Tropical Lakes from 18–0 kyr B.P.', *Nature* **317**, 130–134.

Manabe, S. and Wetherald, R. T.: 1980, 'On the Distribution of Climate Change Resulting from an Increase of CO_2 Content of the Atmosphere', *J. Atmos. Sci.* **37**, 99–118.

Murzayeva, I. E., Konopleva, V. I., Devyatkin, Ye. V., Serebryanny, L. R.: 1984, 'Pluvial Patterns during the Late Pleistocene and Holocene in the Arid Zones of Asia and Africa', *Izv. Acad. Sci. USSR* **N4**, ser. geogr. in Russ.

Petit-Maire, N. and Riser, J. (eds.): 1983, *Sahara au Sahel?*, Quaternaire recent du Bassin de Taoudenni (Mali), Marseille, Imprimerie Lamy, 473 pp.

Poore, R. Z. and PRISM Project Members: 1993, 'Pliocene Climate Reconstruction – Mid-Gauss Warming', Supplement to *EOS*, April 20, p. 49.

Prell, W. L. and Kutzbach, J. E.: 1987, 'Monsoon Variability over the Past 150,000 Years', *J. Geophys. Res.* **92**, 8411–8425.

Rind, D.: 1984, 'The Influence of Vegetation on the Hydrologic Cycle in a Global Climate Model', Climate Processes and Climate Sensitivity, Geophysical Monograph 29, Maurice Ewing Volume 5, 73–91.

Sarthein, M. *et al.*: 1982, 'Atmospheric and Oceanic Circulation Pattern off Northwest Africa during the Past 25 Million Years', in *Geology of the Northwest African Continental Margin*, Springer-Verlag.

Sellers, W. D.: 1965, *Physical Climatology*, University of Chicago Press, 272 pp.

Shine, K. P. *et al.*: 1990, in Houghton, J. T., Jenkins, G., and Ephraums, J. J. (eds.), *Climate Change: The IPCC Scientific Assessment*, Cambridge Univ. Press, pp. 41–68.

Shukla, J. and Mintz, Y.: 1982, 'Influence of Land-Surface Evapotranspiration on the Earth's Climate', *Science* **215**, 1498–1501. Strahler, A. N.: 1969, *Physical Geography*, Wiley, (3d ed.), 733 pp.

Street-Perrot, F. A. and Perrot, R. A.: 'Abrupt Climate Fluctuations in the Tropics: The Influence of Atlantic Ocean Circulation', *Nature* **243**, 6259.

Washington, W. M. and Meehl, G. A.: 1984, 'Seasonal Cycle Experiment on the Climate Sensitivity due to Doubling of CO_2 with an Atmospheric General Circulation Model Coupled to a Simple Mixed Layer Ocean Model', *J. Geophys. Res.* **89**(D6).

Williams, M. A. J., Faure, H. (eds.): 'The Sahara and the Nile', A. A. Balkema, Rotterdam.

van Zinderen Bakker, E. M. and Goetzee, J. A.: 'A Review of Late Quaternary Pollen Studies in East, Central and Southern Africa', *Rev. Paleobot. Palynol.* **55**, 155–174.
Zubenok, L. I.: 1975, 'Evaporation on the Continents', Leningrad, Gidrometeoizdat, in Russ.

(Received 18 February, 1992; in revised form 20 December, 1993)

Pergamon

0360-5442(94)E0003-K

Energy Vol. 19, No. 9, pp. 967–974, 1994
Elsevier Science Ltd. Printed in Great Britain.

THE EFFECTIVENESS OF MARINE CO_2 DISPOSAL

H. S. KHESHGI,†‡ B. P. FLANNERY,† M. I. HOFFERT,§ and A. G. LAPENIS§

†Corporate Research, Exxon Research and Engineering Company, Annandale, NJ 08801 and §Department of Applied Science, New York University, New York, NY 10003, U.S.A.

(Received 22 September 1993; received for publication 21 January 1994)

Abstract—We consider the relative effectiveness of engineered systems for collection and marine disposal of CO_2 from fossil-fuel-fired power plants using comparisons of the trend with time of CO_2 concentration in the atmosphere from systems with and without marine disposal. Since the retention time for CO_2 increases rapidly with depth of disposal, from a few years in the mixed layer up to several centuries as the depth increases beyond 1000 m, deep oceans have been considered as potential storage sites for CO_2. However, CO_2 collection and disposal consumes energy and produces extra CO_2. We show that some of this extra CO_2 reaches the atmosphere, so that atmospheric concentrations from systems with marine disposal ultimately exceed those from systems without controls. In some circumstances, they do so rapidly, making marine disposal less favorable than direct atmospheric release. The Global Warming Potential (GWP) can be used to compare different systems, but results for GWP highlight the difficult issue of determining what time scales are important in considering options to reduce concerns about global warming.

INTRODUCTION

To slow or reduce human emissions of greenhouse gases, Marchetti[1] proposed that CO_2 emissions from power plants might be captured and stored in deep oceans. Relative to human emissions, oceans have an immense capacity for CO_2 and it appears that CO_2 retention times could exceed 1000 years for marine disposal below a few km.[2] Here we develop a systems approach to assess the relative effectiveness of options to limit emissions of CO_2 based on capture and marine disposal. Control of CO_2 emissions by capture and marine disposal entails, e.g., processes for gas separation, compression, chemical reaction, transportation and power conversion. We evaluate the relative buildup of atmospheric CO_2 with time for a base case system (a fossil-fuel-powered electrical utility) producing a given amount of power plus accompanying CO_2 that is released directly to the atmosphere, and for various configurations with CO_2 emission controls that produce identical amounts of delivered power. In our analysis, we use simple parameterized models to consider the behavior of two systems: the first describes the fate of CO_2 emitted into the atmosphere or deposited in the ocean and the second describes the engineered system for capture and disposal of CO_2. We combine the models to describe the buildup of atmospheric CO_2 and to evaluate relative Global Warming Potential as a diagnostic of performance.

MODELLING THE OCEAN–ATMOSPHERE CARBON CYCLE

We develop two simple approximate relations to represent air–sea exchange of CO_2. The first describes transfer of excess CO_2 from the atmosphere to the oceans. The second describes mixing and transfer of excess CO_2 from a limited region of the ocean to the atmosphere. However, it is important to note two major caveats. In spite of intense effort to understand the observed growth of atmospheric CO_2 with time and to predict future trends, carbon-cycle models cannot explain why the observed growth rate of atmospheric CO_2 is 3–4 GtC/yr (1 GtC = 1×10^9 tonnes C = 1×10^{15} gC) when the total human emission rate appears to be ~8 GtC/yr.[3] Absorption of CO_2 by the oceans accounts for only a part of the 4–5 GtC/yr difference. Apparently, the terrestrial biosphere plays an important role, which has not been incorporated into the models we describe below. Also, we assume that the behavior of the ocean–atmosphere system is insensitive to the small additions of CO_2 that we consider. However, our results apply over times long enough that climate change might alter the system. Ultimately, long

‡To whom all correspondence should be addressed.

term assessment of disposal schemes will require scenarios of climate change itself, but that is beyond current capability.

The carbon cycle describes processes that store and transfer carbon among various reservoirs, including the atmosphere and marine and terrestrial biosphere. CO_2 migrates across the air–sea interface driven by differences in its partial pressure in the two media. Ocean dynamic mixing of water masses and biospheric activity in the water column establish the distribution of marine carbon. Over longer periods, creation and dissolution of calcareous sediments become important. Carbon cycle models of the ocean–atmosphere system have been used to simulate the fate of excess atmospheric CO_2. Maier-Reimer and Hasselmann[4] simulated a well-mixed global atmosphere and a three dimensional ocean with complex circulation and mixing, although without sediment or biospheric interaction, and fit the simulated decay of the perturbation $C(t)$ of atmospheric CO_2 resulting from a pulse emission of CO_2 of size $C_A(0)$ into the atmosphere at time $t = 0$ to the function

$$f_A(t) \equiv C(t)/C_A(0) = \left[a_0 + \sum_{j=1}^{4} a_j \exp(-t/\tau_j)\right]. \tag{1}$$

For the smallest pulse considered to an atmosphere containing CO_2 at a preindustrial level of 265 ppm, Maier-Reimer and Hasselmann[4] found the parameters $a_j = (0.131, 0.201, 0.321, 0.249, 0.098)$ and $\tau_j = (362.9, 73.6, 17.3, 1.9)$ yr. For a pulse to an atmosphere containing 360 ppm CO_2 (near current value), the lower solubility (higher "buffer factor")[5] of CO_2 in seawater is accounted for by using the parameter $a_j = (0.160, 0.194, 0.310, 0.241, 0.095)$ to match the response seen in models with contemporary levels of atmospheric CO_2.[6] After a time $t \gg \tau_1$, the airborne fraction $f_A(t)$ approaches the asymptotic limit $C(\infty)/C_A(0) \to a_0$. From Eq. (1), the transfer rate of excess CO_2 into the ocean is

$$-E_{atm}(t) \equiv -\frac{dC(t)}{dt} = C_A(0) \sum_{j=1}^{4} a_j \frac{\exp(-t/\tau_j)}{\tau_j}, \tag{2}$$

where E is the evasion rate of CO_2 from the oceans to the atmosphere.

Next we develop an approximate model to describe time dependent evasion to the atmosphere of a pulse $C_M(0)$ of CO_2 injected into the ocean at time ($t = 0$). For evaluation of specific disposal sites this requires a three dimensional ocean model. However, for this analysis of a generic disposal site we utilize a much simpler one dimensional, upwelling-diffusion (UD) model. Such models have been used in the past to study many features of the global ocean in an approximate fashion.[7,8] Alternatively, the model might approximate the behavior of a small, isolated ocean basin. Our goal is to obtain results in a form that is convenient for analysis of marine disposal schemes, and we believe the functional form for CO_2 evasion from an UD model suits that objective. In an UD model, the time evolution of the distribution of carbon $c(z,t)$ as a function of depth z and time t is described by

$$\frac{dc}{dt} = \frac{d(wc)}{dz} + \frac{d}{dz}\left(K \frac{dc}{dz}\right). \tag{3}$$

The values of upwelling velocity $w = 4$ m/yr and diffusion coefficient $K = 2000$ m^2/yr used here match global average distributions of temperature and chemical tracers.[7,8] Given a pulse injection $C_M(0)$ of carbon at depth d into a small isolated ocean region,† the solution to Eq. (3) for the surface evasion rate at time t is

†Application of the UD model to an isolated ocean basin of small surface area compared to the total ocean area implies the boundary condition $c|_{z=0} = 0$ (our analysis deviates from a preliminary report[9] of this work at this point). In this case a significant amount of CO_2 does not penetrate to the ocean floor, allowing the convenient boundary condition $c|_{z=\infty} = 0$. Injection into a small ocean basin implies that all CO_2 will eventually evade, and then a fraction will be reabsorbed at the total ocean surface according to (2). While the upwelling isolated basin is a limiting case, evasion functions of time-profile similar to Eq. (4) can be constructed from detailed mixing studies[10] for specific disposal sites. Evasion functions estimated from detailed studies[10] do encompass (for disposal at the same depth but different locations) the evasion function given by Eq. (4).

$$E_{UD}(t) = C_M(0)\frac{d}{2\sqrt{\pi K t^3}}\exp\left[-\frac{(wt-d)^2}{4Kt}\right]. \quad (4)$$

Evasion occurs in a profile that peaks prior to the time $t = d/w$ required for upwelling to reach the surface, and extends over a time period $\sim(Kd/w^3)^{1/2}$ available for diffusion to occur. Also, in this model all injected carbon ultimately escapes to the atmosphere. We assume that CO_2 enters the atmosphere according to E_{UD} and then decays from the atmosphere into the global ocean according to E_{atm}, Eq. (2).

Different behavior is expected if CO_2 is injected into the well-mixed layer of depth $h \sim 100$ m that caps the ocean. Previously, we considered injection at depths $d \gg h$, such that upwelling dominated the time for CO_2 to reach the surface. However, for shallow injection, e.g. into the mixed layer itself, evasion is limited by mass transfer at the air/sea interface which is modeled as

$$E_m(t) = \frac{C_M(0)}{\tau_m}\exp\left(-\frac{t}{\tau_m}\right), \quad \tau_m \approx \left(\frac{h}{100\ m}\right)\text{ yr}, \quad (5)$$

and is far more rapid than at greater depths where Eq. (4) is applicable.

While this treatment neglects many factors,† the evasion profile captures essential characteristics expected from marine disposal. Evasion follows a depth dependent delay and occurs in a smooth profile broadened by diffusive effects. Analysis of an actual disposal scheme would require consideration of the appropriate mixing pattern. For instance, settling schemes[11] would produce a distribution of CO_2 with depth, rather than a localized point injection. This could be modeled by a broader evasion function. Alternatively, disposal might be into a downwelling current in which case analysis would require a 3D simulation.[10]

EFFICIENCY OF SYSTEMS FOR CO_2 CAPTURE AND DISPOSAL

In our base case of an engineered system for CO_2 capture and disposal, a coal-fired power plant emits an amount of C_0 of CO_2 to deliver an amount Q_0 of electricity. We scale systems with CO_2 capture and disposal to produce the same amount of electricity, and we assume that CO_2 produced scales with energy as in the base case. Table 1 lists parameters that we use to describe inefficiencies in systems with CO_2 emission controls. We use values in the last column of Table 1 for illustrative evaluation of system factors, defined below. In proposed designs these parameters span considerable

Table 1. Parameters describing system efficiency for CO_2 capture and disposal.

Reference amount	Inefficiency as a fraction of reference amount	Symbol	Illustrative Value
Energy produced in plant	Energy used for CO_2 capture	ε	0.2
CO_2 produced in plant	CO_2 emitted at plant	ξ	0.1
CO_2 collected for disposal	CO_2 lost to air during disposal	σ	0.01
CO_2 collected for disposal	CO_2 from energy used in disposal	δ	0.05

†To evaluate a specific disposal option will require not only details of mixing behavior and the engineered system for CO_2 capture and disposal, but also the expected CO_2 concentration when implemented, since this will affect the ocean's ability (the "buffer factor") to absorb more CO_2. These effects can be included in this model framework by modification of evasion functions (2) and (4).

ranges depending on actual technologies implemented. The first two depend on the approach by which CO_2 is captured during power generation. The final two parameters depend sensitively on the location of the plant and disposal site. These require individual analysis, but certainly depend on distance to and depth of disposal. A plant consumes additional energy Q_1 to carry out CO_2 capture, while continuing to deliver usable energy Q_0. The produced CO_2 scales as total energy, $Q = Q_1 + Q_0$, and

$$Q = Q_0 + Q_1, \quad Q_1 = \varepsilon Q, \quad Q_0 = (1 - \varepsilon)Q,$$
$$\frac{C}{C_0} = \frac{Q}{Q_0} = \frac{1}{1 - \varepsilon}. \tag{6}$$

The parameter ε represents the fraction of energy Q required to capture and compress CO_2. Of the produced CO_2, some fraction ξ is emitted immediately at the plant as a result of incomplete capture of CO_2, leaving an amount $(1 - \xi)C$ of CO_2 actually collected and available for disposal. Additional immediate emissions to air, $\sigma(1 - \xi)C$, occur as losses during transport and disposal. Finally, we allow for emissions associated with the use of external energy sources during transport and disposal, and we assume that these can be represented as some proportion, $\delta(1 - \xi)C$, of the amount being disposed. With these definitions, for the modified plant, the amounts of CO_2 which directly enter the atmosphere (C_A) and the ocean (C_M) at time $t = 0$ are

$$C_A = [\xi + (\sigma + \delta)(1 - \xi)]C, \quad C_M = [1 - \xi - \sigma(1 - \xi)]C$$
$$\alpha \equiv \frac{C_A}{C_0} = \frac{\xi + (\sigma + \delta)(1 - \xi)}{1 - \varepsilon}, \quad \Omega \equiv \frac{C_A + C_M}{C_0} = \frac{1 + \delta(1 - \xi)}{1 - \varepsilon}. \tag{7}$$

The total amount of CO_2 generated by a plant with marine disposal always exceeds that in the base case (for values in Table 1, $\Omega = 1.3$), and there is some immediate emission of CO_2 to the atmosphere (for values in Table 1, $\alpha = 0.19$). Figure 1 shows contours of Ω and α, as a function of ε and ξ, for a range of values representative of estimates for performance of various systems for capture of CO_2 from power plants.

ATMOSPHERIC CO_2 RESULTING FROM PULSE INJECTIONS

The variation with time of the airborne fraction of CO_2 normalized relative to atmospheric emission in the base case implied by the previous discussion on carbon cycle is given by

$$\frac{C_A(t)}{C_0} = \alpha f_A(t) + \int_0^t \frac{E_{UD}(\chi)}{C_0} f_A(t - \chi)\mathrm{d}\chi. \tag{8}$$

Figure 2 shows the time dependent perturbation of atmospheric CO_2 resulting from pulse injection for representative cases of disposal at various ocean depths, and with various inefficiency factors. In all cases the airborne fraction with emission controls ultimately exceeds that in the base case. Benefits, in terms of time when atmospheric concentrations are reduced below those in the base case, may persist for only short periods (< 10 years) for unfavorable combinations of inefficiency and injection at shallow depth. Benefits over only short periods are of little or no value to mitigate greenhouse concerns. With deeper marine disposal the cross over time depends sensitively on both depth of injection and engineering efficiency. Note that in all cases, after a time, $t \gg \tau_1$, the airborne fraction approaches an asymptotic limit, $C_A(\infty)/C_0 \rightarrow \Omega a_0$. Clearly, since $\Omega > 1$, capture and disposal ultimately lead to higher levels of CO_2 than would occur in the base case without emission controls.

USING GLOBAL WARMING POTENTIAL TO INTERPRET RESULTS

The concept of a Global Warming Potential (GWP) provides one basis to interpret these results. GWP has been under development as a policy tool to compare the relative contributions of different

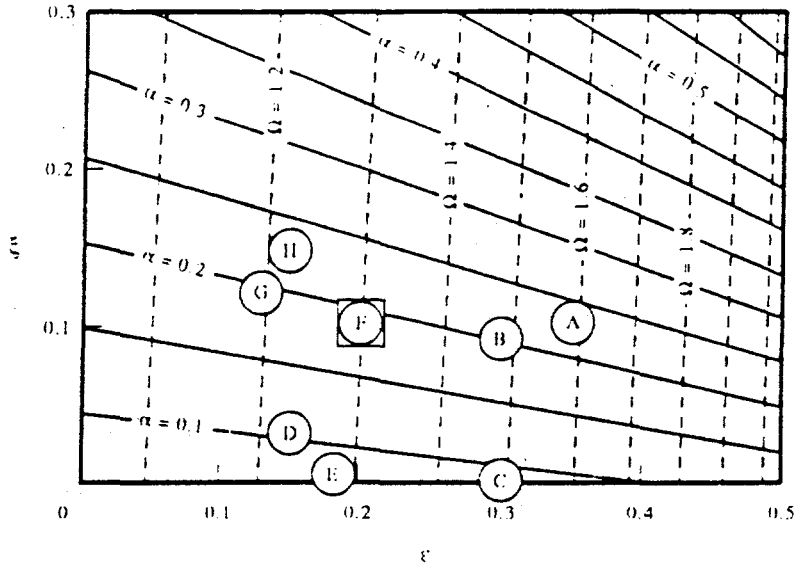


Fig. 1. Variations of α and Ω as functions of ε and ξ (with $\sigma = 0.01$ and $\delta = 0.05$). Relative to a base case power plant producing given amounts of electricity and CO_2, α (with contours drawn as solid lines) is the amount of CO_2 immediately released to the atmosphere, and Ω (with contours drawn as dashed lines) is the total amount of CO_2 produced by the system. Characteristics of the capture system are fraction ε of total energy Q produced that is required for capture, and fraction ξ of CO_2 produced at the power plant that is emitted immediately to the atmosphere as the result of incomplete CO_2 capture. Circles indicate values for plant designs estimated in studies of specific systems, while the square is the illustrative value given in Table 1. Point A (Ref. 14) and B (Ref. 15) are for retrofitted coal-fired plants with separation of CO_2 from flue gases by chemical absorption using amines. C (Ref. 16) is for air separation to produce oxygen followed by coal combustion with flue gas recycle. D (Ref. 15) and E (Ref. 15) use an integrated-gasifier, combined-cycle (IGCC) power plant to make synthesis gas which is combusted with oxygen separated from air to drive a turbine (unproved technology). F (Ref. 17), G (Ref. 18) and H (Ref. 15) use IGCC where CO_2 is recovered from synthesis gas by performing a shift reaction to make CO_2/H_2, and then physical adsorption (by selexol) to capture the CO_2 from H_2.

gases to global warming. Basically, the driver for enhancement of the natural greenhouse effect is the cumulative radiative forcing associated with increasing gas concentration. For a given gas GWP(Δt) is evaluated as the ratio of the integral over the time period extending to the *Time Horizon* Δt of radiative forcing from a pulse injection of that gas relative to the integrated forcing from an equal pulse injection (by weight) of CO_2. As applied to the relative ranking of different gases, there are many reasons to criticize GWP. Relative radiative forcing is computed with simple models at a single point in the atmosphere, the tropopause, which may not adequately represent effects on climate or ground level temperature. For some gases, e.g. methane, it is not yet possible to evaluate indirect effects, such as influence on atmospheric chemistry and structure, which strongly influence their radiative forcing. Nor is it possible to evaluate the integral over time reliably since factors that control atmospheric residence time are uncertain, especially for CO_2 the reference gas. Even if GWP could be evaluated reliably, from a policy perspective one needs to assess relative costs (and benefits) of implementing controls on emissions as a function of time, not just their forcing. However, perhaps the simplest criticism is the difficulty, in a policy setting, of accounting for the variation of GWP with Δt.

In spite of these criticisms GWP has become a conventional diagnostic. We can readily extend the idea of GWP to compare power-plant systems with and without CO_2 emission controls by defining

$$\mathrm{GWP}(\Delta t) = \int_0^{\Delta t} \frac{C_A(\tau)}{C_0}\,\mathrm{d}\tau \Big/ \int_0^{\Delta t} f_A(\tau)\mathrm{d}\tau. \tag{9}$$

GWP for a plant with controls is given by the time integral of its (airborne) fraction, as shown in Fig. 2, relative to the same integral for the base case. Results are shown in Fig. 3 for the same cases as in

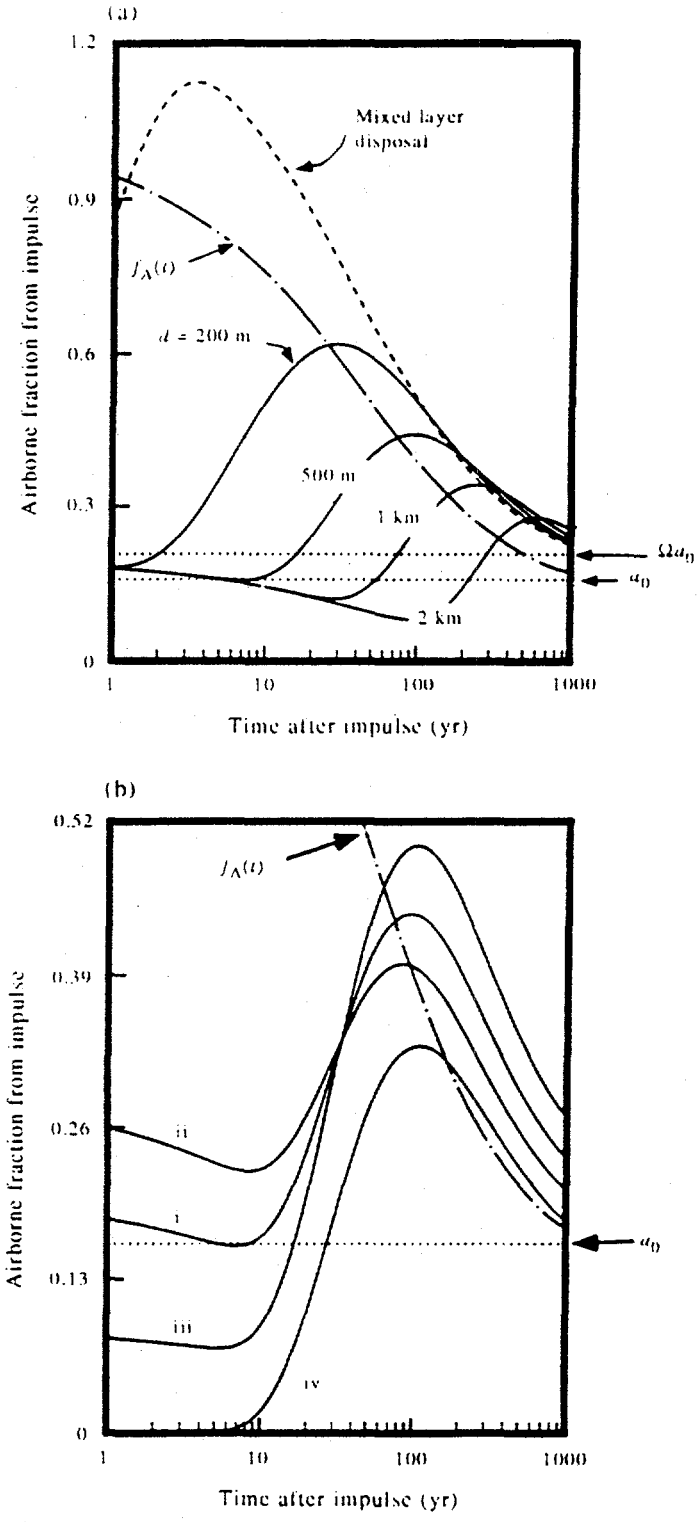


Fig. 2. Variation of atmospheric CO_2 with time for pulse injection. CO_2 amounts are normalized relative to emission in the base case. Panel (a) shows dependence on injection depth with other parameters at illustrative values (as cited in Table 1). The airborne fraction asymptotically approaches $a_0 = 0.16$ for direct release of CO_2 into the atmosphere, and $\Omega a_0 = 0.21$ for marine disposal with illustrative parameters values. Panel (b) shows dependence on engineering inefficiencies for injection at 500 m: (i) illustrative values (see Table 1); (ii) capture with low energy requirements and less complete capture ($\varepsilon = 0.1$ and $\xi = 0.2$); (iii) capture with high energy requirements and complete capture ($\varepsilon = 0.3$ and $\xi = 0.$); (iv) no energy requirements or immediate CO_2 emission for capture and marine disposal ($\varepsilon = \xi = \sigma = \delta = 0$).

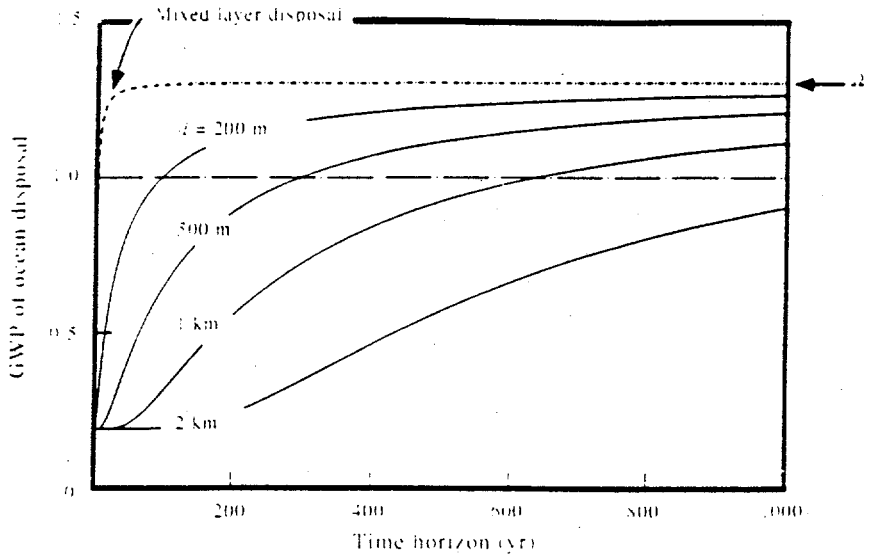


Fig. 3. Variation of Global Warming Potential with time. Curves show dependence of GWP on injection depth with other parameters at illustrative values, as in Fig. 2(a). For marine disposal, the GWP curves asymptotically approach Ω.

Fig. 2(a). In all cases considered here, after some time, GWP with controls ultimately exceeds 1, i.e. with controls cumulative impact on climate change ultimately is larger.†

Although the application of GWP here to compare different systems emitting the same gas, CO_2, avoids many of the technical difficulties encountered in evaluating GWP for different gases, we remain faced with the interpretative questions concerning the Time Horizon. In CO_2 emission-control systems for power plants, installation costs occur early, benefits disappear after a time. It will be important to determine how long a time period of benefits might be required for this option, capture and marine disposal of CO_2, to make sense. Such considerations also establish performance criteria for system design.

Systems for capture and marine disposal of power plant CO_2 emissions require considerable amounts of energy to delay or reduce the appearance of CO_2 in the atmosphere. They do not appear to offer associated economic or environmental benefits. Indeed, environmental concerns remain to be evaluated. Moreover, their utilization would accelerate depletion of finite fossil-fuel reserves. Thus, as options to mitigate CO_2 concerns, they do not appear to be as attractive in the near term as other approaches, such as energy efficiency and forest protection.

REFERENCES

1. L. Marchetti, *Clim. Change* **1**, 59 (1977).
2. M. I. Hoffert, Y. C. Wey, A. J. Callegari, and W. S. Broecker, *Clim. Change* **2**, 53 (1979).
3. U. Siegenthaler and J. L. Sarmiento, *Nature* **365**, 119 (1993).
4. E. Maier-Reimer and K. Hasselmann, *Clim. Dynam.* **2**, 63 (1987).
5. M. I. Hoffert, A. J. Callegari, and C. T. Hseih, in *Carbon Cycle Modelling, Scope 16*, pp. 287–305, B. Bolin ed., Wiley, New York, NY (1981).
6. H. S. Kheshgi, B. P. Flannery, and M. I. Hoffert, *J. Geophys. Res.* **86**, 4957 (1991).
7. W. D. Broecker and T.-H. Peng, *Tracers in the Sea*, Eldigeo Press, Lamont-Doherty Geological Observatory, Palisades, New York, NY (1982).
8. A. K. Jain, D. J. Wuebbles, H. S. Kheshgi, and M. I. Hoffert, 1993 Fall AGU Meeting, *EOS* **74**, 163 (1993).
9. B. P. Flannery, H. S. Kheshgi, M. I. Hoffert, and A. G. Lapenis, *Energy Conversion and Management* **34**, 983 (1993).

†Marine disposal concepts leading to permanent storage of CO_2 have been suggested. One concept discussed by, e.g., Wilson[12] relies on enhanced dissolution of carbonate sediments on the sea floor accomplished by contacting sediments with CO_2, this could increase the solubility of CO_2 in the oceans. Another concept discussed by, e.g., Shindo et al[13] is to create a stable liquid pool of CO_2 on the sea floor which would be somehow protected from the effects of ocean mixing. While there are great benefits to permanent storage, the feasibility of these concepts has not yet been proven.

10. G. Steegan, K. Cole, and R. Bacastow, *Energy Conversion and Management* **34**, 857 (1993).
11. P. M. Haugan and H. Drange, *Nature* **357**, 318 (1992).
12. T. R. S. Wilson, *Energy Conversion and Management* **33**, 627 (1992).
13. Y. Shindo, P. C. Lund, Y. Fujioka, and H. Komiyama, *Energy Conversion and Management* **34**, 1073 (1993).
14. W. Fulkerson, R. R. Judkins, and M. K. Sanghvi, *Scientific American* **263**, 128 (1990).
15. K. Blok, in *Greenhouse Issues* **6**, IEA Coal Research, London (1992).
16. H. Herzog, D. Golomb, and S. Zemba, *Env. Prog.* **10**, 64 (1991).
17. G. S. Booras and S. C. Smelser, *Energy—The International Journal* **16**, 1295 (1991).
18. C. A. Hendriks, K. Blok, and W. C. Turkenburg, *Energy—The International Journal* **16**, 1277 (1991).

GLOBAL BIOGEOCHEMICAL CYCLES, VOL. 9, NO. 1, PAGES 153–166, MARCH 1995

Distribution of radiocarbon as a test of global carbon cycle models

Atul K. Jain,[1,4] Haroon S. Kheshgi,[2] Martin I. Hoffert,[3] and Donald J. Wuebbles[1,4]

Abstract. Accurate global carbon cycle models are needed to estimate the future change of atmospheric CO_2 for specified scenarios of CO_2 emissions. Model accuracy cannot be tested directly because of the difficulty in estimating the carbon flux to the oceans and the terrestrial biosphere. However, one test of model consistency is the requirement that the model reproduce past changes and spatial distributions of ^{14}C. A model for carbon exchange within and among the atmosphere, oceans, and terrestrial biosphere is found to satisfy this test. The ocean is modeled as an upwelling-diffusion column capped by a mixed layer with recirculation of the polar bottom water to complete the thermohaline circulation. This ocean advection scheme contains only two key dynamic parameters, the vertical eddy diffusivity κ and the upwelling velocity w, which are calibrated to match the vertical distribution of preanthropogenic ^{14}C. The thermocline depth scale κ/w = 1343 m found by calibration is considerably deeper than that required to match the steady vertical temperature profile (500 m). This is consistent with the hypothesis that isopycnal mixing, which is much more rapid than diapycnal mixing, has a stronger effect on ^{14}C than on temperature since isopycnals are nearly isothermal. This model is found to match measured values, within measurement error, of the prebomb decrease in ^{14}C in the atmosphere and the mixed layer due to the Suess effect, the bomb ^{14}C in the mixed layer, the bomb ^{14}C penetration depth, the bomb ^{14}C ocean inventory, and the vertical distribution of total carbon. Results are compared to those of other schematic carbon cycle models as well as those of ocean general circulation models.

Introduction

Ocean carbon cycle models calibrated to reproduce radiocarbon concentrations are playing an important role in estimating the future buildup of fossil fuel CO_2 in the atmosphere [*Toggweiler et al.*, 1989a,b; *Siegenthaler and Joos*, 1992; *Siegenthaler and Sarmiento*, 1993]. A primary motivation for tracer calibration of ocean carbon cycle models is the difficulty (associated with buffering by ocean carbonate chemistry) of directly measuring the uptake of carbon dioxide by the oceans. A hierarchy of such models has developed over the years. Computational considerations, internal consistency, and a given model's ability to reproduce observations are key factors in selecting a model type for a given application.

The earliest carbon cycle models were of the classical multibox type [*Revelle and Suess*, 1957; *Bolin and Eriksson*, 1959; *Broecker et al.*, 1971; *Machta*, 1973; *Hoffert*, 1974; *Keeling*, 1973, 1977; *Bacastow and Keeling*, 1973; *Bjorkstrom*, 1979; *Broecker and Peng*, 1982]. These offer an intuitively appealing approach to finding the time variation of the mass $N_i(t)$ of stable carbon isotopes (^{12}C or ^{13}C) or radiocarbon (^{14}C) in each of a number of well-mixed reservoirs in response to a specified carbon emission rate $N_i(t)$. Multibox models normally assume first-order transport kinetics in which the outflux from box i to some adjacent box j is $k_{ij}N_i$ leading to reservoir differential equations of the form

$$\frac{dN_i}{dt} = \sum_{i=1}^{n}(k_{ji}N_j - k_{ij}N_i) + \dot{N}_i - \lambda N_i$$

where n is the number of boxes. The term $-\lambda N_i$ applies to radiocarbon decay, where $\lambda = 1.21 \times 10^{-4}$ yr^{-1} is the decay constant for $^{14}C \rightarrow {}^{14}N + \beta^- + \nu$. Exchange coefficients k_{ij} ($i \neq j$) have inverse time units and are normally calibrated to recover observed carbon masses (or concentrations) of the reservoirs in the steady state ($dN_i/dt = 0$). The number of boxes and the disaggregation scheme for the oceans are somewhat at the discretion of the modeler. The minimum box number is three: an atmosphere ($i = a$), an ocean mixed layer ($i = m$), and a deep ocean ($i = d$). Single box oceans cannot make realistic projections of future CO_2 levels because carbon uptake is not limited by CO_2 gas exchange at the atmosphere-mixed layer interface but by deeper thermocline mixing. This is modeled to a first approximation by the k_{md} coefficient.

High-latitude box models [*Sarmiento and Toggweiler*, 1984; *Siegenthaler and Wenk*, 1984; *Knox and McElroy*, 1984; *Volk and Liu*, 1988; *Toggweiler and Sarmiento*, 1985] with separate polar- and nonpolar-mixed layer boxes were proposed to explain atmospheric CO_2 variations during glacial-interglacial transitions thought to arise from ocean composition changes pumped by circulation, solubility, or biology changes [*Volk and Hoffert*, 1985]. Inclusion of a polar box addresses this problem by

[1] Lawrence Livermore National Laboratory, Livermore, California.
[2] Exxon Research and Engineering Company, Annandale, New Jersey.
[3] Department of Earth Systems Science, New York University, New York.
[4] Now at University of Illinois, Urbana

Copyright 1995 by the American Geophysical Union.

Paper number 94GB02394.
0886-6236/95/94GB-02394$10.00

estimating the effect of composition variations in polar surface waters on the strength of the ocean CO_2 sink. However, present models show too small an effect. How did atmospheric CO_2 increase from the midst of the last glaciation to the preindustrial concentration of 278 parts per million by volume (ppmv) by the ice-core-recorded 80 ppmv? And how were ocean carbon pump changes triggered by Milankovitch changes in solar radiation over the seasons and latitudes? These remain open questions that are being actively studied.

If one divides the oceans into enough boxes, the result mimics a finite difference model of the three-dimensional (3-D) oceans. Multiocean, multibox models [*Broecker and Peng*, 1986, 1987; *Bjorkstrom*, 1986; *Kier*, 1988, 1989; *Walker*, 1991; *Walker and Kasting*, 1992] are intermediate steps, distinguishing different water masses and ocean basins but falling short of full 3-D models. They typically use tracer-calibrated exchange coefficients as opposed to physical diffusivities, although they attempt to represent volumetric internal flows correctly. Box-diffusion models [*Oeschger et al.*, 1975; *Siegenthaler and Oeschger*, 1978; *Hoffert et al.*, 1979; *Keeling et al.*, 1989a,b] are a transition between discrete mutilayer ocean box models and vertically resolved continuous diffusion models. They do not include vertical advection or paths between the high-latitude surface ocean and the deep sea. The world's oceans are modeled by vertically stacked boxes of thickness Δz with two-way mass exchange between interior layers and their top and bottom neighbors. If the exchange coefficient between adjacent layers is constant, $k = k_{i,i+1} = k_{i,i-1}$, the effective diffusivity is $\kappa = k\Delta z^2$. Outcrop-diffusion and high-latitude exchange/interior diffusion-advection (HILDA) models [*Siegenthaler*, 1983; *Siegenthaler and Joos*, 1992; *Shaffer and Sarmiento*, 1994] combine diffusive models with discrete boxes in a manner meant to mimic the exchange of high-latitude surface ocean with the deep sea along constant potential density (isopycnal) outcrops. Upwelling-diffusion (UD) models [*Wyrtki*, 1962; *Munk*, 1966; *Hoffert et al.*, 1981; *Volk*, 1984; *Shaffer*, 1989; *Kheshgi et al.*, 1991] are vertically resolved ocean models including both upwelling and diffusion. Recent incarnations of UD models, including the one we develop in the present paper, allow for exchange with polar surface, deep oceans, and marine biosphere effects.

Ocean general circulation models of (OGCM) carbon cycle [*Sarmiento*, 1986; *Maier-Reimer and Hasselmann*, 1987; *Toggweiler et al.*, 1989a,b; *Maier-Reimer and Bacastow*, 1990; *Bacastow and Maier-Reimer*, 1990] are the most spatially detailed and computationally demanding treatments. They transport carbon within individual basins and around the world's oceans using 3-D current fields generated by the OGCMs; the internal 3-D circulation of oceans is very sparsely mapped by direct measurements. Sophisticated versions include a marine biosphere at a level of detail sufficient to track ecosystem dynamics [*Sarmiento et al.*, 1993]. The potential advantage of using a general circulation model (GCM) is that spatial distributions of both the data and model results can be compared to better constrain model parameters. However, OGCM circulations can depend on numerical integration schemes, model resolution, and input subgrid diffusivities [*Bryan*, 1987]. Their carbon uptake rates are also affected by mixing by unresolved eddies (which must be parameterized) and perhaps by spurious "numerical" diffusion.

A convolution integral of oceancarbon cycle model response to an impulsive injection of carbon can be used to find the response of the ocean carbon sink to changes in atmospheric concentration of CO_2 [*Maier-Reimer and Hasselmann*, 1987]. This approach can also be used to relate CO_2 emissions to atmospheric CO_2 concentrations in both direct and inverse modes [*Harvey*, 1989; *Wigley*, 1991]. However, it is accurate only so long as the governing equations remain nearly linear in the carbon variable. Variations in the oceanic carbonate buffer factor as emissions proceed may limit the applicability to small perturbations from the initial state (see below). In any case, one needs a believable carbon model to start with to generate the impulse response.

Prior comparisons have shown that simpler schematic models can predict present-day carbon uptake at rates comparable to OGCM-based carbon cycle models when normalized to recover the global mean penetration depth (in 1974) of bomb radiocarbon [*Siegenthaler and Sarmiento*, 1993]. These intercomparisons were made in preparation for an upcoming round of Intergovernmental Panel on Climate Change (IPCC) estimates of atmospheric CO_2 buildup projections from various carbon emission scenarios. However, the "simple" ocean carbon cycle models reported were limited to the box-diffusion and HILDA type.

An UD ocean carbon cycle model is one component of the global carbon cycle model shown in Figure 1 and developed in this paper. This model offers computational economy and a parameterization of ocean mixing consistent with the UD ocean energy balance model used by the IPCC [*Hoffert et al.*, 1980; *Bretherton et al.*, 1990] to estimate global greenhouse warming. In what follows, we discuss the physical significance of the model's mixing parameterization and develop the governing equations of an UD model for both stable and radioactive carbon isotopes. The model transport coefficients are calibrated so that model results for ^{14}C in the deep oceans match Geochemical Ocean Sections Study (GEOSECS) data. Model results are found to be consistent with the measured dilution of atmospheric ^{14}C by fossil fuel emissions (the Suess effect) and the penetration of the oceans by bomb-produced ^{14}C. The model is then used to calculate the historical buildup of atmospheric CO_2 including the effects of the land biosphere. *Siegenthaler and*

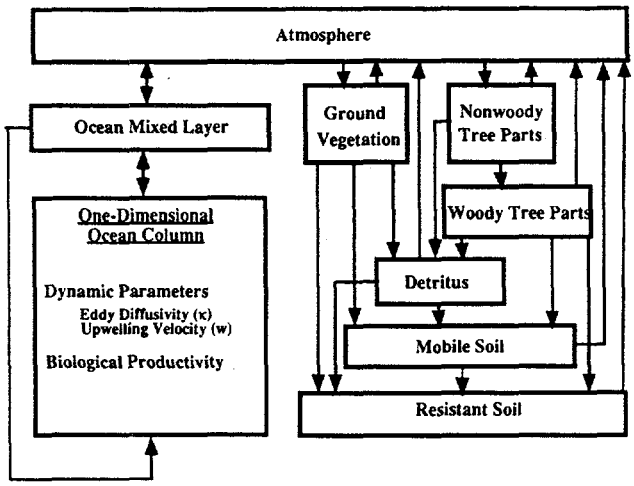


Figure 1. Diagram of the atmosphere-ocean-biosphere system model of the global carbon cycle.

Sarmiento's [1993] ocean carbon cycle uptake intercomparison is then extended to include our model results.

Ocean Chemistry and Transport

Uncertainties in oceanic carbon uptake rates arise mainly from uncertainties in how carbon mass is exchanged between the surface and the deep oceans. Carbon entering the oceans is "buffered" by rapid aqueous reactions to bicarbonate (HCO_3^-) and carbonate ions (CO_3^{-2}) in addition to dissolved carbon dioxide (CO_2). The relative concentrations of these constituents varies with the shifting equilibrium of the carbonate reactions. However, the concentration of dissolved inorganic carbon (DIC), $[CO_2] + [HCO_3^-] + [CO_3^{-2}]$, is conserved. Table 1 indicates the surface ocean DIC concentration and its partitioning for the preindustrial and present atmospheric CO_2 concentrations. Although atmospheric CO_2 (and atmospheric carbon mass) increased by 26.9% since preindustrial times, the change in DIC concentration of surface water is only 2.47%. The relative change in total ocean DIC is too small to measure over the entire ocean volume to construct an accurate estimate of the net ocean carbon uptake. For this reason carbon cycle models are tested by requiring that the model reproduce ^{14}C uptake and distribution. The ratio of these percentage changes is the present buffer factor, $\xi = 26.9/2.47 \approx 10.9$, defined in the appendix.

An effect of the biological productivity of surface-dwelling phytoplankton is the downward transport of carbon through the thermocline by the sinking of undecomposed particulate organic carbon (POC) and calcium carbonate shells (particulate inorganic carbon, PIC). Sinking POC is oxidized on the way down with the help of heterotropic bacteria and other organisms such as zooplankton to release CO_2 which appears as a DIC source. Excluding deepwater-forming regions at high-latitudes, the horizontally averaged oceans upwell at vertical velocity w. The downward biological carbon flux must be balanced in the steady state by upwelling and turbulent carbon fluxes, wC and $\langle w'C' \rangle$, where w' and C' are fluctuations in upwelling and concentration. A major parameter in our model is the eddy diffusivity, $\kappa = \langle w'C' \rangle / (\partial C/\partial z)$. When the surface concentration of DIC rises due to CO_2 transfer from the atmosphere, upward diffusion decreases and there is a net downward flux from the biological carbon rain.

Eddy diffusivity is not isotropic. The vertical component κ of eddy diffusivity is much small than the horizontal components. Order of magnitude calculations [*Hoffert and Flannery*, 1985] suggest that ubiquitous 50 to 100 km eddies [*Semtner and Chervin*, 1992] oriented along slightly inclined constant density surfaces (isopycnals) can produce eddy fluxes by isopycnal mixing whose vertical components are comparable to those of much smaller scale mixing across isopycnals (diapycnal mixing) by breaking internal gravity waves [*Sarmiento et al.*, 1982; *Redi*, 1982; *Kraus*, 1990). Slow vertical mixing, compared with CO_2 gas exchange at the surface, across the stably stratified thermocline is almost certainly the main bottleneck to oceanic fossil fuel carbon uptake. However, it is an open question whether isopycnal mixing, diapycnal mixing, or some combination is the dominant process responsible for vertical mixing.

A fundamental issue for schematic models is that the different values of eddy diffusivity can be found when the model is calibrated to reproduce the distribution of different tracers [*Siegenthaler and Joos*, 1992]. A possible explanation is that the relative contribution of isopycnal versus diapycnal mixing is different for different tracers if they have different patterns of surface penetration. For simplicity, and for consistency with prior ocean climate models, we take both κ and w as depth-independent constants at this point. Apart from mixing and upwelling through the thermocline, we also allow for a downward path for carbon through the polar sea as described below.

Table 1. Preindustrial and Present Concentrations of Dissolved Inorganic Carbon

	Preindustrial	Present	Change	Percent Change
pCO_2, μatm	279.0	354.1	75.1	26.9
$[CO_2]$, μmol kg^{-1}	9.3	11.8	2.5	26.9
$[CO_3^{--}]$, μmol kg^{-1}	215.8	185.5	-30.3	-14.0
$[HCO_3^-]$, μmol kg^{-1}	1711.9	1787.7	75.8	4.4
[DIC], μmol kg^{-1}	1937.0	1985.0	48.0	2.47

DIC = $[CO_2] + [CO_3^{--}] + [HCO_3^-]$ are shown for the surface ocean. Concentrations assumed to be at equilibrium with atmospheric CO_2 are computed for constant total alkalinity $[TA] = [HCO3^-] + 2[CO_3^{--}] + [H_2BO_3^-] + ... + [OH^-] - [H^+]$ = 2252 meqmol kg^{-1}, water temperature T_m = 19.2 °C and salinity S = 34.262‰ using the carbonate chemistry equilibrium constants given by *Peng et al.* [1987].

Model

The globally averaged model for carbon cycle depicted in Figure 1 consists of four reservoirs: the atmosphere, the terrestrial biosphere, the mixed layer of the ocean, and the deep ocean. The atmosphere and mixed layers are modeled as well-mixed reservoirs. The deep ocean, however, is treated as an advective-diffusive medium with a continuous distribution of DIC described by a one-dimensional (1-D) conservation of mass equation characterized by eddy diffusivity κ and upwelling velocity w [*Hoffert et al.*, 1981]. Although the oceanic thermohaline circulation is complex and not well understood [*Broecker and Peng*, 1982; *Toggweiler and Samuels*, 1993; *Caldeira et al.*, 1994], carbon cycle models which seek to investigate carbon storage in the ocean must characterize the effective transport properties of the circulating water masses. The approach we advocate in this paper is to utilize information from oceanographic tracer studies to calibrate transport coefficients for deep-sea transport. In our model the thermohaline circulation is schematically represented by polar deep water formation, with the return flow upwelling through the 1-D water column to the surface ocean from where it is returned, through the polar sea, as bottom water to the bottom of the ocean column thereby completing the thermohaline circulation. Bottom water formation and sinking have been observed in relatively restricted regions of the polar seas, while the return flow has not been

measured and must be inferred from ocean tracers or temperature. The effects of the return flow are expected to spread rapidly in the horizontal direction and are assumed to occur uniformly across the world ocean. This picture of large-scale thermohaline circulation of the world ocean (in which cold, dense seawater sinks at high-latitudes, upwells over the world ocean and returns to polar bottom water forming zones via the meridional current flowing essentially in the surface ocean) was proposed by *Wyrtki* [1961] at least 30 years ago and has roots in physical oceanography going back considerably before that time. Air-sea exchange is modeled by an air sea exchange coefficient in combination with the buffer factor ξ that summarizes the chemical reequilibration of seawater in response to CO_2 variations. The buffer factor is calculated from the set of equations for borate, silicate, phosphate, and carbonate equilibrium chemistry and the temperature-dependent equilibrium constants as given by *Peng et al.* [1987]. An additional carbon source term is added to the deep ocean model to account for the oxidation of particulate organic carbon (POC). The sink of carbon due to new production of POC and the source of carbon due to particulate oxidation at intermediate depths contribute to a 5-10% lower concentration of ΣCO_2 in the surface waters compared to the deep waters. The observed range of the particulate fluxes is quite broad, ranging from 2 to 20 Gt C/yr [*Sundquist*, 1985; *Siegenthaler and Sarmiento*, 1993]. Flux calculations based on models, of course, depend inherently on model structure and assumptions. The marine biogenic flux of carbon in our model is 8.5 Gt C/yr which lies at the middle of the observed range.

The bottom water formation in the higher-latitude belts is quite well known [*Broecker and Peng*, 1982]. Most of the deep bottom water forms in the Weddell Sea in the southern hemisphere. In the UD model, the downwelling flow in a polar sea zone represents the bottom water forming regions of the world oceans. There is an important feedback for the atmospheric CO_2, namely, the direct interaction of atmospheric CO_2 with the deep ocean which is nearly free of excess CO_2. In the past, rapid vertical exchange between the polar surface waters and the deep ocean has been considered in several carbon cycle models [*Crane*, 1982; *Siegenthaler*, 1983; *Kheshgi et al.*, 1991]. In our model, bottom water concentration is controlled by newly introduced bottom water parameter π_c which is the change in the concentration of the bottom water N_b relative to that in the nonpolar region:

$$\pi_c = \frac{N_b - N_{b0}}{N_m - N_{m0}} \quad (1)$$

A value of $\pi_c = 1$ implies that the change in bottom water concentration is the same as the change in mixed layer concentration. However, other effects such as exposure of cold bottom water to the atmosphere could alter bottom water composition. The effect of uncertainties in this parameter can be quantified by using a range of values. When the model is run from 1765 to 1990 by specifying the atmospheric CO_2 concentration record, the model-estimated ocean uptake for the period 1980-1989 is 2.1± 0.1 Gt C/yr, for π_c = 0.5± 0.5. This is compared with the independent estimates [*Sarmiento and Sundquist*, 1992] of 2.0 ± 0.8 Gt C/yr for the same period. We use π_c=0.5 for total carbon and ^{14}C in our model results which follow. It is worth noting that the ratio of $^{14}C/^{12}C$ is insensitive to the value of π_c (the model-estimated change in the bomb ^{14}C inventories and penetration depth for the range of π_c =0.0-1.0 is <1%) and so measured ratios of $^{14}C/^{12}C$ cannot be used to choose the parameter π_c.

To estimate terrestrial biospheric fluxes, a six-box globally aggregated terrestrial biosphere submodel is coupled to the atmosphere box. The six boxes represent ground vegetation, nonwoody tree parts, woody tree parts, detritus, mobile soil (turnover time 70 years), and resistant soil (turnover time 500 years). The mass of carbon contained in the different reservoirs and the rate of exchange between them have been based on the analysis by *Harvey* [1989]. The model estimate of the global emission rate of CO_2 from changes in land use, from 1765 (preindustrial time) to 1990, is determined by the deconvolution of the atmospheric CO_2 concentration history (H.S. Kheshgi, et al., Accounting for the missing carbon sink with the CO_2 fertilization factor, submitted to *Climate Change*, 1994; hereinafter refereed to as Kheshgi et al., submitted manuscript, 1994). The model-estimated land use emissions are well within the error bounds (±1 Gt C/yr) of estimates of land use emissions summarized by *Schimel et al.* [1994].

The rate of photosynthesis by terrestrial biota is thought to be stimulated by increasing atmospheric carbon dioxide concentration. The increase in the rate of photosynthesis, relative to preindustrial times, is modeled to be proportional to the logarithm of the relative increase in atmospheric CO_2 concentration from its preindustrial value of 278 parts per million (ppm) [*Bacastow and Keeling*, 1973]. The proportionality constant β, known as the CO_2 fertilization factor, is chosen to be 0.42. The magnitude of the modeled biospheric sink is dependent on the chosen value of the CO_2 fertilization factor [*Harvey*, 1989; *Wigley*, 1993]. By varying the ambient CO_2 partial pressure for individual plants in controlled experiments under ideal growing conditions, a CO_2 fertilization effect has been demonstrated and found to lead to CO_2 fertilization factors ranging from 0.15 to 0.60 [*Kohlmaier et al.*, 1991]. Nutrient limitation and community competition for resources may, however, severely diminish ecosystem response to changing atmospheric CO_2 partial pressure [*Bazzaz and Fajer*, 1992]. The value of β = 0.42, however, is not prescribed by CO_2 fertilization experiments. The modeled magnitude of the CO_2 fertilization effect is chosen here to lead to a reconstruction of past carbon cycle, described by Kheshgi et al., submitted manuscript [1994], that matches the land use emission estimate [*Schimel et al.*, 1994] of 16.0 GtC for the decade 1980-1989. The rate coefficients for exchange to and from terrestrial biosphere boxes are temperature dependent according to an Arrhenius law.

In the appendix we describe the model equations for the ocean and atmosphere reservoirs of CO_2 and ^{14}C. These equations have been solved using a fourth-order Runga-Kutta method. A detailed description of various terrestrial components of the model is given by Kheshgi et al., submitted manuscript [1994].

Model Calibration to Preindustrial ^{14}C

There are two dynamic parameters in our model (the eddy diffusivity κ, and upwelling velocity w) which are calibrated by

matching the natural ^{14}C distribution in the deep ocean. Carbon 14 concentrations in this study are expressed as the deviations of $^{14}C/^{12}C$ ratio in parts per thousand (ppt) from the standard ratio that is expressed as $\Delta^{14}C$ [*Stuiver and Pollach*, 1977]. For the model calculations we use an arbitrary scale for $^{14}C/^{12}C$ ratio such that the preindustrial atmospheric ^{14}C ratio is 1 unit. When presenting the model results, we convert our arbitrary units into standard $\Delta^{14}C$ units as follows:

$$\Delta\,^{14}C = (\text{model units-1}) \times 10^3 \qquad (2)$$

Therefore, in the preindustrial atmosphere, $\Delta^{14}C$ was equal to 0‰.

In the steady state, the rate of changes in atmosphere, mixed layer, and deep ocean concentrations are zero. Therefore the ocean ^{14}C ratio can be calculated by fixing the surface ocean and the polar sea values of total inorganic carbon and ^{14}C ratio. The preindustrial DIC concentrations in the mixed layer and the polar boxes of 2.03 mol/m^3 and the ^{14}C ratios in the corresponding boxes of 0.95 (-50‰) and 0.85 (-150‰) were estimated by *Broecker et al.* [1985] and *Broecker* [1963] and are used here to generate our initial conditions. The bomb-produced ^{14}C did not penetrate below the thermocline at the time of GEOSECS survey. Therefore the vertical profile of preindustrial ^{14}C ratio below 1000 m is well approximated by the global-average values estimated by *Shaffer and Sarmiento* [1994] from GEOSECS data. The values of κ and w required to match the preindustrial, vertical profile of ^{14}C are 4700 m^2/yr and 3.5 m/yr. The observed vertical profile of DIC for the oceans has also been estimated from the GEOSECS data by *Takahashi et al.* [1981]. The model estimated deep ocean profiles for ^{14}C concentration and total inorganic carbon are compared in Figures 2 and 3 with the GEOSECS-derived data. The $\Delta^{14}C$ value decreases across the main thermohaline, reaching a minimum of about -190‰ at a depth of 2.5 km; below this depth it rises again reaching the value of -178‰ at the ocean floor. The model-estimated mean ^{14}C ratio of -164‰ is in good agreement with the estimated global average value of -160‰ [*Oeschger et al.*, 1975]. The steady state concentration of DIC increases as one descends from the surface to reach a maximum and then decreases toward the ocean floor. The model-estimated mean DIC concentration of 2.30 mol/m^3 is close to the GEOSECS-derived value of 2.31 mol/m^3 [*Takahashi et al.*, 1981].

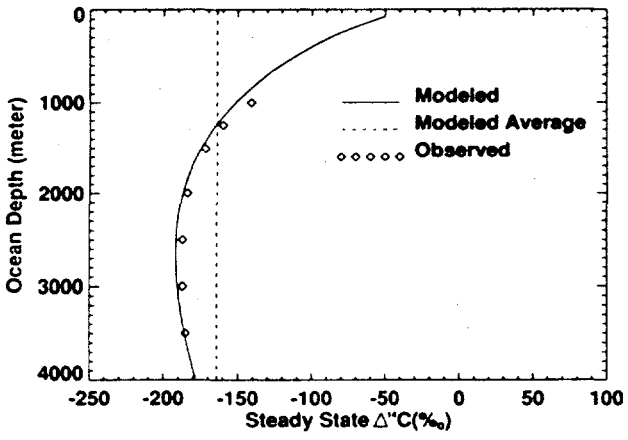


Figure 2. Comparison of the model simulated steady state ^{14}C (in per mil) in the ocean with the observed data. The estimated observed global mean values are taken from *Shaffer and Sarmiento* [1994].

The gas exchange flux was estimated (for the values of κ and w calibrated to match deep ocean ^{14}C) from the ^{14}C ratio of the surface reservoirs. The resulting global mean gas exchange rate at preindustrial CO_2 concentration of 278 ppm is 17.8 mol m^{-2} yr^{-1}, in close agreement with the *Sarmiento et al.*'s [1992] estimated gas exchange rate values 17.9 mol m^{-2} yr^{-1} at 275 ppm and with the value of 17.5 mol m^{-2} yr^{-1} at 280 ppm obtained by *Broecker et al.* [1985] from observed data. The bomb radiocarbon study of *Toggweiler et al.* [1989b] used a slightly lower value of the gas exchange rate of 16.6 mol m^{-2} yr^{-1}. Our flux value is also higher than the value of 15.2 mol m^{-2} yr^{-1} obtained by *Siegenthaler and Joos* [1992]. The difference in values could be due to the fact that we have used the actual profile of total inorganic carbon in the deep ocean, while *Siegenthaler and Joos* [1992] assumed constant concentration profile for the whole ocean equal to the surface ocean value, which is about 10% less than the mean ocean value.

Suess Effect

Fossil fuels are depleted of ^{14}C. The input of CO_2 from these sources lowers the atmospheric ratio of ^{14}C to total C by dilution. This effect was first determined by *Suess* [1955]. On the other hand, forest and soil carbon have ^{14}C ratio nearly the same as the atmospheric ratio; hence input or exchange of CO_2 from these sources has little effect on the atmospheric ^{14}C ratio if atmospheric ^{14}C is not changing. *Stuiver and Quay* [1980] suggest that changes in atmospheric ^{14}C could also be caused by changes in the production rate of ^{14}C which are correlated to the sunspot index. However, because of the uncertainty in the variation of production rate of ^{14}C and its negligible effect on ^{14}C ratio [*Bacastow and Keeling*, 1973] over the timescale of the Suess effect, we do not account for this effect in our calculation and thus assume constant ^{14}C production. The Suess effect is apparent in measured atmospheric ^{14}C before the 1950s, at which time the ^{14}C ratio increased drastically due to emissions from nuclear bomb detonations. High-precision tree ring measurements of atmospheric ^{14}C levels between 1840 and 1950 show a decrease of about 24‰ as shown in Figure 4 [*Stuiver and Quay*, 1981]. In another analysis of tree ring data, *Keeling* [1973] found a slightly higher Suess effect of -25‰ over the same time period.

Carbon cycle models have been used to predict the size of the atmospheric Suess effect up to 1950. *Bacastow and Keeling* [1973] estimated a Suess effect of about -23 ± 4‰ using their one-dimensional (1-D) model, while *Oeschger et al.* [1975] estimated a value of -20‰. Our UD model estimates an atmospheric ^{14}C change of -25‰. In Figure 4, the Suess effect estimated by our UD model over the time period from 1840 to 1950 is seen to be consistent with tree ring data.

The surface ocean also shows a Suess effect. *Druffel and Linick* [1978] and *Druffel and Suess* [1983] measured a change of -9 ± 3‰ in the surface ocean ^{14}C ratios over the period 1850-1950. This estimate is based on the average of the radiocarbon measurements for the Pacific and the Atlantic corals which

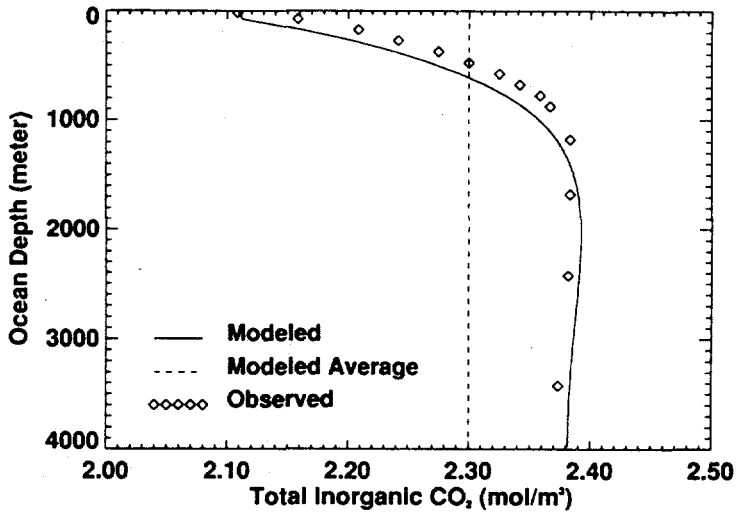


Figure 3. Comparison of the model estimated total inorganic carbon (in moles per cubic meter) in 1974 with observed ocean data. The observed data are a global average of the GEOSECS data which were taken over the period 1972-1978 [*Takahashi et al.*, 1981].

includes, for example, Florida coral (-11‰), Belize coral (-12‰), and Galapagos coral (-6‰). The UD model predicts a mixed layer Suess effect of -9‰ over the same period, again consistent with observations.

Bomb Produced ^{14}C

Bomb tests carried out during the period 1954-1963 produced about 700 x 10^{26} ^{14}C atoms [*Broecker and Peng*, 1982]. The bomb ^{14}C tracer interrogates mixing processes capable of removing ^{14}C (or CO_2) from the atmosphere on the decade timescale. It has been hypothesized [*Broecker et al.*, 1980] that models capable of predicting bomb ^{14}C distributions will also be capable of predicting the uptake of anthropogenic CO_2 which also has a similar characteristic timescale (≈ 20 years). This is in contrast to the preindustrial distribution of ^{14}C, which was used to calibrate the ocean mixing parameters κ and w, and is sensitive to slow mixing processes (≈ 500 years) approaching the half-life time of radiocarbon (5730 years). It is therefore important to check whether our model can also reproduce the measured distribution of the bomb ^{14}C. This check tests whether the ratio of air-sea exchange to vertical mixing resistance just beneath the mixed layer is of the right magnitude. Since the production data of ^{14}C produced by nuclear weapons tests are not precisely known, the average observed atmospheric concentrations, shown in Figure 5, are prescribed from 1950 to 1975 in order to estimate the uptake of bomb ^{14}C by the ocean. Figure 6 shows the observed- and model-estimated ^{14}C distributions in the ocean for the year 1974 which clearly show the effects of bomb-produced ^{14}C. The observed values were derived from GEOSECS data by *Shaffer and Sarmiento* [1994]. Although the model slightly underestimates the bomb ^{14}C concentration in the upper ocean, elsewhere the model is able to reproduced the observed bomb ^{14}C distribution.

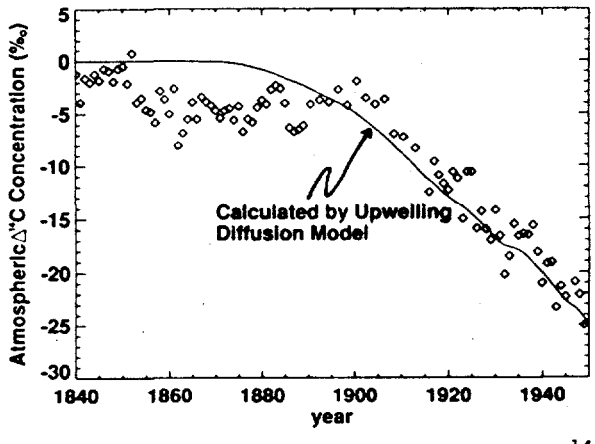


Figure 4. Comparison of the model estimated atmospheric ^{14}C decrease caused by the dilution of ^{14}C by fossil fuel CO_2 (i.e., the Suess effect) with the tree rings measurements of *Stuiver and Quay* [1981] from 1840 to 1950.

The observed mean concentration of bomb ^{14}C in surface water, defined as the difference between the postbomb $\Delta^{14}C$ and the prebomb $\Delta^{14}C$, was 160±15‰ [*Broecker et al.*, 1985]. The UD model predicts 172‰ for the period 1950-1974 which lies within the limit of uncertainties.

The average oceanic inventory of bomb ^{14}C ($\Sigma^{14}C$) defined as the excess ^{14}C amount over the prebomb level, is determined [*Broecker et al.*, 1985] by integrating the difference between the prebomb and postbomb vertical profiles as

$$\Sigma^{14}C = \int_0^{4000m} \left[\Sigma CO_2(\Delta^{14}C - \Delta^{14}C_{1950})\right] dz \qquad (3)$$

where $\Delta^{14}C_{1950}$ and $\Delta^{14}C$ are the pre and post-bomb ocean surface concentration and ΣCO_2 is the DIC concentration. The penetration depth Z is

$$Z = \int_0^{4000m} (\Delta^{14}C - \Delta^{14}C_{1950}) dz \Big/ (\Delta^{14}C - \Delta^{14}C_{1950})\Big|_{z=0} \qquad (4)$$

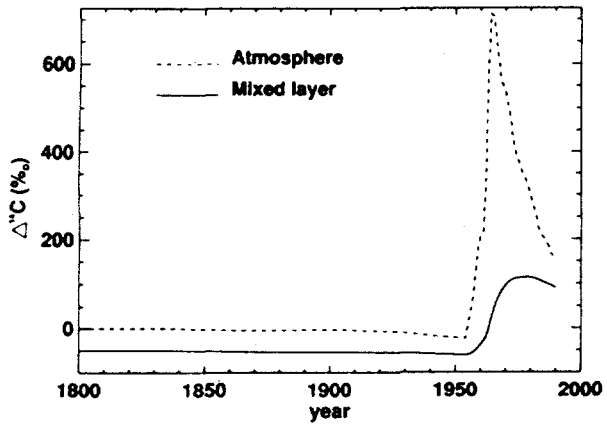


Figure 5. Model estimated ^{14}C concentration (per mil) in mixed layer from 1800 to 1990. The atmospheric ^{14}C concentration changes until 1950, due to dilution by fossil fuel CO_2, are calculated with the carbon cycle model. Thereafter, concentrations are prescribed according to observation [*Tans*, 1981].

An ocean average model-estimated DIC concentration of 2.1 mol/m^3 is used for ΣCO_2 when evaluating (3) and (4). The model-estimated bomb CO_2 inventories and penetration depths from 1955 to 1975 are shown in Figures 7 and 8. For 1974 (time of the GEOSECS survey) the model-estimated bomb ^{14}C inventory and the penetration depth are 8.3 x 10^{13} atoms/m^2 and 308 m, respectively. From the GEOSECS data, *Broecker et al.* [1985] estimated the total-ocean inventory to be about 2.9x10^{28} atoms. If we divide this value by the world's ocean area (3.6 x 10^{14} m^2), the average bomb ^{14}C inventory is 8.0±0.5 x 10^{13} atoms/m^2 and the resulting penetration depth is 328±20 m. Our model-estimated value for bomb ^{14}C inventory is about 4% higher as compared to observed data, whereas the estimated penetration depth is about 7% lower. There could be two reasons for these minor differences. First, the total CO_2 concentration value used in our calculation is higher than the value of 2.15 mol/m^3 [*Toggweiler et al.*, 1989b] used by *Broecker et al.* [1985]. Second, the prenuclear (1950) surface water concentration in our calculations is lower (-59‰) than that (-50‰) used by *Broecker et al.* [1985]. This is due to the fact that we start the calculation from preindustrial time with prescribed biospheric and fossil emissions, which results in a Suess effect of -25‰ in surface water, while *Broecker et al.* [1985] start the calculations in 1950 with the preindustrial atmospheric and ocean ^{14}C ratios and ΣCO_2.

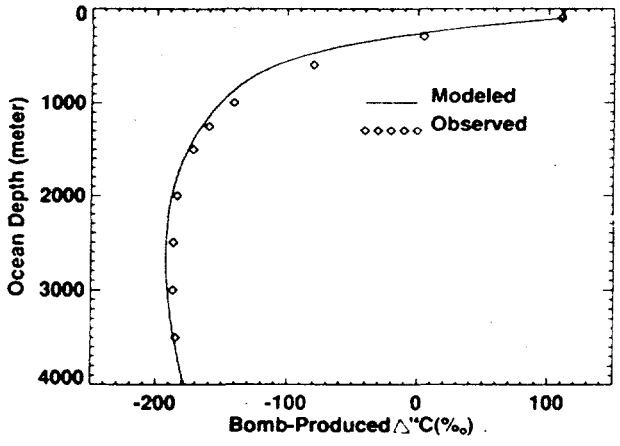


Figure 6. Comparison of the model-estimated bomb-produced ^{14}C (in per mil) in the deep ocean with the observed data. The observed values, estimated from the GEOSECS data, are taken from *Shaffer and Sarmiento* [1994].

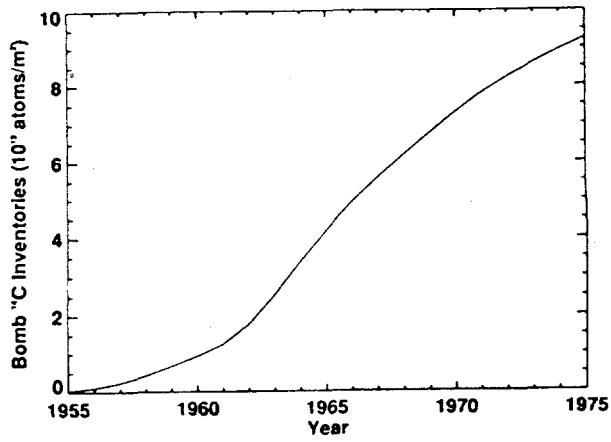


Figure 7. Model-estimated vertically integrated inventories of bomb ^{14}C (10^{13} atoms/m^2) from 1955 to 1975.

Model Intercomparison

The results of the UD model are also compared in Table 2 with those of a pure diffusion model [*Siegenthaler*, 1983], high-latitude exchange/interior diffusion advection (HILDA) [*Siegenthaler and Joos*, 1992], and ocean general circulation models (OGCMs) [*Toggweiler et al.*, 1989b; and M. Heimann, personal communication, 1993].

Many of the simple models have been calibrated using the distribution of ^{14}C tracer. *Siegenthaler* [1983] calibrated their model parameters by means of the postbomb ^{14}C distribution in the deep sea; however, the estimated mean deep ocean value for

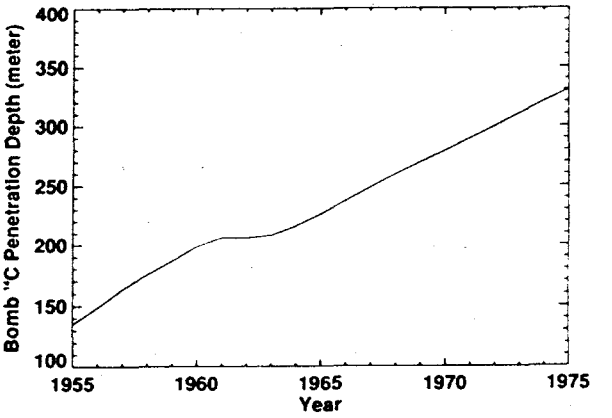


Figure 8. Model-estimated mean penetration depth (in meters) for excess ^{14}C from 1955 to 1975.

the preindustrial time was about 30% higher than the observed values (Table 2). Using the HILDA model, *Siegenthaler and Joos* [1992] found that to produce natural as well as bomb ^{14}C distribution in the ocean, the vertical diffusivity κ in their model must be reduced with depth. In contrast, the UD model is able to reproduce both the natural and bomb ^{14}C distribution in ocean with the constant diffusivity, as indicated in previous sections. It should be noted that the structure of the UD model differs from the HILDA model in that the HILDA model requires, in addition to upwelling velocity (w) and diffusivity (κ), two more tunable parameters: water exchange between the high-latitude surface and deep ocean boxes (u), and between the high- and low-latitude deep ocean boxes (q). Also, the upwelling velocity used in the HILDA model is very small (w = 0.7 m/yr) as compared to our value (w = 3.5 m/yr). Our higher value is in closer agreement with the upwelling velocity derived from ocean overturning rates (w = 4 m/yr) [*Gordon and Taylor*, 1975]. It must also be pointed out that contrary to *Siegenthaler and Joos* [1992], we also take into account the downward transport of ^{14}C by the rain of organic particles produced by marine organisms. Neglecting this biological effect can introduce errors up to 10% in the steady state profile of ^{14}C ratios [*Fiadeiro*, 1982]. Even so, Table 2 shows that the changes predicted by the HILDA and UD models are quite close to the observed values.

Table 2 also shows that the results of the two OGCMs presented do not lie within the range of uncertainty of the observed data, as did the UD results. The *Toggweiler et al.* [1989b] results are determined from the Geophysical Fluid Dynamics Laboratory (GFDL) primitive equation OGCM [*Bryan*, 1987]. M. Heimann's (personal communication, 1993) results are based on the Hamburg Large Scale Geostrophic OGCM [*Maier-Reimer et al.*, 1993]. While the GFDL model predicts a mean

Table 2. Comparison of Upwelling-Diffusion Model (UD) Results with Observations and Other Different Model Results

					OGCM	
	Observation	UD Model, This Study	HILDA Model, *Siegenthaler and Joos* [1992][a]	Diffusion Model, *Siegenthaler* [1983]	*Toggweiler et al.*[1989a,b][b]	M. Heimann, (personal communication, 1994)
Preindustrial (1765) ocean surface $\Delta^{14}C$ ‰	-49 ± 3[c]	−50	−55	−46	—	—
Prebomb (1950) ocean surface $\Delta^{14}C$ ‰	-58 ± 3[c]	−59	−62	−53	−50	—
Bomb $\Delta^{14}C$ in ocean surface ‰[d]	160 ± 15[e]	171	155	163	160	206
Mean ^{14}C for the deep-sea $\Delta^{14}C$, ‰	−160[f]	−163	—	−110	—	—
Total CO_2 concentration ΣC, mol/m^3[g]	2.15[h]	2.18	2.13	2.05	2.15	—
Bomb ^{14}C inventories, 10^{13} atoms/m^2	8.00 ± 0.5[e]	8.30	7.95	8.42	6.90	9.27
Penetration depth, m	328 ± 20[i]	308	338	355	283	320
Oceanic CO_2 uptake, 1980s, Gt C/yr	2.0 ± 0.8[j]	2.10	2.15	3.33	1.67	1.45

a) Values for HILDA model are area weighted mean from low- and high-latitude belts.

b) Assume prebomb atmosphere(1950) in steady state atmosphere.

c) Estimates are based on the Pacific and Atlantic corals [*Druffel and Suess*, 1983].

d) Defined as the difference between the $\Delta^{14}C$ at the time of GEOSECS and the prebomb $\Delta^{14}C$.

e) *Broecker et al.* [1985].

f) Weighted mean of Atlantic and Pacific data [*Oeschger et al.*, 1975].

g) Average value for that part of the water column which is contaminated with bomb ^{14}C.

h) Estimated by *Toggweiler et al.* [1989a,b] from the GEOSECS observations.

i) Estimated from the formula of *Broecker et al.* [1985].

j) Estimate [*Sarmiento and Sundquist*, 1992] of ocean carbon uptake based on measurements outlined by *Tans et al.* [1990].

surface water bomb ^{14}C concentration of 160‰, in agreement with the GEOSECS record, the prediction of the bomb ^{14}C inventory and penetration depth are too low as compared to the observations. On the other hand, the Hamburg model predicts very high values of bomb ^{14}C surface concentrations, about 23% higher than the observed value. The bomb ^{14}C inventory is higher by about 16% than the observed value, even though it does reproduce the observed bomb ^{14}C inventory. These results suggest that the current published OGCMs do not correctly simulate the atmospheric-ocean exchange and the vertical mixing in the upper ocean. Therefore such models in their current state should not be used to predict either current or future uptake of CO_2 by the oceans.

The models listed in Table 2 yield an ocean CO_2 uptake rate for the 1980s of between 1.45 and 3.33 Gt C/yr. Both GCMs [*Toggweiler et al.*, 1989a; *Maier-Reimer et al.*, 1993] underpredict the ocean CO_2 uptake rate, while the diffusion model [*Siegenthaler*, 1983] overpredicts the ocean CO_2 uptake. The UD model and the HILDA model results lie near the center of the uncertainty band of *Sarmiento and Sundquist's* [1992] estimate for the current ocean uptake of 2.0± 0.8 Gt C/yr.

Radiocarbon Estimates of Terrestrial Biosphere

The magnitude and the timing of the terrestrial ecosystems response to changes in the land use and climate are a large source of uncertainty in global carbon cycle models. An understanding of these uncertainties requires knowledge of both the inventory of carbon in the biosphere (including soils) and the turnover rate of carbon in the carbon reservoirs, in particular, the soil reservoir. An estimated 1300 to 1500 Gt C are stored globally as organic matter in the upper meter of mineral soils [*Schlesinger*, 1991] which is roughly twice the estimated storage of carbon in atmosphere or in the living biosphere [*Houghton et al.*, 1990]. Attempts have been made to model the turnover times of soil carbon reservoirs with the aid of soil ^{14}C measurements [*Harrison et al.*, 1993; *Trumbore*, 1993]. In this paper, however, we do not attempt to estimate the soil carbon turnover rates from ^{14}C data, realizing that there are not only the uncertainties in the observed data but also in the soil carbon dynamics [*Stevenson and Elliott*, 1990]. Instead, we use the turnover times reported by *Harvey* [1989].

To estimate the terrestrial biospheric uptake of radiocarbon, we use a six-reservoir terrestrial model, as shown in Figure 1. Turnover times for the terrestrial reservoirs range from approximately 1 year for nonwoody tree parts to 20 years for woody tree parts and 70 and 500 years for mobile (fast cycling) and resistant (slow cycling) soil reservoirs (Table 3). We assume a fertilization factor ß =0.42. We also take into account a temperature-respiration feedback in our calculations. However, the global mean temperature over the time period considered here is too small to lead to an appreciable change in the modeled respiration rates. In steady state the model-estimated ^{14}C concentrations of terrestrial reservoirs are -36‰ for ground vegetation, -36‰ for nonwoody tree parts, -38‰ for woody tree parts, -37‰ for detritus, -45‰ for mobile soil, and -96‰ for resistant soil (see Table 3). Changes in atmospheric ^{14}C concentrations have an impact first on terrestrial carbon reservoirs with short turnover times and have little influence on terrestrial carbon reservoirs with long turnover times. We have studied two ^{14}C effects: ^{14}C dilution (Suess effect), and ^{14}C enhancement (bomb-produced). The model estimated dilution of atmospheric ^{14}C caused by fossil fuel CO_2 from 1765 to 1950 was about -25 ‰ (Table 3) and the observed ^{14}C increase due to nuclear bomb tests from 1950 to 1964 was about +737 ‰. In 1964, nuclear tests were banned and since then the atmospheric ^{14}C concentration has decreased exponentially (Figure 5). The carbon reservoir having a shortest turnover time, 0.7 years, follows the atmospheric signal; ^{14}C concentration reduction for this reservoir is 20‰ from 1765 to 1950 and increase is 622‰ from 1950 to 1964 (Table 3). In contrast, the carbon reservoir

Table 3. Turnover Times and ^{14}C Concentrations for Terrestrial Reservoirs for Preindustrial (1765), Prebomb (1950), and Postbomb (1964) Years

Reservoir	Turnover Time, years	^{14}C Concentration, ‰				
		1765	1950	1964	1950-1765	1964-1950
Atmosphere		0	-25	712	-25	737
Ground vegetation	2.0	-36	-56	398	-20	454
Nonwoody tree parts	0.7	-36	-56	566	-20	622
Woody tree parts	20.0	-38	-53	23	-15	76
Detritus	1.8	-37	-55	227	-18	82
Mobile soil	70.0	-45	-51	-43	-6	8
Resistant soil	500.0	-96	-96	-95	0	1

Also given are the ^{14}C concentration differences for the prebomb and preindustrial time period (1950-1765) and post and prebomb time period (1964-1950).

having longest turnover time, 500 years, shows the least changes. The resistant soil reservoir experiences almost no appreciable change from 1765 to 1950 and only 1‰ increase from 1950 to 1964 (Table 3).

The global terrestrial biospheric uptake of bomb radiocarbon can not be compared with observations, because there are no observational estimates available in the literature. However, we have compared our model results with other recently published model estimates. *Joos* [1994] has estimated the terrestrial biospheric bomb radiocarbon inventories for the period 1965 to 1989 to be $75 \pm 30 \times 10^{26}$ atoms. This estimate is simply based on the average of four carbon cycle model results. The individual estimates of four models are 60×10^{26} [*Hesshaimer et al.*, 1994], 45×10^{26} [*Broecker et al.*, 1985], 89×10^{26} [*Siegenthaler and Oeschger*, 1987], and 99×10^{26} atoms [*Siegenthaler and Joos*, 1992], respectively. Our model estimated bomb radiocarbon inventory for the terrestrial biosphere reservoir for the same period (1965-1989) is 80×10^{26} atoms which is well within the range of *Joos'* [1994] estimates.

Temperature Change

We now examine whether the model dynamic parameters κ and w that are derived from the ^{14}C calibration are able to simulate vertical temperature distribution. Figure 9 shows that the UD model does not accurately reproduce the *Levitus* [1982] horizontal average distribution of potential temperature in the deep ocean. The results from the prognostic OGCM [*Toggweiler et al.*, 1989a] and HILDA [*Siegenthaler and Joos*, 1992] models are also shown in Figure 9. The HILDA model yields temperatures that are too high overall. The results found by both our UD model and by *Toggweiler et al.* [1989a] in their OGCM are remarkably similar. We find that both models give temperatures near the thermocline which are too warm, although the bottom water and mixed layer temperatures are in agreement with observations. It seems that the apparent upper ocean diffusivity for the uptake of the bomb radiocarbon is considerably more intense than that needed to match the steady state vertical temperature profile. This is consistent with the hypothesis that tracers, such as ^{14}C, with concentration contours crossing isopycnal contours will see enhanced vertical mixing relative to temperature which has isotherms virtually parallel to isopycnals. Therefore we reach a conclusion similar to that of *Siegenthaler and Joos* [1992] in that the UD model is not able to predict the distribution of both temperature and ^{14}C with one set of dynamic parameters.

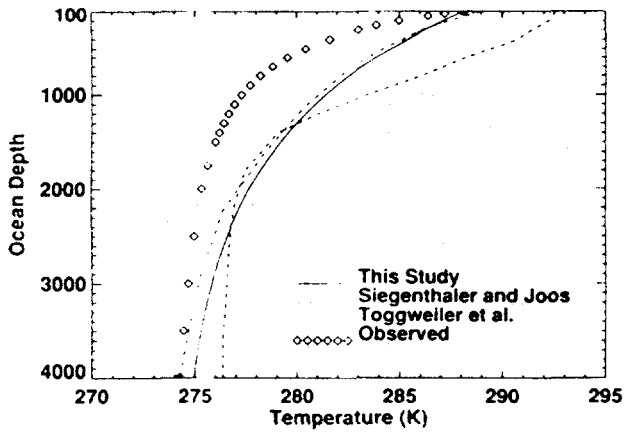


Figure 9. Comparison of the depth dependence of equilibrium mean deep ocean temperature compared between observations [*Levitus*, 1982], this model result, OGCM carbon cycle model results [*Toggweiler et al.*, 1989a], and a HILDA model result [*Siegenthaler and Joos*, 1992].

Concluding Discussion

A carbon cycle model describing the exchange of carbon among the atmosphere, ocean, and terrestrial biosphere is found to be consistent with measured values (within measurement error) of the prebomb rise in ^{14}C in the atmosphere and the mixed layer due to the Suess effect, the bomb ^{14}C in the mixed layer, the bomb ^{14}C penetration depth, the bomb ^{14}C ocean inventory, and the vertical distribution of total carbon. This is the first study to recognize that a simple upwelling-diffusion model for the deep ocean can reproduce all of these features by calibration of only two parameters, κ and w. The thermocline depth scale κ/w = 1343 m found by calibration to steady state ^{14}C is considerably deeper than that required to match the steady vertical temperature profile (500 m); this is consistent with the hypothesis that isopycnal mixing, which is much more rapid than diapycnal mixing, has a stronger effect on ^{14}C than on temperature since isopycnals are nearly isothermal [*Hoffert and Flannery*, 1985]. Results are compared to those of one other schematic carbon cycle model as well as those of two ocean general circulation models which have also attempted to reproduce the ^{14}C records. The GCMs' results lie outside the range of uncertainties of the records, while the schematic model (HILDA) proposed by *Siegenthaler and Joos* [1992] succeeded. The HILDA model, however, required the use of an extremely low upwelling velocity as well as a vertical eddy diffusivity that decreases with depth in contradiction to our understanding of bottom water formation and ocean mixing [*Gordon and Taylor*, 1975; *Gargett and Holloway*, 1984]. The UD and HILDA models use different approximations to represent the effects of bottom water formation and rapid mixing in the high-latitude oceans.

We introduce the parameter π_C, which is the change in the concentration of the bottom water relative to that in the mixed layer. Model ^{14}C results are not sensitive to this parameter, although the ultimate response of the system to emissions is sensitive to π_C. We choose the value of $\pi_C = 0.5$ from independent estimates of ocean carbon uptake. Thus calibrated, this model for the global carbon cycle is intended for use in predicting the rise of atmospheric CO_2 due to emissions; future studies will address such analyses.

Appendix

The rate of change of carbon in the model atmosphere is governed by

$$\frac{dN_a}{dt} = k_a \left(\frac{N_{ao}}{N_{mo}} \right) (N_{mo} + \xi n_m) - k_a n_a + F_e + B_e \qquad \text{(A1)}$$

where the mass of carbon (in gigatons carbon) in the atmosphere is $N_a = N_{ao} + n_a$ and $N_{ao} = 590$ Gt C is the preindustrial atmospheric carbon mass. The mass of dissolved inorganic carbon (Gt C) in the mixed layer is $N_m = N_{mo} + n_m$ and $N_{mo} = 676$ Gt C is the preindustrial mixed layer carbon mass. The atmospheric rate of exchange is $k_a = 0.11$ yr^{-1}. The mass rate of fossil fuel emissions to the atmosphere is F_e. The net terrestrial mass rate of emissions to the atmosphere is B_e. The buffering action of seawater is taken into account by multiplying the excess CO_2 content of the mixed layer (n_m) by buffer factor ξ, where

$$\xi = \frac{(P_m - P_{m0})/P_{m0}}{(N_m - N_{m0})/N_{m0}} \tag{A2}$$

P_m is the equilibrium CO_2 partial pressure of seawater in the mixed layer, and $P_{m\,0} = 769.79$ μatm is preindustrial value. Partial pressure P_m is calculated from the set of chemical equilibrium equations for borate (Tb), silicate (Tsi), phosphate (Tp), total alkalinity (Talk), silicate (*s*), and carbonate. Concentrations Tb, Tsi, Tp, Talk, and *s* are assumed to remain constant in the mixed layer. Following *Peng et al.* [1987] these concentrations are taken to be Tb=409.07 μ*M* kg^{-1}, Tsi=46.5 μ*M* kg^{-1}, Tp=1.43 μ*M* kg^{-1}, Talk=2333 μ*M* kg^{-1}, and s=34.26‰, respectively.

The rate of change of carbon in the model mixed layer is governed by

$$\frac{dN_m}{dt} = -k_a\left(\frac{N_{ao}}{N_{mo}}\right)(N_{mo} + \xi n_m) + k_a N_a + A\Omega\kappa\left.\frac{\partial N_d}{\partial z}\right|_{z=0} + A\Omega w[N_d(0) - N_b] - F_g \tag{A3}$$

where the surface area of the oceans is $A = 3.62 \times 10^{12}$ m^2. The vertical eddy diffusivity in the deep ocean column $\kappa = 4700$ m^2/yr and upwelling velocity $w = 3.5$ m/yr are model parameters calibrated by matching the depth variation of preindustrial ^{14}C. The conversion factor $\Omega = 1.2 \times 10^{-14}$ Gt C/mol C. In the above equations, F_g is the net marine biogenic flux of carbon leaving the mixed layer, i.e., the new production, and is taken to be 8.5 Gt C/yr [*Bacastow and Maier-Reimer*, 1991; *Najjar et al.*, 1992].

The rate of change of carbon in the model deep ocean column is governed by

$$\frac{\partial N_d}{\partial t} = \kappa\frac{\partial^2 N_d}{\partial^2 z} + w\frac{\partial N_d}{\partial z} + J \tag{A4}$$

where N_d is the concentration of dissolved inorganic carbon (in moles C per cubic meter) and J is the source of carbon by from respiration of particulate organic carbon. The strength of the source $J(z)$ is assumed to decrease exponentially with depth with an exponential depth scale of 750 m [*Volk and Hoffert*, 1985], which roughly matches the scattered data on organic particulate flux derived from sediment trap measurements which best fits a profile proportional to $1/z$ [*Berger et al.*, 1987]. This partial differential equation has the boundary conditions

$$N_d = \frac{N_m}{A\Omega h_m} \qquad \text{at } z = h_m \tag{A5}$$

$$\kappa\frac{\partial N_d}{\partial z} + wN_d = wN_b \qquad \text{at } z = h_d \tag{A6}$$

where $h_d = 4{,}000$ m is the mean depth of the ocean and $h_m = 75$ m is the mean depth of the mixed layer. The bottom water concentration N_b is determined by (1). The depth below the mixed layer is z.

The amount of ^{14}C is given by the product of carbon mass or concentration N and the ^{14}C ratio R. The equation for the mass balance of atmospheric ^{14}C is

$$\frac{dN_aR_a}{dt} = k_a\left(\frac{N_{ao}}{N_{mo}}\right)(N_{mo} + \xi n_m)\alpha_{ma}R_m - k_aN_a\alpha_{am}R_a + \bar{Q}_p + B_{er} - \lambda N_aR_a \tag{A7}$$

where R_a, R_m and R_b are the mole ratios of ^{14}C to carbon in the atmosphere, mixed layer, and terrestrial biospheric emissions. Preindustrial values for R_a and R_m are 0‰ and -50‰. Here α_{am} and α_{ma} are the ^{14}C fractionation factors for CO_2 uptake by the surface ocean and CO_2 evasion from the surface ocean. The factors α_{am} and α_{ma}, according to *Keeling* [1973], are 0.972, and 0.955. Note that no term appears for fossil fuel emissions since fossil fuels are depleted in ^{14}C. The decay constant λ of ^{14}C is 1/(8267 years). The mean production rate $\bar{Q}_p$ of ^{14}C from cosmic ray flux is taken to be equal to 2 atom cm^{-2} s-1. The transfer of ^{14}C from the terrestrial biosphere to the atmosphere is B_{er}, where

$$B_{er} = R_a\alpha_{ab}(-F_{at,gv} - F_{at,nw}) + \sum_{i=1}^{6}(R_{bi}F_{i,at} + R_{bi}Flu_i) \tag{A8}$$

and where R_{bi} is the ^{14}C ratio of the ith terrestrial reservoir. The preindustrial values of R_{bi} are given in Table 3. Flu_i is the net land use emission contribution of the ith biosphere reservoir. $F_{i,at}$ is the flow of carbon from terrestrial reservoir i to atmosphere. Similarly, $F_{at,gv}$ and $F_{at,nw}$ are the carbon flows from the atmosphere reservoir to the ground vegetation and nonwoody tree reservoirs. The steady state values of $F_{i,at}$ and $F_{at,i}$ are given by *Harvey* [1989]; α_{ab} is the $^{14}C/^{12}C$ fractionation factor for CO_2 uptake by the terrestrial biosphere which is approximately square of that for $^{13}C/^{12}C$. According to *Keeling et al.* [1989a] the fractionation factor for $^{13}C/^{12}C$ is 0.982, and hence for $^{14}C/^{12}C$, $\alpha_{ab} = 0.964$.

The equation for the mass balance of mixed layer ^{14}C is

$$\frac{dN_mR_m}{dt} = k_aN_a\alpha_{am}R_a - k_a\left(\frac{N_{ao}}{N_{mo}}\right)(N_{mo} + \xi n_m)\alpha_{ma}R_m - \lambda N_mR_m + A\Omega\kappa\left.\frac{\partial N_dR_d}{\partial z}\right|_{z=h_m} + A\Omega w[(N_dR_d)|_{z=h_m} - N_bR_b] - F_gF_g^* \tag{A9}$$

where R_d is the mole ratio of ^{14}C to carbon in the ocean column.

The equation for the mass balance of ocean column ^{14}C is

$$\frac{\partial N_d R_d}{\partial t} = \kappa \frac{\partial^2}{\partial^2 z}(N_d R_d) + w\frac{\partial}{\partial z}(N_d R_d) + J_d J_d^* - \lambda N_d R_d \quad \text{(A10)}$$

with boundary conditions

$$N_d R_d = \frac{N_m R_m}{A\Omega h_m} \qquad \text{at } z = h_m \quad \text{(A11)}$$

$$\kappa \frac{\partial}{\partial z}(N_d R_d) + w N_d R_d = w N_b R_b \qquad \text{at } z = h_d \quad \text{(A12)}$$

The magnitude of the marine biosphere carbon source and the sink terms are not affected by the addition of fossil fuel CO_2 to the oceans because the biological productivity, according to the present understanding, is primarily limited by the nutrients such as phosphate and nitrate, and not CO_2 [*Siegenthaler and Sarmiento*, 1993]. The biological flux in our model acts as a natural background process and does not change with time. The corresponding ratios F_g^* and J^* of ^{14}C, however, follow the time-varying mixed layer ratio R_m times a fractionation factor of 0.954 [*Toggweiler and Sarmiento*, 1985].

Acknowledgments. Work was performed under the auspices of the U.S. Department of Energy at the Lawrence Livermore National Laboratory under contract W-7405-Eng-48 and was supported in part by the Department of Energy's Environmental Science Division. We thank J. R. Toggweiler and M. Heimann for contributing their GCM results for comparison.

References

Bacastow, R., and C. D. Keeling, Atmospheric carbon dioxide and radiocarbon in the natural carbon cycle, II, Changes from A. D. 1700 to 2070 as deduced from a geochemical model, in *Carbon in the Biosphere, AEC Symp. Ser.*, vol. 30, edited by G. M. Woodwell, and E. V. Pecan, pp. 86-135, National Technical Information Service, U.S. Department of Commerce, Springfield, Va., 1973.

Bacastow, R., and E. Maier-Reimer, Ocean circulation model of the carbon cycle, *Clim. Dyn.*, *4*, 95-125, 1990.

Bacastow, R. B., and E. Maier-Reimer, Dissolved organic carbon in modeling oceanic new production, *Global Biogeochem. Cycles*, *5*, 71-85, 1991.

Bazzaz, F. A., and E. D. Fajer, Plant life in a CO_2-rich world, *Sci. Am.*, *264*, 68-74, 1992.

Berger, W. H., K. Fischer, C. Lai, and G. Wu, Ocean productivity and oceanic carbon flux, I, Overview and maps of primary production and export production, *SIO Ref. 87-30*, Scripps Inst. of Oceanogr., Univ. of Calif., San Diego, La Jolla, 1987.

Bjorkstrom, A., A model of CO_2 interaction between atmosphere, oceans, and land biota, in *The Global Carbon Cycle*, Scope 13, edited by B. Bolin et al., pp. 403-457, John Wiley, New York, 1979.

Bjorkstrom, A., One-dimensional and two-dimensional ocean models for predicting the distribution of CO_2 between the ocean and the atmosphere, in *The Changing Carbon Cycle*, edited by J.R. Trabalka and D.E. Reichle, pp. 258-278, Springer-Verlag, New York, 1986.

Boden, T.A., R.J. Sepanski, and F.W. Stoss, *Trends '91: A Compendium of Data on Global Change*, 665 p., Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tenn., 1991.

Bolin, B., and E. Eriksson, Changes in the carbon content of the atmosphere and sea due to fossil fuel combustion, in *The Atmosphere and Sea in Motion*, edited by B. Bolin, Rockefeller Institute Press, New York, 1959.

Bretherton, F.P., K. Bryan, and J.D. Woods, Time-dependent greenhouse-gas-induced climate change, in *Climate Change: The IPCC Assessment*, edited by J.T. Houghton et al., pp. 173-193, Cambridge University Press, New York, 1990.

Broecker, W. S., Radioisotopes and large-scale oceanic mixing, in *The Sea*, vol. 2, edited by N.M. Hill, pp. 88-108, Wiley-Interscience, New York, 1963.

Broecker, W.S., and T.-H. Peng, *Tracers in the Sea*. 691 p., Lamont-Doherty Earth Observatory, Palisades, N.Y., 1982.

Broecker, W.S., and T.H. Peng, Carbon cycle 1985, glacial to interglacial changes in the operation of the global carbon cycle, *Radiocarbon*, *28*(2A), 309-327, 1986.

Broecker, W.S., and T.-H. Peng, The role of $CaCO_3$ compensation in the glacial to interglacial atmospheric CO_2 change, *Global Biogeochem. Cycles*, *1*, 15-39, 1987.

Broecker, W.S., Y.-H. Li, and T.-H. Peng, Carbon dioxide – man's unseen artifact, in *Impingment of Man on the Oceans*, edited by D.W. Hood, pp. 287-324, Wiley-Interscience, New York, 1971.

Broecker, W. S., T.-H. Peng, and R. Eng, Modeling the carbon system, *Radiocarbon*, *22*, 565-580, 1980.

Broecker, W.S., T.-H. Peng, G. Ostlund, and M. Stuiver, The distribution of bomb radiocarbon in the ocean, *J. Geophys. Res.*, *90*, 6953-6970, 1985.

Bryan, F., Parameter sensitivity of primitive equation ocean general circulation models, *J. Phys. Oceanogr.*, *17*, 970-985, 1987.

Caldeira, K., M.I. Hoffert, and A.K. Jain, Simple ocean carbon cycle models, in *Proceeding of the 1993 Global Change Institute on the Carbon Cycle*, edited by T.M.L. Wigley, and D. Schimal, Cambridge University Press, in press, 1994.

Crane, A. J., The partitioning of excess CO_2 in a five-reservoir atmosphere-ocean model, *Tellus*, *34*, 398-405, 1982.

Druffel, E. M., and T.W. Linick, Radiocarbon in annual coral rings of Florida, *Geophys. Res. Lett.*, *5*, 913-916, 1978.

Druffel, E. M. and H.E. Suess, On the radiocarbon record in banded corals, *J. Geophys. Res.*, *88*, 1271-1280, 1983.

Fiadeiro, M. E., Three dimensional modeling of tracers in the deep pacific ocean, II, Radiocarbon and circulation, *J. Mar. Res.*, *40*, 537-550, 1982.

Gargett, A.E., and G. Holloway, Dissipation and diffusion by internal wave breaking, *J. Mar. Res.*, *42*, 15-27, 1984.

Gordon, A. L., and H.W. Taylor, Heat and salt balance within the cold waters of the world oceans, in *Numerical Models of Ocean Circulation*, pp. 54-56, National Academy of Science, Washington, D.C., 1975.

Harrison, K., W. Broecker, and G. Bonani, A strategy for estimating the impact of CO_2 fertilization on soil carbon storage, *Global Biogeochem. Cycles*, *7*, 69-80, 1993.

Harvey, L. D. D., Effect of model structure on response of terrestrial biosphere models to CO_2 and temperature increases, *Global Biogeochem. Cycles*, *3*, 137-153, 1989.

Hesshaimer, V., M. Heimann, and I. Levine, Radiocarbon evidence for a smaller oceanic carbon dioxide sink than previously believed, *Nature*, *370*, 201-203, 1994.

Hoffert, M.I., Global distributions of atmospheric carbon dioxide in the fossil fuel era: A projection, *Atmos. Environ.*, *8*, 1225-1249, 1974.

Hoffert, M.I., and B.P. Flannery, Model projections of the time-dependent response to increasing carbon dioxide, in Projecting the Climatic Effects of Increasing Carbon Dioxide, *Rep. DOE/ER-0237*, edited by M.C. MacCracken and F.M. Luther, pp. 149-190, U.S. Dept. of Energy, Off. of Energy Res., Washington, D.C., 1985.

Hoffert, M.I., Y.-C. Wey, A.J. Callegari, and W.S. Broecker, Atmospheric response to deep-sea injections of fossil-fuel carbon dioxide, *Clim. Change*, *2*, 53-68, 1979.

Hoffert, M.I., A.J. Callegari, and C.-T. Hsieh, The role of deep sea storage in the secular response to climatic forcing, *J. Geophys. Res.*, *85*, 6667-6679, 1980.

Hoffert, M. I., A. J. Callegari, and C. T. Hseih, A box-diffusion carbon cycle model with upwelling, polar bottom water formation and a marine biosphere, in *Carbon Cycle Modeling*, Scope 16, edited by B. Bolin, pp. 287-305, John Wiley, New York, 1981.

Houghton, J. T., G.J. Jenkins, and J.-J. Ephraums, (Eds.), *Climatic Change: The IPCC Scientific Asessment*, Cambridge University Press, New York, 1990.

Joos, F., Imbalance in the budget, *Nature*, *370*, 181-182, 1994.

Keeling, C.D., The carbon dioxide cycle: Reservoir models to depict the exchange of atmospheric carbon dioxide with the oceans and land plants, in *Chemistry of the Lower Atmosphere*, edited by S.I. Rasool, pp. 251-329, Plenum, New York, 1973.

Keeling, C. D., Impact of industrial gases on climate, in *Energy and Climate*, pp. 72-91, Geophysics Study Committee, National Academy of Sciences, Washington, D.C., 1977.

Keeling, C. D., R.B. Bacastow, A.F. Carter, S.C. Piper, T.P. Whorf, M. Heimann, W.G. Mook, and H. Roeloffzen, A three-dimensional model of atmospheric CO_2 transport based on observed winds, 1, Analysis of observational data, in *Aspects of Climate Variability in the Pacific and Western Americas*, *Geophys. Monogr. Ser.*, vol. 55, edited by D.H. Peterson, pp. 165-236, AGU, Washington, DC, 1989a.

Keeling, C. D., S.C. Piper, and M. Heimann, A three-dimensional model of atmospheric CO_2 transport based on observed winds, 4, Mean annual gradients and interannual variations, in *Aspects of Climate Variability in the Pacific and Western Americas*, *Geophys. Monogr. Ser.*, vol. 55, edited by D.H. Peterson, pp. 305-363, AGU, Washington, D.C., 1989b.

Keir, R.S., On the late Pleistocene ocean geochemistry and circulation, *Paleoceanography*, *3*, 443-445, 1988.

Keir, R.S., Paleoproduction and atmospheric CO_2 based on ocean modeling, in *Productivity of the Ocean: Past and Present*, edited by W.H. Berger, V.S. Smetacek and G. Wefer, pp. 395-406, John Wiley, New York, 1989.

Kheshgi, H. S., B.P. Flannery, and M.I. Hoffert, Marine biota effects on the compositional structure of the world oceans, *J. Geophys. Res.*, *96*, 4957-4969, 1991.

Knox, F., and M. McElroy, Changes in atmospheric CO_2: Influence of the marine biota at high-latitudes, *J. Geophys. Res.*, *89*, 4629-4637, 1984.

Kohlmaier, G. H., R. Revelle, C.D. Keeling, and S.C. Piper, Reply to Idso, *Tellus*, *43B*, 342-346, 1991.

Kraus, E.B., Diapycnal mixing, in *Climate-Ocean Interaction*, edited by M.E. Schlesinger, pp. 269-293, Kluwer, Norwell, Mass., 1990.

Levitus, S., Climatological Atlas of the World Ocean, *NOAA Prof.* Paper *13*, Geophy. Fluid Dyn. Lab., Princeton, N.J., 1982.

Machta, L., The role of the oceans and biosphere in the carbon cycle, in *The Changing Chemistry of the Oceans*, edited by D. Dryssen and D. Jagner, pp. 121-145, Wiley-Interscience, New York, 1973.

Maier-Reimer, E., and R. Bacastow, Modeling of geochemical tracers in the ocean, in *Climate-Ocean Interaction*, edited by M.E. Schlesinger, pp. 233-267, Kluwer Academic Norwell, Mass., 1990.

Maier-Reimer, E., and K. Hasselmann, Transport and storage of CO_2 in the ocean–An inorganic ocean-circulation carbon cycle model, *Clim. Dyn.*, *2*, 63-90, 1987.

Maier-Reimer, E., U. Mikolajewicz, and K. Hasselmann, Mean circulation of the Hamburg LSG OGCM and its sensitivity to the thermohaline surface forcing, *J. Phys. Oceanogr.*, *23*, 731-757, 1993.

Munk, W.H., Abyssal recipes, *Deep Sea Res., Oregnogr. Abstr.*, *13*, 707-736, 1966.

Najjar, R. G., J.L. Sarmiento, and J.R. Toggweiler, Downward transport and fate of organic matter in the ocean: Simulations with the general circulation model, *Global Biogeochem. Cycles*, *6*, 45-76, 1992.

Oeschger, H., U. Siegenthaler, U. Schotterer, and A. Guglemann, A box-diffusion model to study the carbon dioxide exchange in nature, *Tellus*, *27*, 168-192, 1975.

Peng, T.-H., T. Takahashi, W.S. Broecker, and J. Olafsson, Seasonal variability of carbon dioxide, nutrients and oxygen in the northern North Atlantic surface water, *Tellus, Ser. B*, *39*, 439-458, 1987.

Redi, M.N., Oceanic isopycnal mixing by coordinate rotation, *J. Phys. Oceanogr.*, *12*, 1154-1158, 1982.

Revelle, R., and H.E. Suess, Carbon dioxide exchange bewteen the atmosphere and the ocean and the question of an increase of CO_2 during the past decades, *Tellus*, *9*, 18-27, 1957.

Sarmiento, J. L., Three-dimensional ocean models for predicting the distribution of CO_2 between the ocean and atmosphere, in *The Changing Carbon Cycle*, edited by J.R. Trabalka and D.E. Reichle, pp. 279-294, Springer-Verlag, New York, 1986.

Sarmiento, J. L., and E.T. Sundquist, Revised budget for the oceanic uptake of anthropogenic carbon dioxide, *Nature*, *356*, 589-593, 1992.

Sarmiento, J. L., and J.R. Toggweiler, A new model for the role of the oceans in determining atmospheric PCO_2, *Nature*, *308*, 621-624, 1984.

Sarmiento, J. L., C.G.H. Rooth, and W.S. Broecker, Radium 228 as a tracer of basin wide processes in the abyssal ocean, *J. Geophys. Res.*, *87*, 9694-9698, 1982.

Sarmiento, J.L., J.C. Orr, and U. Siegenthaler, A perturbation simulation of CO_2 uptake in an ocean general circulation model, *J. Geophys. Res.*, *97*, 3621-3645, 1992.

Sarmiento, J. L., R.D. Slater, M.J.R. Fasham, H.W. Ducklow, J.R. Toggweiler and G.T. Evans, A seasonal three-dimensional ecosystem model of nitrogen cycling in the North Atlantic euphotic zone, *Global Biogeochem. Cycles*, *7*, 417-450, 1993.

Schimel, D., I. Enting, M. Heimann, T. Wigley, D. Raynaud, D. Alves, and U. Siegenthaler, The carbon cycle, in *The IPCC 1994 Report on Radiative Forcing of Climate Change*, edited by J.T. Houghton, B.A. Callander, and S.K. Varney, in press, 1994.

Schlesinger, W. H., *Biogeochemistry: An Analysis of Global Change*, Academic, San Diego, Calif., 1991.

Semtner, A.J., and R.M. Chervin, Ocean general circulation from a global eddy-resolving model, *J. Geophys. Res.*, *97*, 5493-5550, 1992.

Shaffer, G., A model of biogeochemical cycling of phosphorous, nitrogen, oxygen, and sulfur in the ocean: One step toward a global climate model, *J. Geophys. Res.*, *94*, 1979-2004, 1989.

Shaffer, G. and J.L. Sarmiento, Biogeochemical cycling in the global ocean, 1, A new analytical model with continuous vertical resolution and high-latitude dynamics, *J. Geophys. Res.*, in press, 1994.

Siegenthaler, U., Uptake of excess CO_2 by an Outcrop-Diffusion Model of the ocean, *J. Geophys. Res.*, *88*, 3599-3608, 1983.

Siegenthaler, U. and F. Joos, Use of a simple model for studying oceanic tracer distribution and the global carbon cycle, *Tellus, Ser. B*, 44, 186-207, 1992.

Siegenthaler, U., and H. Oeschger, Predicting future atmospheric carbon dioxide levels, *Science*, *199*, 388-395, 1978.

Siegenthaler, U., and H. Oeschger, Biospheric CO_2 emissions during the past 2000 years reconstructed by deconvolution of ice coredata, *Tellus, Ser. B*, *39*, 140–154, 1987.

Siegenthaler, U. and J.L. Sarmiento, Atmospheric carbon dioxide and ocean, *Nature*, *365*, 119-125, 1993.

Siegenthaler, U., and T. Wenk, Rapid atmospheric CO_2 variations and ocean circulation, *Nature*, *308*, 624-626, 1984.

Stevenson, F.J., and E.T. Elliott, Methodologies for assessing the quantity and quality of soil organic matter, in *Dynamics of Soil Organic Matter in Tropical Ecosystems*, edited by D.C. Coleman, J.M. Oades, and G. Vehara, pp. 173–242, University of Hawaii Press, Honolulu, 1990.

Stuiver, M., and H. Pollach, Discussion reporting of ^{14}C data, *Radiocarbon*, *19*, 355-363, 1977.

Stuiver, M., and P.D. Quay, Changes in atmospheric ^{14}C attributed to variable Sun, *Science*, *207*, 11-19, 1980.

Stuiver, M., and P.D. Quay, Atmospheric ^{14}C changes resulting from fossil fuel CO_2 release and cosmic ray flux variability, *Earth Planet. Sci. Lett.*, *53*, 349-362, 1981.

Suess, H. E., Radiocarbon concentration in modern wood, *Science*, *122*, 415-417, 1955.

Sundquist, E.T., Geological perspectives on carbon dioxide and the carbon cycle, in *The Carbon Cycle and Atmospheric CO_2: Natural Variations Archean to Present, Geophys. Monogr. Ser.*, vol. 32, edited by E.T. Sundquist, W.S. Broecker, pp. 5-59, AGU, Washington, D.C., 1985.

Takahashi, T., W. S. Broecker, and A. E. Bainbridge, Supplement to the alkalinity and total carbon dioxide concentration in the world oceans, in *Carbon Cycle Modeling*, Scope 16, edited by B. Bolin, pp. 159-199, John Wiley, New York, 1981.

Tans, P. P., A compilation of bomb ^{14}C data for use in global carbon cycle models, in *Carbon Cycle Modeling*, Scope 16, edited by B. Bolin, pp. 131-158, John Wiley, New York, 1981.

Tans, P. P., I. Y. Fung, and T. Takahashi, Observational constraints on the global atmsopheric CO_2 budget, *Science*, *247*, 1431-1438, 1990.

Toggweiler, J. R., and B. Samuels, New radiocarbon constraints on the upwelling of abyssal water to the ocean's surface, in *The Global Carbon Cycle, NATO ASI Ser.*, vol. I-15, edited by M. Heimann, pp 333-366, Spring-Verlag, New York, 1993.

Toggweiler, J.R., and J.L. Sarmiento, Glacial and interglacial changes in atmospheric carbon dioxide: The critical Role of ocean surface water in high-latitudes, in *The Carbon Cycle and Atmospheric CO_2: Natural Variations Archean to Present, Geophys. Monogr. Ser.*, vol. 32, pp. 163-184, AGU, Washington, D.C., 1985.

Toggweiler, J.R., K. Dixon, and K. Bryan, Simulations of radiocarbon in a coarse resolution world ocean model, 1, Steady state prebomb distributions, *J. Geophys. Res.*, *94*, 8217-8242, 1989a.

Toggweiler, J.R., K. Dixon, and K. Bryan, Simulations of radiocarbon in a coarse resolution world ocean model, 2, Distributions of bomb-produced carbon 14, *J. Geophys. Res.*, *94*, 8243-8264, 1989b.

Trumbore, S.E., Comparison of carbon dynamics in tropical and temprate soils using radiocarbon measurements, *Global Biogeochem. Cycles*, *7*, 275-290, 1993.

Volk, T., Multi-property modeling of the marine biosphere in relation to global climate and carbon cycles, ph.D. thesis, 348 pp., New York Univ., New York, 1984.

Volk, T., and M. I. Hoffert, Ocean carbon pumps: Analysis of relative strengths and efficiencies in ocean-driven atmospheric CO2 changes, in *The Carbon Cycle and Atmospheric CO2: Natural Variations Archean to Present, Geophys. Monogr. Ser.*, vol. 32, pp. 91-110, AGU, Washington, D.C., 1985.

Volk, T., and Z. Liu, Controls of CO_2 sources and sinks in the Earth scale surface ocean: Temperature and nutrients, *Global Biogeochem. Cycles*, 2, 73-89, 1988.

Walker, J.C.G., *Numerical Adventures with Geochemical Cycles*, 192 pp., Oxford University Press, New York, 1991.

Walker, J. C. G., and J.F. Kasting, Effect of fuel and forest conservation on future levels of atmospheric carbon dioxide, *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, *97*, 151-189, 1992.

Wigley, T.M.L., A simple inverse carbon cycle model, *Global Biogeochem. Cycles*, *5*, 373–381, 1991.

Wigley, T. M. L., Balancing the carbon budget: Implications for projections of future carbon dioxide concentration changes, *Tellus, Ser. B.*, 45B, 409-425, 1993.

Wyrtki, K., The oxygen minimum in relation to ocean circulation, *Deep Sea Res. Oceanogr. Abstr.*, 9, 11-23, 1961.

M.I. Hoffert, Department of Earth Systems Science, New York University, New York, NY 10003. (e-mail: hoffert@acf.nyu. edu)

A.K. Jain, and D.J. Wuebbles, University of Illinois, Department of Atmospheric Sciences, 105 S. Gregory Avenue, Urbana, IL 61801. (e-mail: jain@uiatma.atmos.uiuc.edu;e-mail: wuebbles@uiatma.atmos.uiuc.edu)

H.S. Kheshgi, Exxon Research and Engineering Company, Route 22 East, Annandale, NJ 08801.

(Received March 30, 1994; revised September 7, 1994; accepted Sepetember 14, 1994.)

ELSEVIER

Palaeogeography, Palaeoclimatology, Palaeoecology 121 (1996) 221–237

PALAEO

Estimating the accuracy of Russian paleotemperature reconstructions

Haroon S. Kheshgi [a], Andrei G. Lapenis [b]

[a] *Exxon Research and Engineering Company, Annandale, New Jersey 08801, USA*
[b] *Earth Systems Group, Department of Physics, New York University, New York, NY 10003, USA*

Received 19 October 1994; revised and accepted 23 June 1995

Abstract

A framework is developed for the estimation of errors in paleotemperature reconstructions, such as those compiled by Borzenkova (Borzenkova, 1992; Borzenkova et al., 1992). Three primary sources of error are considered: (1) methodological error—the difference between the temperature inferred (method dependent) from a sample and the actual temperature when the sample was formed, (2) time-dating error—the error in temperature caused by a sample formed at a time different than intended, and (3) spatial interpolation error—the error caused by interpolating from point data to create an average or map of paleotemperatures. Each of these error sources can be the primary source of uncertainty, depending on the epoch, location, and methods applied. Preliminary error estimates are given for Borzenkova's Holocene Climatic Optimum reconstruction of the zonal–mean–annual paleotemperature distribution. For this reconstruction, methodological error is found to be the dominant source of errors at low latitudes, while spatial interpolation error is found to be the dominant source at high latitudes. This finding is due to higher expected spatial variability of temperature at high latitudes which results in an increase in spatial interpolation error with latitude relative to methodological error which is insensitive to latitude. At mid-latitudes, where the reconstruction contained a higher density of data, the temperature was found to be significantly warmer than present for the Holocene Climatic Optimum. For epochs earlier than the Quaternary, the epochs were defined by time slices that span as much as millions of years, which suggests that time-dating error could be the primary source of error for paleotemperature reconstructions which are intended to represent snapshots in time. Natural climate variability for the Cenozoic and Cretaceous over time scales of millions of years is, however, not well quantified; therefore, it remains ambiguous how to quantify errors in temperature caused by inaccurate time dating for pre-Quaternary epochs.

1. Introduction

Three applications of ancient climate reconstructions have been proposed to better understand the potential magnitude and effects of global warming induced by the increase in greenhouse gases. The first proposal is to use paleoclimate data on warm epochs as analogs for future global warming (Borzenkova and Zubakov, 1984; Budyko and Izrael, 1991). However, the paleo-analog concept is seriously flawed because potential future climate change due to an enhanced greenhouse effect would likely occur over time scales shorter than are focused on in paleoclimate reconstructions, which presumably reflect near-equilibrium responses (Crowley, 1990). The second is to use

Elsevier Science B.V.
SSDI 0031-0182(96)00075-5

paleoclimate data as a target or boundary condition in numerical experiments with climate models (Barron and Washington, 1982; Barron and Washington, 1985). And the third is to use paleoclimate data to derive a global climate sensitivity to radiative forcing (Hoffert and Covey, 1992). All of these applications require accurate estimates of the paleotemperature distribution over the surface of the Earth at snapshots in time; however, there is concern (Webb et al., 1993) that, at present, the accuracy of paleotemperature reconstructions is insufficient for these uses. And there is even greater concern with using Russian paleoclimate reconstructions for which there is limited information available on the data and methods used to produce the reconstructions. While the accuracy of methods used to estimate the paleotemperature at which, for example, a sediment sample was formed has been addressed (Bradley et al., 1993), there has not been a comprehensive analysis of errors for paleoclimate reconstructions. In this report we form a framework to include errors that contribute to paleoclimate reconstructions, and make a preliminary estimate of errors in Russian paleotemperature reconstructions.

The recent compendium of Russian paleoclimate reconstructions published by Borzenkova (1992) and Borzenkova et al. (1992) contain contour maps of paleotemperatures and tables of zonal average temperatures for nine paleoclimatic epochs. The time ranges for the nine paleoclimatic epochs, as well as the main sources of data upon which the reconstructions are based, are summarized in Table 1. A wide variety of methods, which are reviewed in this paper, were employed to estimate paleotemperatures used to make the reconstructions. These methods exhibit varying degrees of accuracy. Furthermore, some information concerning the sources and methods used in constructing maps and averages is not available. These complications have led to doubt concerning the accuracy of the reconstructions which we begin to address in this paper.

The accuracy of paleoclimate reconstructions depends on more than just the accuracy of the method used to assess the temperature at which a sample was formed. Errors in paleotemperature reconstructions arise from three sources: methodological error, dating error, and spatial interpolation error.

Methodological errors account for the difference in temperature inferred from the measured value of some proxy or indicator, and the actual temperature at one location and time. This type of error contains contributions from both error in correlating the indicator to temperature and instrumental error in measuring the indicator. We define σ_m as the magnitude (one standard deviation) of methodological error. Usually, the majority of methodological error is due to the inaccuracy of correlations constructed from present temperatures and indicators which are applied to ancient climates.

Dating error for one location is the error in temperature due to the difference in time between the true age of a sample and the desired age of the reconstruction time snapshot. Error in estimation of sample age is due primarily to inaccurate stratigraphic estimations and natural mixing processes: e.g., bioturbation, turbulence mixing in the water/sediment boundary layer, and event stratification. In many reconstructions, however, samples with ages within a prescribed time range—which can be longer than the age-estimation error of a sample—are accepted and compiled in the reconstruction, adding to dating error. A statistical estimate of the dating error in terms of temperature can be constructed from an estimate of the dating error in terms of time given a statistical model for the temporal variability of paleotemperature. The magnitude σ_t of the standard dating error in units of temperature is shown below to be related to the amplitude of climate variability for time-scales equal in magnitude to the dating error in units of years. Climate variability, however, is difficult to assess for epochs prior to the Quaternary. During the Quaternary, analysis of paleotemperature variability gleaned from ice core (Lovejoy and Schertzer, 1986) and ocean sediment (Kheshgi et al., 1993) data has shown that the amplitude of variability increases with time-scale.

Methodological and dating errors lead to errors in estimating temperature at each location. If methodological and dating errors are statistically independent, then their effects add geometrically:

$$\sigma_p^2 = \sigma_m^2 + \sigma_t^2 \qquad (1)$$

Table 1
Main sources of paleoclimatic data used by Borzenkova (1992) and Borzenkova et al. (1992) for paleotemperature reconstructions

Past epoch	Time period (B.P.)	Main sources of paleotemperature data
1. Holocene Climatic Optimum	6.2–5.3 k.y.	Vipper et al., 1981; Burashnikova et al., 1982; Klimanov, 1982; Abramova, 1985; Dolukhanov, 1985; Barash, 1988; Barash et al., 1989; Borzenkova, 1992
2. Last Glacial Maximum	20–18 k.y.	CLIMAP Project Members, 1976; Barash et al., 1989; Borzenkova, 1992
3. Last Interglacial (Eemian)	127–125 k.y.	CLIMAP Project Members, 1984; Velichko et al., 1984; Barash et al., 1989; Borzenkova et al., 1992; Frenzel et al., 1992
4. Pliocene Climatic Optimum	4.3–3.3 m.y.	Budyko and Izrael, 1991; Borzenkova et al., 1992
5. Middle Eocene	48–45 m.y.	Hubbard and Boulter, 1983; Boersma, 1984; Kulcova, 1987; Velichko and Chapaluga, 1987; Borzenkova, 1992
6. Maastrichtian	70–66 m.y.	Teis and Naidin, 1973; Boersma, 1984; Parrish and Spicer, 1988; Borzenkova, 1992
7. Santonian–Coniacian	88–84 m.y.	Lloyd, 1984; Krasheninnikov and Basov, 1985; Vakhrameev, 1988; Borzenkova, 1992
8. Cenomanian	95–92 m.y.	Yasomanov, 1985; Parrish and Spicer, 1988; Borzenkova, 1992
9. Albian	105–95 m.y.	Ronov and Balukhovsky, 1981; Axelrod, 1984; Goldbert, 1984; Yasomanov, 1985; Vakhrameev, 1988; Borzenkova, 1992

where σ_p is the standard error of a paleotemperature estimate at one location.

One method of paleotemperature reconstruction is the use of many paleotemperature estimates to form an average over some region. Table 2 contains Borzenkova's estimations of annual-mean paleotemperatures averaged over 10° latitude bands for the Northern Hemisphere shown as anomalies from the present temperature. This data has been used (Houghton et al., 1990; Hoffert and Covey, 1992) to make inferences about the climate system, even though the accuracy of these estimates has not been addressed. In this paper we consider the accuracy of these zonal paleotemperature averages.

Poor spatial coverage leads to *spatial interpolation error* in zonal averages. This error increases with increasing spatial variability of temperature in a zone. This error is reduced by uniform coverage of a zone by data at an increased number of locations. Redundancy of closely spaced locations, while not improving spatial coverage, nevertheless, does reduce the error in that region by taking the appropriate mean. By considering the number of data locations, their distribution, the amplitude of spatial temperature variability, and their standard

Table 2
Zonal-mean-annual paleotemperature anomalies estimated by Borzenkova (1992) and Borzenkova et al. (1992) for the nine epochs listed in Table 1. The anomalies are defined as the mean-annual average over a latitude belt of the difference in temperature between that of the past epoch minus that of the present climate

Latitude belts	Temperature anomalies (°C) for past epochs[1]								
	1[1]	2[1,3]	3[1]	4[1]	5[1]	6[1]	7[1]	8[1]	9[1]
5°S–5°N	0.2	−2.0	0.2	1.5	2.9	1.8	2.9	1.9	2.8
5°N–15°N	0.3	−2.0	0.2	1.5	3.0	1.6	1.9	2.0	2.8
15°N–25°N	0.3	−2.0	0.4	1.6	3.0	2.1	3.5	4.1	4.1
25°N–35°N	0.4	−2.3	0.8	2.5	5.7	5.6	7.8	6.7	9.4
35°N–45°N	0.8	−5.4	1.6	3.6	10.6	8.6	13.0	9.4	14.1
45°N–55°N	1.5	−5.4	3.4	4.9	15.4	13.8	19.4	14.1	20.3
55°N–65°N	2.3	−8.3	4.5	7.5	19.3	15.2	21.1	15.0	21.7
65°N–75°N	3.0	−11.5	6.2	10.6	23.2	19.7	27.4	22.6	26.4
75°N–85°N	3.4	[2]	6.2	11.8	26.2	21.8	29.9	24.1	24.7
Northern Hemisphere mean	1.1	−4.4	1.9	3.6	9.7	6.9	9.8	7.7	10.3

[1] Numbers correspond to epochs listed in Table 1.
[2] No entry given by Borzenkova (1992).
[3] The listed anomaly for the Last Glacial Maximum is for the summer (August), not the annual average.

error σ_p, an estimate the standard error of the zonal averages can be made. The magnitude of spatial interpolation error and its dependence on the distribution of locations has been analyzed for the instrumental record of temperature change (see e.g. Madden et al., 1993; Shen et al., 1994), and we recommend the application of this type of analysis to paleoclimate reconstructions. In the preliminary estimates of errors of Russian paleoclimate reconstructions given in this paper, however, we make the simplifying approximation that the data from each location is far enough away from neighboring locations that each location can be approximated to be independent.

The reconstructed temperature anomalies listed in Table 2 for the past epochs that are warmer than present were created by first estimating the difference (anomaly) between the annual-average paleo-temperature and the current climate temperatures at locations where paleodata were available. Next the zonal average temperature anomaly for each latitudinal belt was formed by taking a weighted average of the anomaly values for data within each of the zones. The reconstructed zonal temperature anomalies for the Last Glacial Maximum, on the other hand, represent the paleotemperature anomaly of the summer only (August) surface temperatures. This was done because the data for the winter months of the Last Glacial Maximum was available over smaller regions than was the data for the summer months. The anomaly of warm season, cold season, and annual average temperature anomaly should not be expected to be the same, and averaging the warm and cold season anomalies to give the annual average anomaly can lead to biasing of the annual anomaly as discussed by COHMAP Members (1988). We, however, do not consider this source of bias explicitly in this study.

Independent compilation of paleoclimate reconstructions, such as the Russian reconstructions, can provide, alternatively, either greater spatial coverage (if the reconstructions are based on data at different locations), or a reliability check (if the reconstructions contain data at similar locations). For example, Table 2 lists the mean August temperature of the Northern Hemisphere during the Last Glacial Maximum to be 4.4°C colder than present. This anomaly value is three times the annual-global-average anomaly of −1.6°C estimated by

CLIMAP (1981). This difference is mostly due to the greater decrease of equatorial temperatures according to Russian data that was included in the reconstruction summarized in Table 2. In this reconstruction Borzenkova (1992) did use the CLIMAP data as well. However, Borzenkova (1992) used additional information derived from 120 piston cores (Barash, 1988) taken by Russian research vessels. The difference between Borzenkova's (1992) reconstruction and CLIMAP could be caused by different versions of micropaleontological analysis applied by the authors of Russian marine data (Barash, 1988) and CLIMAP (Imbrie and Kipp, 1971). Another possible explanation can be the fact that significant part of the Russian data are from the upwelling regions of the tropical Atlantic Ocean—e.g. 7 cores from the area of Benguela upwelling, compare to 2 cores from this region which were analyzed by CLIMAP (Barash, 1988). Recently, Guilderson et al. (1994) found that the sea surface temperatures derived from Barbados corals give a greater tropical temperature decrease for the Last Glacial Maximum than estimates made by CLIMAP. The spatial variation in sea surface temperature as reconstructed by CLIMAP (1981), differs from both the Russian reconstructions and the Guilderson et al. (1994) result. One possible ways to understand these apparent differences is to intercompare CLIMAP data with the independent Russian data. Before this can be done, however, specialists must agreed on the Russian taxonomy of foraminifera species, the size of the fraction used for analysis, and a detailed explanation of stratigraphic schemes.

2. Estimates of error sources

2.1. Methodological error

First, we consider the methodological errors applicable to the reconstructions shown in Table 2. The zonal averages listed in Table 2 were constructed from paleotemperature estimates from several independent methods: lithological, palaeoecological and palaeontological, and isotopic and geochemical. Each method for paleotemperature estimation differs in its accuracy and applicability.

Lithological methods are based on the current relationship between climatic zones and sedimentation type or weathering intensity (Strakhov, 1963; Fairbridge, 1970; Gordon, 1975; Shopf, 1980; Hallam, 1986). These methods are used to estimate long-term (persistent) changes in climate.

Over land, this group of methods can be used to determine three climatic regions: cold/cool-temperate, humid/semi-humid, and warm/arid climates. A good marker of cold and cool-temperate climates are low weathering rates and angular, unstratified ice-rafted sediments. The mean-annual temperature, which correspond to such conditions, cannot be higher than 10–15°C (Goldbert, 1984). The clearest marker of humid climates is the presence of bauxite—the final product of the laterite process—which forms in tropical and subtropical climates with dry seasons of one to three months duration and with a mean-annual temperature greater than 22–26°C (Strakhov, 1963). Hot arid climates are indicated by the presence of the evaporates containing layers of gypsum and dolomites. These minerals form in basins similar to the current Persian Gulf where the air temperature is from 25 to 27°C (Yasomanov, 1978).

Over the oceans, similar climatic belts can be found through the lithological and mineralogical composition of ocean sediments (Lisitsyn, 1981). For example, the accumulation region for the mineral glauconite in ocean sediments occurs when the mean-annual surface temperature is greater than 12–18°C. Hence, the presence of glauconite in sediment cores was interpreted as evidence for surface waters warmer than 12–18°C (Yasomanov, 1978). Sedimentary evidence of reef-building organisms shows that the surface waters in a region, at the time corresponding to reef building, were warmer than 18–20°C (Ronov and Balukhovsky, 1981). A marker for cold sea surface temperature is the presence of diatom oozes (silica) in ocean sediments in amounts greater than 30% SiO_2 by dry weight of non-carbonate deposits (Levitan and Bogdanov, 1980). The region of diatom oozes corresponds to the maximum reaches of the polar fronts with mean-annual sea surface temperature lower than 5–10°C (Lisitsyn, 1981).

Lithological methods define upper and/or lower bounds on temperature when a marker is present. These methods can, therefore, be applied to mark the locations of specific ranges of surface temperatures, but cannot be used to indicate temperatures at every given location. As summarized in Table 3, the standard methodological error σ_m of the bounds on temperature range from ± 1 to ± 3°C (assumed to be equal to half the range of bounds discussed for each of the methods in this section). However, since the range of temperatures indicated by an upper or lower bound is quite large (much larger than the error in the bounds), this group of lithological methods allows only a low accuracy estimate of long term climate change. These methods are useful, however, as additional and independent ways of linking together other more accurate, but geographically sparse, indicators.

Palaeoecological and palaeontological methods can be divided into three groups (Borzenkova,

Table 3
Estimates of the standard methodological error σ_m for methods used in constructing Russian paleotemperature maps and zonal averages. Table structure follows that in the review article by Bradley et al. (1993)

Methods	Sources	Measured Properties	Standard Errors, σ_m	Method Limitations	References
1.Isotopic and Geochemical	Benthic & Planctonic Foraminifera, and Corals	Oxygen isotopes	± 1°C	Difficult to separate salinity, precipitation and sea level effects	Bradley et al. (1993)
	Organic matter	37 UK (long-chain alkenones)	± 1°C	Applicable from 0 - 500 Kyr	Bradley et al (1993)
	Belemnites and mollusk shells	Ca/Mg ratio	Error of from ± 2°C to ± 4°C is location dependent.	Temperature estimates affected by changes in sea water chemistry	Teis and Naidin (1973)
2.Palaeoecological and Palaeontological	Foraminifera, Radiolaria, Coccoliths, and Diatoms	Assemblages	Error of from ± 1.2°C to ± 2°C is location dependent. For the bulk of the oceans the error is ± 1.5°C	Equations are ocean-calibration specific	Barash (1988), Barash et al. (1989), and Bradley et al. (1993)
	Remains of plants and pollen	Morphology, assemblages, and pollen spectrum	± 1°C and ± 50 mm of annual precipitation	Temperature estimates affected by changes in atmospheric CO_2 level	Glichuk (1969), and Klimanov (1976)
	Remains of trees and leaves	Morphology	Less than ± 2°C, but more than ± 1°C	Separation of temperature signal from precipitation and CO_2	Wolfe (1981), and Borzenkova (1992)
	Skeletons and bones	Morphology and anatomy of organisms	Less than ± 2°C	Applicable for limited temperature ranges	Axelrod (1984), and Borzenkova (1992)
3.Lithological	Unstratified ice-rafted sediments	Geographic distribution	< 12.5 ± 2.5°C	Marker of specific temperature range for long-term climatic changes	Goldbert (1984)
	Bauxites		> 24 ± 2°C		Strakhov (1963)
	Evaporites with layers of gypsum and dolomites		> 26 ± 1°C		Yasomanov (1978)
	Glauconites		> 15 ± 3°C		
	Coral reefs		> 19 ± 1°C		Ronov and Balukhovsky (1981)
	Diatom oozes		< 2.5 ± 2.5°C		Lisitsyn (1981)

1992): methods based on species or family markers, morphological markers, or changes of dominant species in the biocoenoses.

The first group of these methods is based on the correlation between regions occupied by different species or families and specific environmental conditions. For example, the distribution of fan-shaped palm trees is limited to area with temperature no less than 5°C in the coldest month; the accuracy of this correlation is about ±1°C (Brooks, 1949).

The morphological marker method is based on the dependence of morphology and anatomy of organisms on climatic conditions. This method is often applied to plant life. For example, small plants with a thick skin provides evidence of arid climate, while plants with smooth and "drop" ending leaves (which are adapted for fast removal of the water from the plant) indicate a high humidity climate (Axelrod, 1984). If many species are taken into account, the mean size of terrestrial fauna are negatively correlated with the mean-annual temperature (Vereshagin, 1988). This is in contrast to the oceans where the size of Foraminifera and mollusks is positively correlated with mean-annual temperature (Barash et al., 1989). The accuracy of morphological methods is about the same as lithological methods (Borzenkova, 1992). Gregory (1993) estimates the standard error σ_m of Eocene continental surface temperature based on the morphological method to be ±1.5°C. However, Gregory does not take in to account the possibility that higher atmospheric CO_2 levels in the Eocene (Budyko and Ronov, 1978; Freeman and Hayes, 1992; Berner, 1993) which could have affected leaf morphology; higher CO_2 levels could have led to larger fauna, which would result in too low an estimate of temperature. The standard error of this method, therefore, is likely to be larger than that estimated by Gregory. The application of morphological methods is limited by the availability of imprints of ancient organisms with high enough quality to be used for quantitative analysis.

Palaeoecological and palaeontological methods which are based on the relations between climate conditions and dominant species in the biocoenozes can be further subdivided into two different groups. The first group of methods are based on the present day correlations between the spore-pollen ratios (pollen spectrums), surface temperature, and precipitation (Grichuk, 1969). These methods can be applied to determine the shift of geographical zones with climate changes. This results in an error in temperature of about ±1°C and an error in annual precipitation of about ±50 mm (Klimanov, 1982). The second method type is multi-factor analysis combining several palaeomethods. This method is most often applied to oceanic microfaunal data and is based on determining the most temperature dependent Foraminifera species. However, there are a few variations of this method. For example, the factor analysis developed by Russian scientists (Barash, 1971; Barash et al., 1989) differs from the method which was used in the CLIMAP project (Imbrie and Kipp, 1971; CLIMAP Project Members, 1976). To estimate the accuracy of the Russian factor analysis method, Barash calculated the standard deviation between present sea-surface temperature in the Northern Hemisphere and the temperature estimated from microfaunal data taken from the tops of the sediment cores obtained by Russian scientific ships. Comparison at more than 130 locations gave a standard error of Russian microfaunistic method of ±1.2°C (Barash et al., 1989) which is comparable to the errors estimates obtained for CLIMAP data of ±1 to ±2°C (CLIMAP Project Members, 1976).

Isotopic methods give the most accurate estimates of paleotemperatures used in Russian reconstructions. The temperature dependence of oxygen 18 fractionation during the formation of calcite is responsible for the good correlation between temperature and $\delta^{18}O \equiv {}^{18}O/{}^{16}O$ ratio. The difference in the $\delta^{18}O$ ratio between a standard of calcium carbonate deposits and samples was measured with a standard measurement error of ±0.1‰. This error in $\delta^{18}O$ corresponds approximately to a ±0.4 to ±0.5°C error in temperature (Borzenkova, 1992). However, changes of $\delta^{18}O$ can also be caused by changes in continental ice volume, regional changes in salinity, changes in evaporation/precipitation ratio, and diagenetic processes. Therefore, the error σ_m of the oxygen isotopic method exceeds its measurement error. A better

estimate of the method's standard error is ±1 C (Barash, 1988). Another commonly used method in Russian paleoreconstructions is the so-called "Ca/Mg" method. This method is used to estimate mean-annual temperatures in shallow continental basins. The accuracy of this method is limited by uncertainties in past changes of Ca and Mg concentration. The standard error of Ca/Mg method estimated from present day geochemical data is ±1°C (Borzenkova, 1992). However, comparison of Ca/Mg paleotemperature estimates with paleotemperatures reconstructed by $\delta^{18}O$ based on the same samples from Jurassic and Early Cretaceous belemnites from Northern and Eastern Siberia shows that temperatures estimated by the Ca/Mg method differs from those derived from oxygen isotopic data by ±2 to ±4°C depending on location (Teis and Naidin, 1973).

We summarize our estimates of methodological errors σ_m in Table 3. Some of these estimates are used in our assessment of the accuracy of the zonal averages for the Holocene Climatic Optimum (Table 5). Paleotemperature reconstructions are produced using a mixture of relatively accurate and inaccurate methods. For example, in some instances the sea-ice border was marked by the presence of polar bear bones (Borzenkova, 1992); it is not clear how to translate uncertainties in this type of method into uncertainties in paleotemperature, and we have not attempted to do so. For the data at the remaining locations used to create published paleotemperature maps, the methodological error ranges from ±1°C to ±3°C.

2.2. Dating error

The difference in the ages of samples (used to estimate paleotemperatures) and the selected age of the paleotemperature reconstruction leads to an error in paleotemperature which we will call dating error. The accuracy with which the age of a sample can be determined—*measured-age error*—contributes to this error. The size of the measured-age error depends on the dating technique and varies over a wide range from years (e.g. tree rings or corals), to tens of percent of the age of the sample (Bradley et al., 1993). Perhaps a more important contributor is the range of measured ages that is deemed acceptable for a reconstruction—*time-slice error*; this range is equal to the chosen duration of the epoch as, for example, listed in Table 1. In addition, there is uncertainty in the definition of the time range of an epoch which Harland et al. (1990) refers to as *chronogram error*.

There are other sources of time dating error which can, however, have important effects.

A biased error can result if, for example, a sample is chosen to be in the Holocene Climatic Optimum because it exhibits a paleotemperature which is greater than samples at the same location and at nearby times; this was done in Borzenkova's (1992) reconstructions of the first five epochs listed in Table 1. Many of Russian paleoclimatologists used the so-called "climatostratigraphic" method of chronological control. The method of climatostratigraphy is based on the concept that a number of climatic events have global signatures. The extremes of isotopic curves or the positions of biochronological levels are then used as proxies for these global climatic events. This approach does not, however, take into account the transient responses of the different components of the climate system—oceans, glaciers, vegetation cover, etc.—which each have a different time lag. Crowley (1990) suggests that climatochronology of Russian paleoreconstructions for the last interglacial period is no better than a quarter of a precession cycle. We have not attempted to estimate the biased error which the climatostratigraphic method can cause; it's estimation, as well as the estimation of biased error contributions to the overall error in paleoclimate reconstructions remain open questions that will require detailed study.

Contamination of samples can occur through bioturbation by mixing sediment layers from different times (Berger et al., 1977; Goreau, 1980) which significantly decreases the time-resolution in some locations. While these sources of error are not included in estimates given in this study, consideration should be given to these sources of error when generating a paleotemperature reconstruction.

The dating error of Russian paleoreconstructions is listed in Table 4 and do contain contributions from measured-age, time-slice and chronogram errors. These errors are summed with the

Table 4
Estimates of the standard dating error σ_{time} which includes contributions from measured-age error, time-slice error, and chronogram error

Past epoch	Estimates of standard dating error σ_{time}
Holocene Climatic Optimum	±1 k.y.—ocean microfaunistic data (Barash et al., 1989) ±0.6 k.y.—continental paleontological data with ^{14}C dating (Bradley et al., 1993)
Last Glacial Maximum	±3 k.y. (Barash et al., 1989; Borzenkova, 1992; Imbrie et al., 1993)
Last Interglacial (Eemian)	±5 k.y. (Imbrie et al., 1993)—usually, the accuracy of the last-interglacial stratigraphy can be accurate to within a quarter of precession cycle ≈6 k.y. (Crowley, 1990)
Pliocene Climatic Optimum	±1 m.y.
Middle Eocene	±3 m.y.
Maastrichtian	±3 m.y.
Santonian–Coniacian	±3 m.y.
Cenomanian	±3 m.y.
Albian	±5 m.y.

assumption that the error sources are uncorrelated and that reported uncertainties or ranges can be approximated as a standard error.

Time dating error for the Holocene Climatic Optimum contains contributions from both time slice error (±0.5 k.y.) plus measured-age error. Continental data were subject to radiocarbon dating control which is reported (Bradley et al., 1993) to have a standard error of ±0.3 k.y. for samples of 5 k.y. in age. Marine data is reported (Barash et al., 1989) to have a larger standard error of ±1. k.y. due to the poor correlation of oceanic microfaunistic records for this time interval compared to continental palaeontological spectrums (Borzenkova, 1992). The sum of these errors leads to different values of land and ocean time dating error for the Holocene Climatic Optimum. For the remaining epochs it is difficult to find a reliable estimations of the dating errors for the continental paleotemperature reconstructions; therefore, for these epochs we consider only time dating errors for the oceans.

For Last Glacial Maximum and the Last Interglacial, measured-age error is the dominant source of time dating error since these epochs are well defined in time. For epochs preceding the Quaternary, however, the time-slice and chronogram errors dominate. Harland et al. (1990) estimated the chronogram error to be ±1 m.y. for the Pliocene Climatic Optimum and about ±2 m.y. for the five older epochs listed in Tables 1 and 4. The application of paleomagnetic measurements and high-resolution Milankovitch-scale stratigraphy to the Pliocene deposits can potentially produce sample age with a standard error much less than 1 m.y. (Zubakov, 1990). However, the paleotemperature reconstruction of Pliocene Climatic Optimum contains data from samples to which the best dating methods were not applied (Borzenkova, 1992). For this study, we assume that the quality of stratigraphical control is not better than the lowest-resolution method which were used in the reconstructions. Adding the chronogramm error to a time-slice error (equal to half the time range listed in Table 1) for these epochs leads to a total time dating error which is a million years or greater.

As seen Table 4, this estimate of time dating errors of the paleoreconstructions range from 600 years to 5,000,000 years. During such time

intervals, temperature could change significantly, as is evident in, e.g., the Greenland ice cores (Johnsen et al., 1992). Over a time interval of millions of years, climate variability could reach the size of the anomalies in zonal mean temperatures listed in Table 2. While it is not clear, how large exactly the amplitude of natural climate variability was during the Cenozoic or Cretaceous, there is evidence that temperature changes with periods from 20 to 100 k.y. did occur during the Cretaceous (Barron and Washington, 1985) as well as during the Quaternary period.

Estimation of errors incurred by time dating error in paleotemperature reconstructions requires information on temperature variability over the time scale of dating error. In Table 4, the standard time dating errors σ_{time} are listed in units of years. The time dating error can be expressed as an error in temperature if we have information about the structure function $S(\Delta t)$, which is the square root of the expected value of the square of the difference in temperature which can be estimated from information about the time series of temperature:

$$S(\Delta t) = \langle [T(t+\Delta t) - T(t)]^2 \rangle^{1/2}$$

$$\equiv \left\{ \frac{\int_0^\infty [T(t+\Delta t) - T(t)]^2 \, dt}{\int_0^\infty dt} \right\}^{1/2}$$

$$= \sqrt{2}\sigma_T [1 - \rho(\Delta t)] \qquad (2)$$

where the standard deviation of the temperature time series is σ_T and the autocorrelation coefficient of the time series is $\rho(\Delta t)$, and Δt is the time lag. The standard error in temperature due to time dating is related to the expected value of the square of the temperature difference by

$$\sigma_t = \sqrt{\langle [T(t_0+\eta) - T(t_0)]^2 \rangle} \qquad (3)$$

where the selected time of the reconstruction is t_0 and the dating error is η. It is assumed that both the temperature and the dating error are random functions of time. Since the dating error and the temperature are functions of time uncorrelated to each other, the conditional expected value can be applied, which is then related to the structure function:

$$\sigma_t = \sqrt{\langle [T(t_0+\eta) - T(t_0)]^2 | \eta \rangle} = \sqrt{\langle S^2(\eta) \rangle}. \qquad (4)$$

If the dating error has a Gaussian distribution, then the expected value can be evaluated by

$$\sigma_t = \sqrt{\langle S^2(\eta) \rangle}$$

$$= \sqrt{\frac{1}{\sqrt{2\pi\sigma_{\text{time}}^2}} \int_{-\infty}^{\infty} e^{-(\tau^2/2\sigma_{\text{time}}^2)} S^2(\tau) d\tau}. \qquad (5)$$

If the dating error does not have a Gaussian distribution, then the alternative distribution could be substituted in Eq. (5). While an alternative distribution would lead to a slightly different expected value of the temperature error σ_t, we do not attempt to define alternative distributions of dating error in this preliminary analysis. Note that our previous estimates of the structure function showed latitudinal amplification, and so high latitude bands will have a higher value of σ_t than low latitude bands. Land values may also have a different value than sea values, although we will assume the same value in this analysis.

The structure function estimated from the combination of temperature time series from ice cores and the instrumental record was found (Lovejoy and Schertzer, 1986) to exhibit a power law dependence $S(\Delta t) \propto \Delta t^{0.4}$ over the time range from 10 to 20,000 years. Analysis of the SPECMAP record for ocean temperature over the last several hundred thousand years shows (Kheshgi et al., 1993) a latitudinal dependence consistent with that found by Hoffert and Covey's (1992) analysis of paleotemperature reconstructions of past epochs, and also consistent with Lovejoy and Schertzer's (1986) power law. Up to time-scales Δt of 10,000 years we adopt the structure function

$$S(\Delta t) = (0.063\,^{\circ}\text{C/yr}^{0.4}) \times [0.36 + 3.21 \sin^4(\text{latitude})] \Delta t^{0.4}. \qquad (6)$$

The estimate of the structure function can be used

to find the standard error due to dating:

$$\sigma_t = \pi^{-0.25}(0.063^{\circ}\mathrm{C/yr}^{0.4})[0.36 + 3.21\sin^4(\text{latitude}))](2\sigma_{\text{time}}^2)^{0.2}\Gamma^{0.5}(0.9) \quad (7)$$

where the gamma function $\Gamma(0.9) \approx 1.069$. Using this equation, the dating error in units of temperature with standard error σ_t can be calculated from the dating error in units of time with standard error σ_{time}.

2.3. Error estimates of zonal averages

The construction of spatial averages of temperature anomalies modifies error estimates. Averages can be made over the entire globe, a hemisphere, or latitudinal belts as reported in Table 2. The average is made by taking a weighted sum of point data over a zone

$$T_{avg} = \sum_{i=1}^{n} w_i T_i \quad (8)$$

where w_i are the weighting factors which sum to one, T_i is the temperature anomaly at a point, and n is the number of points. Since the errors for individual points are assumed to be uncorrelated, the averaging of point data results in a reduction of the methodological and dating errors. However, an additional error is added that is caused by approximating the temperature field by a set of discrete points and interpolating to find an average. Clearly this spatial interpolation error will depend on a number of factors: the number of points, the spatial distribution of points, the spatial correlation of the temperature anomaly, and the spatial variability of temperature. There is, however, limited knowledge of these factors, and so we will consider a limiting case.

Consider the limiting case where there is no spatial correlation of temperature anomaly, and the temperature anomaly varies from point to point with a Gaussian distribution with variance σ^2_{space}. In this case, the position of the points will not matter since there is no spatial correlation. If the methodological, dating, spatial averaging errors are assumed to be uncorrelated, then the error of the zonal average temperature anomaly is

$$\sigma_{avg}^2 = \sum_{i=1}^{n} w_i^2(\sigma_m^2 + \sigma_t^2 + \sigma_{space}^2)_i \,. \quad (9)$$

In reality, there is a spatial correlation between points. The implications of spatial correlation relative to error estimate Eq. (9) are now examined. Consider the case where variance σ^2_{space} is determined from the distribution of temperature over the zone. If the points at which paleotemperatures are measured are distributed either over a uniform grid or are uniformly random in location, then the size of the error σ_{avg} will be less than that estimated by Eq. (9), on average. If, however, the points are not uniformly distributed, then the size of the error will increase. Analysis of errors becomes significantly more complicated if we are to include the effects of spatial correlation. Nevertheless, a study of the effects of spatial sampling on the size of the error of zonal average temperature by Madden et al. (1993) does show that the simple model (9) for spatial interpolation error is a fair approximation.

The size of the spatial interpolation error's contribution to the error of the average will be reduced in averaging regions where the spatial variability of temperature σ^2_{space} is small. And so, the size of the errors can be decreased by judicious choice of the region. Since one primary dependence of annual mean temperature over the surface of the earth is on latitude, the choice of latitudinal belts for averaging regions will lead to a smaller error of the regional average than most other choices of region.

3. Holocene climatic optimum error estimates

Of the reconstructions described by Borzenkova (1992), the information describing the paleotemperature reconstruction of the Holocene Climatic Optimum is the most complete. According to radiocarbon dating, the warm and stable climate that marks the Holocene Climatic Optimum followed the end of the last glacial period (8.0–7.9 k.y. B.P.). Borzenkova (1992) and Borzenkova et. al. (1992) show the zonal mean temperature recon-

struction for the warmest part of this interval to be from 6.2 to 5.5 k.y. B.P.

A map of the positions at which paleotemperature estimations with radiocarbon dating were made is given on page 151 of Borzenkova's (1992) monograph. From this map we estimate the number of points n_{land} listed in Table 5 that Borzenkova used to reconstruct surface temperature on the land for each latitudinal belt. While there is no explicit information on the *method* used to estimate paleotemperature, the methods Borzenkova (1992) most often referenced are landscape reconstructions based on palaeontological data. These reconstructions were made by Klimanov (1982) and Frenzel et al. (1992) with reported methodological error of ± 1°C: see Table 3. Over land we assign a standard time dating error in years of $\sigma_{time,\ land} = 0.6$ *k.y.*: see Table 4.

Borzenkova et. al. (1992) combined the sea surface paleotemperature reconstructions made for this time interval by Barash (1988) and Barash et al. (1989), and a SPECMAP archive (Imbrie et al., 1989). Barash (1988) does show a table (pp. 157, 158) with coordinates of the piston cores which were used to estimate paleotemperature. The SPECMAP archive and the work of Barash (1988) together give us the number of points n_{ocean} listed in Table 5 which Borzenkova (1992) and Borzenkova et al. (1992) used in the paleotemperature reconstruction of the Holocene Climatic Optimum. Most of the paleotemperature estimations from marine data used a factor analysis of the foraminifera species which exhibits an average methodological error of ± 1.5°C: see Table 3. The time dating of Barash's data was made by two methods: (1) by radiocarbon dating, and (2) for samples without radiocarbon control, by the correlation of biostratigraphic layers with an ^{18}O chronological curve (Imbrie et al., 1989). The accuracy of the chronology of the samples without radiocarbon control is roughly ± 1 k.y. at the Holocene Climate Optimum: see Table 4 (Barash et al., 1989) and this is the value used for the estimates made in Table 5.

Since we have information on the number of point paleotemperature data on land and over the oceans, and since we estimate that the size of the methodological and time dating errors are different over land as opposed to over the oceans, we consider land and ocean errors separately. We then form the area weighted average to estimate the error of the zonal paleotemperature. Note that the Holocene may have had a more stable climate than the late Quaternary which was covered by the Specmap archive which Kheshgi et al. (1993) used to construct the model Eq. (6) for climate variability. Lower amplitude climate variability would result in slightly smaller values of time-dating error than appear in Table 5.

We estimate the size of the spatial variability by analogy to the present distribution of temperature. Spatial variability is approximated to be the standard deviation of the annual mean surface temperature over each land and sea on a $5° \times 5°$ grid given by the National Climatic Data Center (1987). The values of $\sigma_{space,land}$ and $\sigma_{space,ocean}$ listed in Table 5 show that there is an amplification of spatial variability at moderately high latitudes and that the spatial variability over the land is greater than over the oceans.

To make the estimates of the average zonal temperature over land $\sigma_{avg,land}$ and sea $\sigma_{avg,ocean}$ we have assumed that each point was weighted equally. Methodological, time dating, and spatial coverage errors all contributed to the land and sea average errors listed in Table 5. To calculate the standard error in the zonal average over land $\sigma_{avg,land}$ and oceans $\sigma_{avg,ocean}$, we evaluate Eq. (9) with the weighting factor $w_{i,land}$ and $w_{i,ocean} = 1/n_{ocean}$ along with the estimates for methodological error, time-dating error, and spatial variability over land and ocean.

We can see the relative importance of methodological, time dating, and spatial coverage errors by comparing the values of σ_m, σ_t and σ_{space} for both land and oceans for the latitude belts. The estimated time-dating error σ_t is not the largest source of error for any of the latitude belts for which there was data. At low latitudes methodological error σ_m was the largest contributor to the error of the average. At high latitudes spatial coverage error σ_{space} is the largest contributor. The latitude dependence of the relative magnitudes of error sources is due to the expected latitudinal amplification of both spatial and temporal varia-

Table 5
Estimates of standard errors of the zonal-mean-annual paleotemperature reconstructions from Borzenkova (1992) and Borzenkova et al. (1992) for the Holocene Climatic Optimum. The standard methodological error used for land data is, $\sigma_{m,land}=1$°C, and for ocean data is $\sigma_{m,ocean}=1.5$°C. The time-dating error for land points $\sigma_{t,land}$ is calculated by evaluating Eq. (7) with the time-dating error in years for land points of $\sigma_{time,land}=0.6$ k.y. The time-dating error for ocean points $\sigma_{t,ocean}$ is calculated by evaluating Eq. (7) with the time-dating error in years for ocean points of $\sigma_{time,ocean}=1$ k.y.

Latitude belts	Ocean fraction f_{ocean}	Land spatial variability $\sigma_{space,land}$ (°C)	Ocean spatial variability $\sigma_{space,ocean}$ (°C)	Number of land points n_{land}	Number of ocean points n_{ocean}	Time dating error for land points $\sigma_{t,land}$ (°C)	Time dating error for ocean points $\sigma_{t,ocean}$ (°C)	Error of zonal average over land $\sigma_{avg,land}$ (°C)	Error zonal average over ocean $\sigma_{avg,ocean}$ (°C)	Error of zonal average σ_{avg} (°C)
5°S–5°N	0.77	0.51	1.09	2	3	0.26	0.32	0.81	1.09	0.86
5°N–15°N	0.77	0.85	0.75	7	15	0.26	0.32	0.51	0.44	0.36
15°N–25°N	0.68	1.37	1.14	9	15	0.29	0.36	0.57	0.50	0.38
25°N–35°N	0.59	1.85	1.32	11	8	0.41	0.50	0.65	0.73	0.50
35°N–45°N	0.54	2.63	1.76	32	13	0.66	0.81	0.51	0.68	0.44
45°N–55°N	0.43	3.5	2.77	44	4	1.06	1.31	0.57	1.71	0.80
55°N–65°N	0.38	3.63	2.83	28	3	1.57	1.93	0.77	2.16	0.95
65°N–75°N	0.47	3.9	3.45	15	2	2.08	2.55	1.17	3.21	1.63
75°N–85°N	0.81	1.99	2.47	0	0	2.45	3.01			

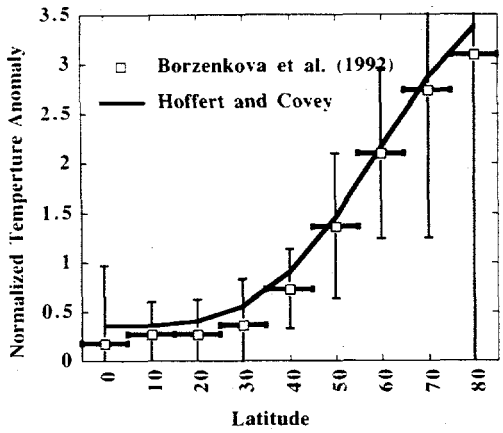


Fig. 1. The dependence of zonal-mean paleotemperature anomaly on latitude. The Borzenkova et al. (1992) reconstruction for the Holocene Climatic Optimum—values listed in Table 2—normalized to the Northern Hemisphere mean are shown by the square symbols with vertical error bars equal to the standard Zonal Error listed in Table 5 (normalized). Results are compared to Hoffert and Covey's (1992) fit of the latitudinal dependence of zonal-mean paleotemperature anomaly to two epochs: the Last Glacial Maximum, and Late Cretaceous.

bility which lead to increases in σ_t and σ_{space} with latitude, whereas σ_m remains independent of latitude.

An area weighted sum of the land and sea error is used to give the standard zonal error

$$\sigma^2_{avg} = (1 - f_{ocean})^2 \times \sigma^2_{avg,\,land} + f^2_{ocean} \times \sigma^2_{avg,\,ocean}. \quad (10)$$

The standard error estimates shown in Table 5 can be compared to the temperature anomaly for the Holocene Climatic Optimum listed in Table 2. In Fig. 1 we show the zonal temperature anomaly normalized (divided by) the Northern Hemisphere temperature anomaly for the Holocene Climatic Optimum. The vertical error bars represent one standard deviation of the zonal error estimate (from Table 5) normalized by the Northern Hemisphere mean paleotemperature anomaly (Table 2). The horizontal error bars represent the width of the latitudinal bands. The temperature anomaly is shown to be more than two standard deviations greater than zero at 60°N latitude, but nowhere else; we conclude that the Holocene Climatic Optimum was significantly, in a statistical sense, warmer than the present climate in this latitudinal belt.

The normalized zonal temperature anomaly for the Holocene Climatic Optimum is compared in Fig. 1 to the latitudinal dependence hypothesized by Hoffert and Covey (1992) to match paleotemperature data from the Last Glacial Maximum and the Late Cretaceous. While the temperature anomaly for the Holocene Climatic Optimum is consistent with Hoffert and Covey's hypothesis, the normalized standard zonal errors are too large to make this comparison a stringent test of their hypothesis.

4. Concluding discussion

In this report we have proposed a framework for estimation of errors in paleotemperature reconstructions like those published by Borzenkova (Borzenkova, 1992; Borzenkova et al., 1992). Three primary sources of error are considered: methodological error, time dating error, and spatial interpolation error. Each of these error sources can be the primary source of uncertainty, dependent on the epoch and location.

Preliminary error estimates are given for Borzenkova's (Borzenkova, 1992; Borzenkova et al., 1992) Holocene Climatic Optimum reconstruction of the zonal mean-annual paleotemperature distribution. For this reconstruction we estimate that methodological errors are the dominant source of errors at low latitudes, while spatial interpolation error is the dominant source at high latitudes. The lowest estimates of error in zonal mean annual paleotemperature are at mid-latitudes where there is the highest concentration of data. At mid-latitudes we estimate that the temperature anomaly was significantly positive for the Holocene Climatic Optimum. Time dating was not found to be a dominant source of errors for this epoch. We did not, however, include biased errors that might have occurred during the reconstruction caused by selection of the samples indicating the warmest periods in the time slice of the epoch chosen to find the warmest "Optimum". Nor did

we include biased methodological errors which would result if, for example, a transfer function had either a positive or negative error over a large part of one zone for a time slice; we recommend that methodological errors be characterized as either biased (spatially correlated error) or random. Biased errors could cause larger zonal errors than are estimated here.

Unfortunately, because of the lack of necessary information, we have not attempted an error analysis for the other eight epochs listed in Table 1. Information on the data locations and methods used to form these reconstructions is needed. For climates older than the Quaternary, the size of the time slices are on the order of millions of years which suggests that time dating error could be the primary source of error for paleotemperature reconstructions which are meant to represent snapshots in time. Since natural climate variability for the Cenozoic and Cretaceous over time scales of millions of years is not well quantified, there remains a problem of how to quantify time dating error for the pre-Quaternary epochs.

Acknowledgments

We thank Drs. B. White, L. Sloan, T. Crowley, K. Caldeira, M. Rampino, and M. Hoffert for their suggestions. Work at New York University was supported by the U.S. Department of Energy (DE-FG02-90ER 61014).

References

Abramova, T.A., 1985. Palaeogeographical conditions of the Aral–Caspian region during the Late Holocene (according to palaeonological data). In: Relief and Climate. Nauka, Moscow, pp. 91–101 (in Russian).

Axelrod, D.I., 1984. An interpretation of Cretaceous and Tertiary biota in polar regions. Palaeogeogr. Palaeoclimatol. Palaeoecol., 45: 106–147.

Barash, M.S., 1971. The vertical and horizontal distribution of planktonic foraminifera in Quaternary sediments of the Atlantic Ocean. In: B.M. Funnel and W.R. Riedel (Editors), The Micropaleontology of Oceans. Cambridge Univ. Press, pp. 433–442.

Barash, M.S., 1988. The Quaternary paleoceanography of the Atlantic Ocean. Nauka, Moscow, 272 pp. (in Russian).

Barash, M.S., Bloom, N.S. and Burmistrova, O.B., 1989. Neogene–Quaternary paleoceanology: according to micropaleontological data. Nauka, Moscow, 286 pp. (in Russian).

Barron, E.J. and Washington, W.M., 1982. Atmospheric circulation during warm geologic periods: Is the equator-to-pole surface-temperature gradient the controlling factor? Geology, 10: 633–636.

Barron, E.J. and Washington, W.M., 1985. Warm Cretaceous climates: High atmospheric CO_2 as a plausible mechanism. In: E.T. Sundquist and W.S. Broecker (Editors), The Carbon Cycle and Atmospheric CO_2: Natural Variations Archean to Present. AGU, Washington, DC, pp. 546–553.

Berger, W.H., Johnson, R.F. and Killingley, J.S., 1977. "Unmixing" of the deep-sea record and deglacial meltwater spike. Nature, 269: 661–663.

Berner, R.A., 1993. Paleozoic atmospheric CO_2: Importance of solar radiation and plant evolution. Science, 261: 68–70.

Boersma, A., 1984. Campanian through Paleocene paleotemperature and carbon isotope sequence and the Cretaceous–Tertiary boundary in the Atlantic ocean. In: W.A. Berger and J.A. Van Couvering (Editors), Catastrophes and Earth History. The New Uniformitarianism. Princeton Univ. Press, pp. 247–276.

Borzenkova, I.I., 1992. The changing climate during the Cenozoic. Hydrometeoizdat, Saint Petersburg, 247 pp. (in Russian, in press in English, Kluwer, Dordrecht).

Borzenkova, I.I. and Zubakov, V.A., 1984. Climatic optimum of the Holocene as a model of global climate of the early 21st century. Meteorol. Gidrol., 8: 69–77 (in Russian).

Borzenkova, I.I., Zubakov, V.A. and Lapenis, A.G., 1992. Global climate changes during the warm epochs of the past. Meteorol. Gidrol., 8: 25–37 (in Russian).

Bradley, R., Bard, E., Farquhar, G., Joussaume, S., Lautenschlager, M., Molfino, B., Raschke, E., Shackleton, N.J., Sirocko, F., Stauffer, B. and White, J., 1993. Group report: evaluating strategies for reconstructing past changes—what and where are the gaps? In: J.A. Eddy and H. Oeschger (Editors), Global Changes in the Perspective of the Past. Wiley, New York, pp. 144–171.

Brooks, C.E.P., 1949. Climate Through the Ages. Benn, London, 358 pp. (republished by Dover, New York, 1970).

Budyko, M.I. and Izrael, Y.A., 1991. Anthropogenic Climate Change. Univ. Arizona Press, Tucson, 275 pp.

Budyko, M.I. and Ronov, A.B., 1978. Evolution of the atmosphere during the Phanerozoic. Geochemistry, 5: 643–653.

Burashnikova, G.A., Muratova, M.A. and Suetova, I.A., 1982. Climatic model of USSR territory during Holocene optimum. In: Development of nature on the territory of the USSR in Late Pleistocene and Holocene. Nauka, Moscow, pp. 245–251.

CLIMAP Project Members, 1976. The surface of the ice age earth. Science, 191: 1131–1137.

CLIMAP Project Members, 1981. Geological Society of America Map Chart Series.

CLIMAP Project Members, 1984. The last interglacial ocean. Quat. Res., 21.

COHMAP Members, 1988. Climatic change of the last 18,000 years: observations and model simulations. Science, 241: 1043–1052.

Crowley, T.J., 1990. Are there any satisfactory geologic analogs for a future greenhouse warming? J. Climate, 3: 1282–1292.

Dolukhanov, P.M., 1985. Arid zone of the Old World during the Late Pleistocene and Holocene. All-Union Geogr. Soc. Bull., 117: 16–23 (in Russian).

Fairbridge, R.W., 1970. An ice-age in the Sahara. Geotimes, 15: 18–20.

Freeman, K.H. and Hayes, J.M., 1992. Fractionation of carbon isotopes by phytoplankton and estimates of ancient CO_2 levels. Global Biogeochem. Cycles, 6: 185–198.

Frenzel, B., Pe'si, M. and Velichko, A.A. (Editors), 1992. Atlas of the Paleoclimates and Paleoenvironment of the Northern Hemisphere. Late Pleistocene and Holocene. Geogr. Res. Inst. HAS, Budapest, 146 pp.

Goldbert, A.V., 1984. The Basic Principles of the Regional Paleoclimatology. Nedra, Moscow, 221 pp. (in Russian).

Gordon, W.A., 1975. Distribution by latitude of Phanerozoic evaporite deposits. J. Geol., 81: 269–284.

Goreau, T.J., 1980. Frequency sensitivity of the deep-sea climate record. Nature, 287: 620–622.

Gregory, K.M., 1993. Paleobotanically derived estimates of Eocene continental climates. EOS, April 20: 49 (suppl.).

Grichuk, V.P., 1969. The reconstruction of some climate elements of the Northern Hemisphere in the Atlantic period of the Holocene. In: Holocene. Nauka, Moscow, pp. 41–57 (in Russian).

Guilderson, T.P., Fairbanks, R.G. and Rubenstone, J.L., 1994. Tropical temperature variations since 20,000 years ago: modulating interhemispheric climate change. Science, 263: 663–665.

Hallam, A., 1986. Origin of minor limestone–shale cycles: Climatically induced or diagenetic? Geology, 14: 609–612.

Harland, W.B., Armstrong, A.V., Cox, A.V., Craig, L.E., Smith, A.G. and Smith, D.G., 1990. A Geologic Time Scale 1989. Cambrige Univ. Press.

Hoffert, M.I. and Covey, C., 1992. Deriving global climate sensitivity from paleoclimate reconstructions. Nature, 360: 573–576.

Houghton, J.T., Jenkins, G.J. and Ephraums, J.J. (Editors), 1990. Climate Change, The IPCC Scientific Assessment. Cambridge Univ. Press.

Hubbard, R.N. and Boulter, M.C., 1983. Reconstruction of Palaeogene climate from palynological evidence. Nature, 301: 147–150.

Imbrie, J. and Mix, A., 1989. Ocean response to orbital forcing in the Late Quaterary: observational and experimental strategies. In: A. Berger (Editor), Climate and Geo-Sciences. Kluwer, Dordrecht, pp. 121–164.

Imbrie, J., Berger, A., Boyle, E., Clemens, S., Duffy, A., Howard, W., Kukla, G., Kutzbach, J., Martinson, D., McIntyre, A., Mix, A., Molfino, B., Morley, J., Peterson, L., Pisias, N., Prell, W., Raymo, M., Shackleton, N. and Toggweiler, J., 1993. On the structure and origin of major glaciation cycles. Paleoceanography, X: 699–735.

Imbrie, J. and Kipp, N.G., 1971. A new micropaleontological method for quantitative paleoclimatology: application to a Late Pleistocene Caribbean core. In: The Late Cenozoic Glacial Ages. New Haven, pp. 71.

Johnsen, S.J., Clausen, H., Dansgaard, X., Fuhrer, K., Gundestrup, N., Hammer, C., Iverson, P., Jouzel, J., Stauffer, B. and Steffensen, J., 1992. Irregular glacial interstadials recorded in a new Greenland ice core. Nature, 359: 311–313.

Kheshgi, H.S., Jayaraman, R. and Lapenis, A.G., 1993. Climate variability from paleoclimate records: Implications for detection of an enhanced greenhouse effect. In: 1993 Fall AGU Meet., 6–10 December, San Francisco, pp. 363.

Klimanov, V.A., 1982. Climate of the Eastern Europe during Holocene climatic optimum (by palaeonological data). In: Development of the Nature on the Territory of the USSR in Late Pleistocene and Holocene. Nauka, Moscow, pp. 251–258 (in Russian).

Krasheninnikov, V.A. and Basov, I.A., 1985. Cretaceous Stratigraphy of the South Ocean. Nauka, Moscow, 170 pp. (in Russian).

Kulcova, I.A., 1987. Stratigraphy of the Paleogene deposits in western Siberia by palynological data. Geol. Geophys., 6: 11–17.

Levitan, M.A. and Bogdanov, Y.A., 1980. History of the biological silica accumulation. In: A.S. Monin and A.P. Lisitsyn (Editors), Geology of the World Ocean. Nauka, Moscow, pp. 232.

Lisitsyn, A.P., 1981. The Climate Zonality and Sedimentation in the Oceans. Nauka, Moscow, 392 pp. (in Russian).

Lloyd, C.R., 1984. Pre-Pleistocene paleoclimates: the geological and paleontological evidence; modeling strategies, boundary conditions, and some preliminary results. Adv. Geophys., 26: 35–140.

Lovejoy, S. and Schertzer, D., 1986. Scale invariance in climatological temperatures and the local spectral plateau. Ann. Geophys., 4B4: 401–410.

Madden, R.A., Shea, D.J., Branstator, G.W., Tribbia, J. and Weber, R., 1993. The effects of imperfect spatial and temporal sampling on estimates of the global mean temperature: experiments with model data. J. Climate, 6: 1057–1066.

National Climatic Data Center, 1987. Monthly climatic data for the world. NOAA, NCDC, Ashville, NC.

Parrish, J.T. and Spicer, R.A., 1988. Late Cretaceous terrestrial vegetation: a near polar temperature curve. Geology, 16: 22–25.

Ronov, A.B. and Balukhovsky, A.N., 1981. The climate zonality of of the continents and climatic changes during Late Mesozoic and Cenozoic. Lithol. Poleznue Iskop., 5: 118–136 (in Russian).

Shen, S., North, G. and Kim, K.-Y., 1994. Spectral approach to optimal estimation of the global average temperature. J. Climate, 7: 1999–2007.

Shopf, T., 1980. Paleoceanography. Harvard Univ. Press, 302 pp.

Strakhov, N.M., 1963. Types of Lithogenesis and Their

Evolution through the History of the Earth. Gosgeolizdat, Moscow, 230 pp. (in Russian).

Teis, R.V. and Naidin, D.P., 1973. Paleothermometry and Isotopic Composition of the Organic Carbonates. Nauka, Moscow, 247 pp. (in Russian).

Vakhrameev, V.A., 1988. Jurassic and Cretaceous Floras and Earth's Climates. Nauka, Moscow, 214 pp. (in Russian).

Velichko, A.A. and Chapaluga, A.L. (Editors), 1987. Climates of the Earth in the Geological Past. Nauka, Moscow, 225 pp. (in Russian).

Velichko, A.A., Grichuk, V., Gurtovaya, E. and Zelikson, E., 1984. The climate of the Northern Hemisphere in the epoch of the Last Mikulino Interglacial (Eemian). Proc. USSR Acad. Sci. Geogr. Ser., 1: 5–18 (in Russian).

Vereshagin, N.K., 1988. Paleogeography and paleoecology of the animals of the mamonths fauna during Quaternary time of the northern Eurasia. In: Global and Regional Territoriography, pp. 14–32 (in Russian).

Vipper, P.A., Meteltseva, E. and Sokolovskaya, V., 1981. Paleogeography of the Holocene and the Upper-Pleistocene of Central Mongolia. Izv. Akad. Nauk Est. SSR, 30: 74–82 (in Russian).

Webb, T.I., Crowley, T.J., Frenzel, B., Gliemeroth, A.-K., Jouzel, J., Labetyrie, L., Prentice, I.C., Rind, D., Ruddiman, W.F., Sarnthein, M. and Zwink, A., 1993. Group report: use of paleoclimatic data as analogs for understanding future global changes. In: J.A. Eddy and H. Oeschger (Editors), Global Changes in the Perspective of the Past. Wiley, New York, pp. 50–71.

Yasomanov, N.A., 1978. Landscapes and Climates during Jurassic, Cretaceous and Paleogene Epochs of the South of the USSR. Nedra, Moscow, 223 pp. (in Russian).

Yasomanov, N.A., 1985. Ancient Climates of the Earth. Gidrometeoizdat, Leningrad, 292 pp. (in Russian).

Zubakov, V.A., 1990. Global Climate Events of the Neogene. Gudrometeoizdat, Leningrad, 221 pp. (in Russian).

PALEOCLIMATE DATA CONSTRAINTS ON CLIMATE SENSITIVITY: THE PALEOCALIBRATION METHOD *

CURT COVEY
Global Climate Research Division, Mail Code L-264, Lawrence Livermore National Laboratory, Livermore, CA 94551, U.S.A.

LISA CIRBUS SLOAN
Institute of Marine Sciences, University of California, Santa Cruz, CA 95064, U.S.A.

and

MARTIN I. HOFFERT
Department of Earth Systems Science, New York University, New York, NY 10003, U.S.A.

Abstract. The relationship between paleoclimates and the future climate, while not as simple as implied in the 'paleoanalog' studies of Budyko and others, nevertheless provides sufficient constraints to broadly confirm the climate sensitivity range of theoretical models and perhaps eventually narrow the model-derived uncertainties. We use a new technique called 'paleocalibration' to calculate the ratio of temperature response to forcing on a global mean scale for three key intervals of Earth history. By examining surface conditions reconstructed from geologic data for the Last Glacial Maximum, the middle Cretaceous and the early Eocene, we can estimate the equilibrium climate sensitivity to radiative forcing changes for different extreme climates. We find that the ratios for these three periods, within error bounds, all lie in the range obtained from general circulation models: 2–5 K global warming for doubled atmospheric carbon dioxide. Paleocalibration thus provides a data-based confirmation of theoretically calculated climate sensitivity. However, when compared with paleodata on regional scales, the models show less agreeement with data. For example, our GCM simulation of the early Eocene fails to obtain the temperature contrasts between the Equator and the Poles (and between land and ocean areas) indicated by the data, even though it agrees with the temperature data in the global average. Similar results have been reported by others for the Cretaceous and for the Last Glacial Maximum.

1. Introduction

Climate sensitivity can be defined as the eventual (or equilibrium) change in global mean surface temperature in response to a prescribed change in global mean radiative forcing. A conventional measure of climate sensitivity is the global warming $\Delta T_{2\times}$ expected from doubling atmospheric CO_2. Although this definition excludes time-dependent effects and regional details, it serves as a first approximation for approaching the issue of future global change. General circulation models of climate obtain $\Delta T_{2\times}$ in the approximate range 2–5 K. Over the last two decades, no GCM with reasonable input assumptions has obtained a sensitivity much outside the 2–5 K range. At the same time it has proved all too easy, by varying a model's assumptions within the bounds of plausibility, to move its sensitivity from one extreme end of the range to the other (e.g., Mitchell *et al.*, 1989).

* The U.S. Government right to retain a non-exclusive royalty-free license in and to any copyright is acknowledged.

Climatic Change **32:** 165–184, 1996.
© 1996 *Kluwer Academic Publishers. Printed in the Netherlands.*

As an alternative to model-based predictions, Kellogg (1977) and Budyko and Izrael (1987) offered a 'paleoanalog' approach based on direct analogy with past warm periods. This method involves modeling the detailed time- and space-dependent details of future climate by analogy with past climatic changes. The problem with this method, however, is that Twenty-First Century global warming would probably involve unprecedented rates of climatic change for which there are no satisfactory geologic analogs (Crowley, 1990). A less ambitious but more justifiable approach – 'paleocalibration' – originated with Lorius *et al.* (1990). These authors used geologic data from the Last Glacial Maximum (LGM, 20 thousand years ago) to infer $\Delta T_{2\times}$ without attempting to forecast the time-evolving behavior or regional details of the future climate. We extended the paleocalibration approach to the warm mid-Cretaceous era of 100 million years ago (Hoffert and Covey, 1992). In this paper we compare our paleocalibration results with more recent results of others, we discuss a fundamental criticism of the technique (Lindzen, 1993), and we introduce a new paleocalibration data point, representing the early Eocene (55 million years ago).

Examination of paleoclimates with GCMs, the same models that predict future global warming, has been pursued for well over a decade (e.g., Hecht, 1985; Crowley and North, 1991). Although the paleocalibration technique is independent of GCMs, it leads naturally to questions of model fidelity in simulating paleoclimates. We conclude this paper with the results of new GCM simulations of the Eocene that address this point.

2. Method

In principle the paleocalibration technique is straightforward. For a given time interval, one obtains both the difference from present-day globally averaged surface temperature (ΔT) and the difference from the present-day globally averaged radiative forcing (ΔQ). ΔT is obtained from whatever geologic proxies are available. ΔQ is derived by calculating or estimating the total effect of the heat trapped by greenhouse gases and the changes in absorption of solar radiation due to changes in solar luminosity, surface albedo and atmospheric aerosol content. The next and final step is simply to define the ratio $\Delta T/\Delta Q$ as the climate sensitivity, which is the global temperature response to the radiative forcing.

As an example, Hansen *et al.* (1993) estimated that for the Last Glacial Maximum, ΔT was –5 K and ΔQ was –7 W m^{-2}. Most of ΔQ arises from continental ice sheets and atmospheric aerosols reflecting more solar energy back to space (ice core samples from the LGM clearly show that the atmosphere then contained much more dust than at present, though the exact amount it contained is controversial). A secondary term is the decreased trapping of infrared radiation due to smaller atmospheric amounts of CO_2 and CH_4. The $\Delta T/\Delta Q$ ratio is 0.7 K (W m^{-2})$^{-1}$. This quantity can be converted to a value for comparison with global warming

estimates by noting that a doubling of atmospheric CO_2 traps about 4 W m^{-2} of infrared radiation. Then, using the values from Hansen *et al.*, the expected global warming due to doubled CO_2 would be $(0.7) \times 4 = 3$ K, in the middle of the range of GCM estimates.

There are important limitations to the paleocalibration technique. First, the climate sensitivity as defined above says nothing about how long the system would take to respond to a given forcing. In the case of future global warming, the heat capacity of the oceans would introduce a lag time, which predictions of future climate would need to account for by means other than paleocalibration. Note, however, that the primary factor determining the lag time is the equilibrium sensitivity value itself (Hansen *et al.*, 1985; Wigley and Schlesinger, 1985). A second limitation of paleocalibration is that it determines only the globally averaged temperature response, not the pattern of regional response (nor the responses of other climatically important quantities like precipitation). As we discuss below, GCMs generally fail to simulate the observed difference in temperature response between Equator and Poles or between land and sea. Thus reliable forecasts of future regional climates are not yet obtainable from either paleocalibration or theoretical climate models.

A further subtlety comes in the definition of ΔQ. 4 W m^{-2}, for example, is the infrared trapping caused by doubled CO_2 *in the absence of other effects or feedbacks* such as changes in temperature, cloudiness or atmospheric water vapor content. ΔQ is defined as the total change in radiative flux at the top of the troposphere due only to changes in greenhouse gases, surface albedo, atmospheric aerosol content and solar luminosity. It may be thought of as the result of hypothetical, instantaneous changes in the above-mentioned factors, before temperature, clouds or water vapor have a chance to respond. Note also that by defining ΔQ as the instantaneous change in flux at the top of the troposphere ($\sim$ 10 km altitude), rather than at the top of the atmosphere, we exclude the stratospheric cooling effect of enhanced greenhouse gases. Instantaneous doubling of CO_2, for example, reduces outgoing infrared flux by about 4 W m^{-2} at the top of the troposphere but by only about 2 W m^{-2} at the top of the atmosphere (e.g., Lindzen, 1995; Hansen *et al.*, 1995). This difference in flux reduction is due to enhanced infrared emission to space by the stratosphere, caused by the increased concentration of IR-emitting CO_2 molecules there. The flux divergence in the stratosphere leads to cooling of that region. Within a few months of instantaneous CO_2 doubling, the stratosphere would thus adjust its temperature so that the flux change would be about 4 W m^{-2} at both the top of the atmosphere and the top of the troposphere (Hansen *et al.*, 1981). For this reason the definition of ΔQ as flux change at the top of the *troposphere* provides an appropriate measure of climate forcing. Another way to put the argument is that the entire troposphere is convectively coupled to the surface, while stratospheric temperature changes are decoupled from tropospheric temperature changes (Cess and Potter, 1988; Shine *et al.*, 1990).

The actual climate would of course change continuously with changing radiative forcing, so ΔQ cannot be measured directly. For example, one cannot expect satellite observations to record a decrease of several W m^{-2} in the infrared flux of Earth to space as greenhouse gases increase. Instead one would expect atmospheric temperature to increase to restore an approximate balance of absorbed solar energy and emitted infrared (i.e., global warming due to an enhanced greenhouse effect). ΔQ is well defined, however, despite its hypothetical nature. Given a specified set of changes in greenhouse gases, aerosols and surface albedo, ΔQ can be found as a straightforward exercise in radiative transfer.

In short, paleocalibration defines the climatic feedbacks involved in cloud and water vapor changes to be part of the response ΔT rather than part of the forcing ΔQ. The technique in effect measures the sum of cloud and water vapor feedbacks by observing ΔT. On the other hand, much slower processes like changes in atmospheric CO_2 and the growth and decay of continental ice sheets are included in ΔQ, the forcing. Distinguishing the fast feedbacks contained in ΔT from the forcing factors in ΔQ is thus a matter of time-scale separation. Paleocalibration does not aim to identify the root causes of past climatic changes, such as the causes of ice sheet growth and decay or of glacial-to-interglacial greenhouse gas variations. Instead the technique aims to measure the feedbacks that translate such root causes into temperature change. Feedbacks due to clouds and water vapor account for most of the uncertainty in the model estimates of future global warming.

3. Review of Results and a Fundamental Criticism

Table I compares our previous results (Hoffert and Covey, 1992) with subsequent estimates of Barron (1993) and Hansen *et al.* (1993). Note that we estimated $\Delta T/\Delta Q$ for both the Cretaceous and the LGM, whereas Barron dealt with the Cretaceous only, and Hansen *et al.* dealt with the LGM only. Both our Cretaceous and LGM estimates gave $\Delta T_{2\times} \sim 2$ K, at the low end of the GCM prediction range. On the other hand, the LGM estimate of Hansen *et al.* gave $\Delta T_{2\times} \sim 3$ K, near the center of the GCM range, and the Cretaceous estimate of Barron gave $\Delta T_{2\times} \sim 4$ K, in the upper half of the range.

The difference between our estimate and that of Hansen *et al.* arises from differing values of ΔT we used for the LGM. We used $\Delta T \approx -3$ K, obtained by taking a global average of the sea surface temperatures compiled by the CLIMAP analysis of LGM data. For land areas we simply assumed that the LGM cooling was identical to that of ocean areas in the same latitude zones. This assumption was based on the general principle that the atmosphere efficiently smoothes out temperature contrasts between land and sea within each latitude zone. It ignores residual land-sea temperature differences in climatic equilibrium, which, according to GCM simulations, would increase the global mean Ice Age cooling by an additional 0.5–1 K (Crowley and North, 1991, p. 79). Furthermore, some terrestrial

TABLE I
Intercomparison of paleocalibration estimates

	ΔT [K]	ΔQ [W m^{-2}]	$\Delta T_{2\times CO_2}$ [K]
Hoffert and Covey (1992)			
LGM	-3 ± 0.6	-6.7 ± 0.9	2 ± 0.5
Hansen *et al.* (1993)			
LGM	-5 ± 1	-7.1 ± 1.5	3 ± 1
Hoffert and Covey (1992)			
Cretaceous	9 ± 2	15.7 ± 6.8	2.5 ± 1.2
Barron (1993)			
Cretaceous	7 ± 2	8 ± 3.5	3.8 ± 2.0

geologic data suggest that the actual Ice Age cooling over land was substantially larger than that over ocean (Rind and Peteet, 1985). Based on this consideration, Hansen *et al.* chose $\Delta T \approx -5$ K for the LGM and obtained a correspondingly higher estimate of $\Delta T_{2\times}$ than ours. It has also been suggested that the tropical ocean ΔT values obtained by CLIMAP are too small in magnitude (Guilderson *et al.*, 1994). If this additional consideration is taken into account, the magnitude of Ice Age ΔT – and the inferred climate sensitivity – could exceed Hansen's estimate.

Barron's discussion of $\Delta T_{2\times}$ provides a further gauge of uncertainty in paleocalibration, although it was not presented as such. (See also Barron *et al.*, 1995, which reaches similar conclusions from a somewhat different point of view.) He estimated Cretaceous $\Delta T \sim 7$ K (an update of his earlier work, which had previously given a best guess of 9 K). He then considered possible combinations of $\Delta T_{2\times}$ and CO_2 amounts that could achieve the Cretaceous warming. For example, 7 K warming could result either from $\Delta T_{2\times} = 5$ K and 2.6 times present-day CO_2, or from $\Delta T_{2\times} = 2$ K and 11 times present-day CO_2 (assuming logarithmic scaling of CO_2 radiative forcing with CO_2 amount). Thus a 2–5 K range of GCM sensitivity to CO_2 doubling, together with the observation that $\Delta T \sim 7$ K, implies Cretaceous CO_2 amounts ranging from 3 to 11 times present. Considering 2–6 times present atmospheric CO_2 as the likely range of actual Cretaceous CO_2 amounts, Barron inferred that the climate's true $\Delta T_{2\times}$ must be significantly greater than 2 K in order to bring the implied Cretaceous CO_2 amounts within reasonable bounds. In the table we have put Barron's numbers into the paleocalibration format, using $\Delta T \approx 7$ K and calculating ΔQ from a 2–6 fold increase in CO_2. Application of the formal error analysis that we used for our own estimates (Hoffert and Covey, 1992, Eqn. 4) results in $\Delta T_{2\times}$ spanning a broad range, as shown in the table. The center of the estimate, however, is noticeably larger than the climate sensitivity that we inferred from the Cretaceous. The table shows the reason for the difference: our

ΔQ was twice as large as that resulting from only a 2–6 fold increase in CO_2. The main reason for this discrepancy is that we considered changes in surface albedo as well as greenhouse gases when we estimated ΔQ. Cretaceous surface albedo that was lower than present makes a contribution to ΔQ of nearly 6 W m^{-2} in the Hoffert-Covey estimate. Our larger ΔQ, together with approximately the same ΔT, implies a smaller climate sensitivity $\Delta T/\Delta Q$.

It should be noted in passing that a recent revision of Cretaceous temperatures by Sellwood *et al.* (1994) obtained 'minimum estimates' somewhat cooler than the Barron's lower limits. Sellwood *et al.*, however, failed to consider the substantial Equator-to-Pole gradient in the background sea water value of oxygen isotope ratio (their proxy for temperature). Taking this gradient into account substantially increases the tropical temperatures inferred from oxygen isotopes (Zachos *et al.*, 1994; Crowley and Kim, 1995).

The most important conclusion from Table I is that paleocalibration gives roughly the same range of possible values for $\Delta T_{2\times}$ as GCMs do. Although it does not change the conventional wisdom about the magnitude of potential human-induced climatic changes, paleocalibration strengthens the GCM-based theory by providing independent confirmation. Of course it would be useful if paleodata could be used to narrow the range of uncertainty in $\Delta T_{2\times}$. Our own previous results (Hoffert and Covey, 1992) and the preliminary Eocene analysis given in the following section imply that $\Delta T_{2\times}$ lies at the low end of GCM predictions. As discussed above, however, alternate interpretations of the paleodata can push $\Delta T_{2\times}$ upward. We must admit that our estimate of 3 K for Ice Age cooling is smaller than the consensus value among paleoclimatologists. Further, our ± 0.6 K LGM error limits accounted only for the scatter of CLIMAP longitude-averaged sea surface temperatures relative to a smooth curve, not the range of different interpretations of the paleodata.

In addition to controversy over the most appropriate input values for the paleocalibration technique, there are fundamental objections to the technique itself. Paleocalibration makes the basic assumption that globally averaged temperature response depends on the globally averaged forcing, i.e., that ΔT is a unique function of ΔQ. Lindzen (1993) has asserted to the contrary that under "an altered *distribution* of heating ... major changes in global climate may occur, even if the sensitivity to changing CO_2 is extremely small [emphasis added]." Although Lindzen did not give the details of his argument, it is evidently based on observations that: (1) tropical temperatures have shown very little increase during warm periods in Earth history; and (2) poleward fluxes of heat outside the tropics are dominated by baroclinic eddy transports (at least in the present-day atmosphere). If the threshold for baroclinic instability depends on meridional temperature gradient, then that gradient may remain largely fixed. Thus, if tropical temperatures do not respond much to increased CO_2, middle and high latitude temperatures also may not.

The paleoclimatic record, however, poses serious problems for dynamical theories in which extratropical heat fluxes are always dominated by baroclinic eddies. Relatively small tropical warmings have been accompanied on many occasions by substantial warmings at higher latitudes. This observation directly contradicts the theoretical prediction that meridional temperature gradient is largely fixed outside the tropics. For example, examination of the data summarized by Hoffert and Covey (1992) demonstrates that during the Cretaceous, the Equator-Pole temperature contrast was about one-half its present value. Increased poleward heat flows are required to explain the warmer high-latitude regions, despite the strong positive dependence of baroclinic heat flux on meridional temperature gradient. The apparent increase in heat flow despite a reduction in temperature gradient (presumably accomplished by processes other than baroclinic eddies) implies that energy fluxes in temperature latitudes can be substantially independent of the temperature structure there.

Related questions about paleocalibration come to mind from the geologic record. The coming and going of Ice Ages are clearly associated with Milankovitch forcing: small changes in the distribution of insolation, caused in turn by small variations in Earth's orbit about the Sun (Imbrie and Imbrie, 1979). Also, the glacial-interglacial CO_2 variations that comprise the main part of Ice Age ΔQ may themselves be caused by glacial-interglacial climate changes. So how can one infer climate sensitivity from them?

The answer to these questions is that paleocalibration does not attempt to identify root causes of climatic change, only to measure the feedbacks that determine the level of response to those root causes and result in the observed paleotemperatures. Thus, Milankovitch forcing may well trigger the growth of continental ice sheets, but it is difficult to explain 3–5 K global cooling during Ice Ages without invoking positive feedbacks that amplify the climate's response, and would also produce significant global warming from doubled CO_2. For example, Kirk-Davidoff and Lindzen (1993) summarized calculations implying that the paleoclimate record is consistent with extremely strong negative feedbacks in the tropics. They presented a simple climate model in which significant global mean temperature changes resulted from merely changing the transport of heat from the Equator toward the Poles. In principle this result is not surprising, because nonlinear feedbacks can create a situation in which moving heat from one location to another will change the global mean temperature (Robock, 1978). Until we see the details of Kirk-Davidoff and Lindzen's calculations, however, we will remain skeptical that the mechanism they advocate can account for the large global changes evident from the geologic record, while at the same time yielding a sensitivity to global mean radiative forcing ($\Delta T_{2\times}$) significantly smaller than conventional wisdom.

Similar comments apply to a study published by Lindzen and Pan (1994). These authors proposed a mechanism to translate Milankovitch forcing into glacial-interglacial variations, via changes in poleward heat flow by the Hadley circulation. They showed that the observed frequencies of glacial-interglacial variations could

be explained by their theory. They concluded that "commonly used notions of climate sensitivity" are therefore "not relevant", and that "the present mechanism can readily produce major changes in climate (including, as a by product, changes in globally averaged temperature) in systems which are profoundly insensitive to a doubling of CO_2." The calculations performed by Lindzen and Pan, however, only aimed at reproducing observations of the frequencies of glacial-interglacial transitions – not their amplitudes. Again, what is lacking is a plausible theory, presented in quantitative detail, that produces 2–5 K global temperature changes from Milankovitch forcing while remaining "profoundly insensitive to a doubling of CO_2."

In our opinion, an important condition for plausibility of a climate theory is that it reproduce observed relationships between surface temperature (T_S) and satellite measurements of infrared radiation flux to space (F_{IR}). These observed relationships have long been used to infer that $\Delta T_{2\times}$ lies in the range of GCM-based conventional wisdom. The argument comes from the high correlation seen between F_{IR} and T_S, with a slope of about 2 W m^{-2} K^{-1} over most of the T_S domain (Warren and Schneider, 1979). Recall that doubled CO_2 produces a radiative forcing of about 4 W m^{-2}. It follows that the observed slope corresponds to $\Delta T_{2\times} \sim 2$ K (without taking albedo feedback into account). Note that 2 W m^{-2} K^{-1} is less than half the slope – implying more than twice the climate sensitivity – of a simple blackbody radiator, i.e., $4\sigma T_S^3 = 5.4$ W m^{-2} K^{-1} for $T_S = 288$ K, where σ is the Stefan-Boltzmann constant. In GCMs the main process boosting $\Delta T_{2\times}$ is positive water vapor feedback. In the models, warmer temperatures lead to greater amounts of atmospheric water vapor, a potent greenhouse gas. Enhanced water vapor, in turn, traps more infrared flux, lowering the slope dF_{IR}/dT_S. Recent studies have shown directly that the GCM-simulated connections among T_S, water vapor and F_{IR} agree with observations (Raval and Ramanathan, 1989; Rind *et al.*, 1991).

In short, GCMs meet our condition of plausibility by agreeing with the satellite data. It seems doubtful to us that such agreement could be achieved in a model dominated by *negative* water vapor feedback, as advocated by Lindzen (1990). The point made by Sun and Lindzen (1993), that "the water vapor content of the air above the trade inversion over the subtropics is not directly related to the sea surface temperature immediately below", is certainly reasonable. Nevertheless the satellite data appear to argue strongly for an indirect effect amounting to positive water vapor greenhouse feedback.

We do not mean to imply that the issue is decided. The satellite observations discussed above relate F_{IR} to T_S by examining different points in space at the same time, or by examining the same points at different times in the cycle of the seasons. There is no guarantee that the correlation thus obtained will be the same as that between globally averaged F_{IR} and globally averaged T_S in a changing climate. The relatively low value of dF_{IR}/dT_S observed from correlations of different latitudes could simply reflect warmer sea surface temperatures lying under the rising branch

of the tropical Hadley or Walker circulations. In these regions there is more water vapor and cloud cover to trap outgoing IR, compared with areas under sinking air. Correlations obtained by observing individual points during the seasonal cycle could arise mainly through seasonal movement of the Hadley cell. Note, however, that such effects would be limited to latitudes equatorward of about $\pm 30°$. GCMs also exhibit positive water vapor feedback poleward of these latitudes, in agreement with satellite data (Del Genio *et al.*, 1994). It is difficult for us to imagine that the relationship between globally averaged F_{IR} and globally averaged T_S will prove to be radically different from that implied by a straightforward interpretation of the satellite data.

The relationship between surface temperature and outgoing IR can be more complex in particular regions, especially in the tropics. Correlations of F_{IR} with T_S in the moist tropics, using the techniques discussed above, imply a positive water vapor feedback that becomes strong enough to initiate a runaway greenhouse effect. In other words, the correlations have $dF_{IR}/dT_S < 0$ for the warmest sea surface temperatures (Warren and Schneider, 1979; Ramanathan and Collins, 1991) or, equivalently, for most locations within 25° latitude of the Equator (Raval *et al.*, 1994). The observed stability of tropical temperature indicates that some compensating negative feedback is operative, but the process or processes stabilizing tropical temperature are still very much at issue. Various possibilities have been considered (Hoffert *et al.*, 1983; Ramanathan and Collins, 1991, 1993; Hartmann and Michelsen, 1993). Alternatively, it may be that near the Equator, the aliasing effects discussed above make the correlations unreliable indicators of the large-scale response to climatic changes. Perhaps a better indication comes from the study of Chou (1994), which examined satellite-observed fluxes over the entire Pacific from 30° S to 30° N. Comparing a selected El Niño with a non-El Niño period, Chou found $dF_{IR}/dT_S \sim 22$ W m^{-2} K^{-1}. If this result is statistically significant, it implies a strong negative feedback from water vapor or other causes in the tropics.

It is worth noting that Chou's results are not inconsistent with the climate sensitivity inferred by paleocalibration. From changes in tropical temperature relative to global mean temperature reconstructed for different paleoclimates (Hoffert and Covey, 1992) we estimate $\Delta T_{tropical}/\Delta T_{mean} \sim 1/5$. For a global mean radiative damping rate of ~ 2 W m^{-2} K^{-1} (as inferred by Hoffert and Covey), the tropical radiative damping would then be ~ 10 W m^{-2} K^{-1}. Combining Chou's value of dF_{IR}/dT_S and his corresponding result for solar flux changes gives 13 W m^{-2} K^{-1}.

To summarize, in our opinion the satellite observations do not directly prove or disprove the conventional wisdom about the magnitude of $\Delta T_{2\times}$. They do, however, provide relevant constraints that models predicting future climate should satisfy. GCMs that yield conventional-wisdom values of $\Delta T_{2\times}$ also generally reproduce the observed pointwise correlations of F_{IR} with T_S. Models such as those developed by Kirk-Davidoff and Lindzen should be published and subjected

to the same tests. By the same token, of course, GCMs should be examined for consistency with Chou's recent results.

4. Paleocalibration for the Eocene Earth

Our inferences from paleocalibration would be put on a firmer foundation if they could be confirmed with more data points representing additional time intervals of Earth history. Here we add a third data point, representing an interval of Earth history that is located temporally between the Last Glacial Maximum and the Cretaceous.

The early Eocene climate was probably the warmest since the Cretaceous (e.g., Shackleton and Boersma, 1981; Wolfe, 1985; Barron, 1987; Crowley and North, 1991; Cerling, 1992; Sloan and Barron, 1992). Surface geography for the early Eocene was similar to today's with the following differences: Australia was in a more southerly position than today, located adjacent to Antarctica; India was located in the tropics and had not yet collided with Asia; the Mediterranean was larger than it currently is; the Himalayas were not as great in elevation; also the Rockies, Andes, and Transantarctic Mountains may have been somewhat reduced in elevation. The polar regions had little or no ice and subtropical plants existed within the Arctic Circle. On the other hand tropical temperatures were similar to today's values (Zachos *et al.*, 1994). As shown below (Figure 1), the Eocene appears to be an extreme case in which tropical temperatures were no warmer, and perhaps even colder, than at present despite significant global mean warming. If any era exemplifies Lindzen's theory of global change through changes in Equator-to-Pole heat distribution, the Eocene should.

Solar output 55 million years ago was similar to that of the present day to within $\sim 0.5\%$ (e.g., Fig. 5 of Crowley *et al.*, 1991). Surface albedo would have been less than present during the Eocene due to: (1) lack of most, if not all, perennial continental and sea ice; (2) higher sea levels, leading to greater coverage of oceans; (3) presence of deciduous forests at high latitudes of both hemispheres; and (4) lack of extensive deserts. Estimates of atmospheric CO_2 during the early Eocene come from theoretical models of Berner (1991; 2× preindustrial) and from geochemical interpretations of organic carbon (Freeman and Hayes, 1992: 2×; Arthur *et al.*, 1991: 3–6×) and soil nodules (Cerling, 1992: 2×). These estimates agree that CO_2 was higher than present, but they vary over a broad range of 2 to 6 times preindustrial values. There is some thought that methane concentrations may have been higher than present during the Eocene due to extensive areas of swamps and wetlands (Sloan *et al.*, 1992), but there is no direct evidence for this. There is no evidence whatsoever regarding atmospheric aerosols during this time (D. Rea, personal communication).

Increased atmospheric carbon dioxide is probably the dominant term in Eocene radiative forcing compared with the present. For the contribution to ΔQ from CO_2

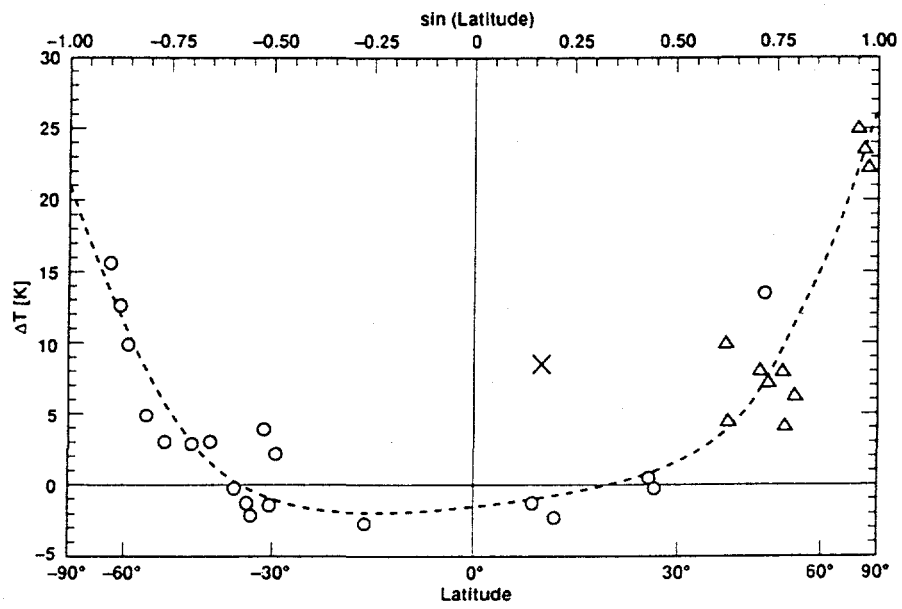


Fig. 1. Paleotemperature data vs. paleolatitude for the early Eocene. ΔT is the difference between Eocene and present-day surface temperature at each point; the dashed line is a least-squares fit to the data points. Circles show marine data (Zachos *et al.*, 1994), triangles show continental data (Sloan and Barron, 1992), and the 'x' denotes a single marine data point excluded from the analysis because it occurs in a modern-day upwelling zone whose anomalously cool temperature overestimates ΔT.

we take the full range of estimates discussed above, 2–6 times the preindustrial value. Using logarithmic scaling from 4.4 W m^{-2} for CO_2 doubling, this gives a contribution of 7.9 ± 3.5 W m^{-2}. We neglect possible contributions from methane. Of the four above-mentioned factors that contribute to a change in surface abledo, lack of ice probably dominates forest growth and lack of deserts (Bonan *et al.*, 1992). Covey *et al.* (1991) estimated that 2 to 3 W m^{-2} radiative forcing would result from the complete disappearance of sea ice from the present-day Earth, but we have excluded sea ice changes from our definition of ΔQ because we want to include such changes in the feedback processes measured by the paleocalibration technique. Accordingly we consider only the remaining contributor to changes in surface albedo, namely higher sea levels and the resulting decrease in the fraction of relatively high-albedo land areas. (Replacing land with ocean could of course lead to increased cloud cover, negating the surface albedo decrease; such an effect is defined as feedback rather forcing in the paleocalibration method). As a crude first estimate we set ΔQ due to this effect at two-thirds the Hoffert and Covey (1992) estimate of Cretaceous ΔQ due to surface albedo effects. Two-thirds is of course a rather arbitrary proportion, but it reflects the evidence that Eocene sea level was not quite as high as Cretaceous (Bluth and Kump, 1991). The result is 3.9 ± 0.6 W m^{-2}. Combining this with the contribution from increased CO_2 (and

taking the square root of the summed squared error estimates, as is appropriate for independent sources of error) gives a total ΔQ of 11.8 ± 3.6 W m^{-2}.

Early Eocene paleoclimate data indicate a world that was warmer than present, with greatest warming at high latitudes and little or no warming at low latitudes. Figure 1 shows the difference between mean annual Eocene and present surface temperatures as a function of latitude. We simply plotted local ΔT inferred from the paleodata at each available point on the globe together with the paleolatitude of each point (data taken from Sloan and Barron, 1992, and Zachos *et al.*, 1994). We then fit a fourth-order polynomial in the sine of latitude to the data, weighting all points equally. The equal-weighting assumption is of course a crude approximation, but the data are so sparse that a more sophisticated treatment, such as interpolation in latitude-longitude space, seems unjustified to us. The globally averaged temperature change obtained from the integral of the fitted curve is $\Delta T = 4.3$ K. (Note that both land and ocean points are used to obtain this value; if land points are excluded the same procedure gives $\Delta T = 3.3$ K.) To obtain an error estimate for this figure we first note that the root-mean-square scatter of points about the fitted curve is 2.6 K. This should be divided by the square root of the number of data points (29) to obtain the contribution to uncertainty in the global average: ± 0.5 K. We must also include errors in translating proxy measurements to temperatures. Although these are more difficult to quantify, there appears to be consensus among those working with the ocean data that ± 2 K is a reasonable estimate for the total error in this category. Errors in the land data may be higher, but we will use ± 2 K because most of the data in Figure 1 come from the ocean (we do not divide 2 K by $\sqrt{29}$ because the errors in proxy-to-temperature conversion may well be systematic). Combining ± 0.5 K and ± 2 K in root-sum-square fashion then gives a final estimate of $\Delta T = 4.3 \pm 2.1$ K.

In Figure 2 our Eocene estimates of ΔT and ΔQ are shown together with our earlier LGM and Cretaceous estimates (Hoffert and Covey, 1992), and the LGM estimate of Hansen *et al.* (1993). The figure illustrates application of the paleocalibration technique by representing the data in graphical terms, as a plot of ΔT versus ΔQ for several climate states in Earth history. By definition the present-day climate is a point at the origin. The three past climates, Eocene, Cretaceous and LGM, provide four more points (counting the independent LGM studies by us and by Hansen *et al.*). Conventional wisdom about Earth's climate sensitivity predicts that all points should lie in the range obtained from general circulation models, corresponding to 2–5 K warming for a doubling of atmospheric CO_2. Within error limits, the points do indeed lie within the theoretically predicted range. The Eocene data, however, stand out as implying a significantly smaller climate sensitivity than the Cretaceous or LGM data. Using Equation (4) in Hoffert and Covey (1992), the Eocene ΔQ and ΔT values presented above imply $\Delta T_{2\times} = 1.6 \pm 0.9$ K. This result must be viewed with caution in light of the preliminary nature of our Eocene numbers.

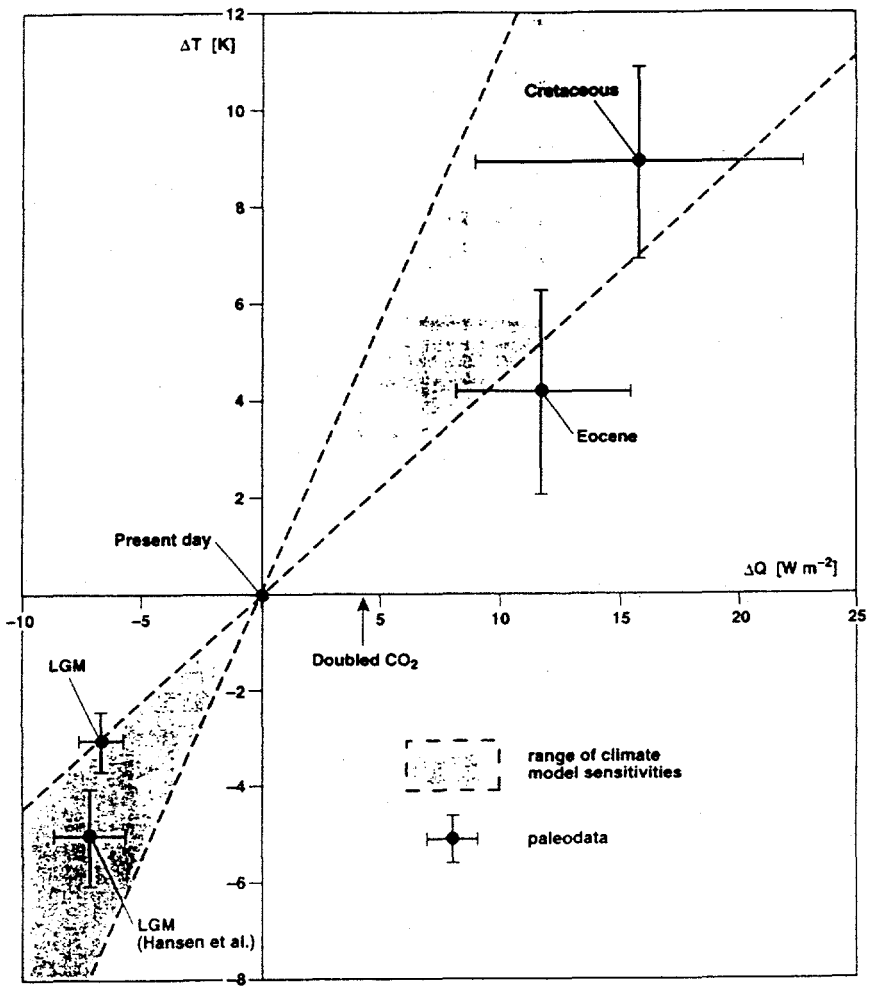


Fig. 2. Globally averaged ΔT vs. globally averaged ΔQ observed for several different paleoclimates. Shown for comparison is a range of climate sensitivity values obtained by theoretical models, equivalent to 2–5 K for CO_2 doubling.

The most robust conclusion from all the paleocalibration results considered together is that the paleodata are approximately consistent with the range of climate sensitivity predicted by theory. The data points imply that Earth's true climate sensitivity lies somewhat in the lower half of the model-predicted range, and they allow the possibility that the actual climate sensitivity is slightly less than the lowest model-predicted value. However, a radical challenge to the GCM-based conventional wisdom, such as a claim that models overestimate $\Delta T_{2\times}$ by an order of magnitude or more, would need to explain why the paleodata points in Figure 2 lie close to the model-predicted range. Correlation of course does not

imply causation, but we would find it surprising if the actual causes of global mean temperature change involved the distribution rather than the global mean of radiative forcing, as suggested by Lindzen (1993). In that case the correlation of data points in Figure 2 would be due to global mean climate changes causing changes in global mean radiative forcing (just coincidentally with the $\Delta T/\Delta Q$ ratio predicted by climate models). But the correlation of the four data points we have compiled – those of Hoffert and Covey (1992), the Eocene point discussed above, and the present-day climate by definition at the origin – is clearly positive ($r = 0.975$; $P < 0.03$). We would expect a negative correlation if the figure were revealing ΔT as the primary cause of ΔQ through the long-term carbon cycle. In that case, increased temperature could lead to increased weathering and hence enhanced removal of CO_2 from the atmosphere, leading to decreased ΔQ (Rampino and Caldeira, 1994).

To reduce error limits and gain more confidence in the placement of data points in Figure 2, we need a more thorough examination of the data for all time intervals plotted. For the Eocene this is especially intriguing because of newly available Russian data for the Eurasian continent that imply a much larger value for ΔT than Western reconstructions indicate. While our compilation of mean annual temperature data produces an Eocene global warming of 4.3 K over present values, a Russian compilation indicates mean global Eocene warming of 9.7 K (Hoffert, 1993). The Russian data span the entire Eocene epoch while our compilation is restricted to the early Eocene (see Sloan and Barron, 1992). However, since the early Eocene is thought to have been the warmest interval, the Russian estimates should have a cold bias and not a warm bias in comparison to our data. The discrepancy between the Eocene temperature estimates is an issue that will have to be clarified in the future.

5. Regional Climate Sensitivity

The foregoing conclusion, that GCMs and paleodata are in rough agreement, generally applies only to the global average of temperature. Considering regional scales in addition to global-mean scales, we note that a GCM simulation of the early Eocene, while obtaining globally averaged ΔT consistent with the data, fails to obtain the sharp equator-to-pole surface ΔT gradient and the proper land-sea thermal contrast that the paleoclimate data suggest (Sloan *et al.*, 1995; Sloan and Rea, 1995). The Eocene GCM results for 1, 2, and 6 times present atmospheric CO_2 show global warming relative to the control simulation of 1.0, 3.1, and 6.3 K respectively, consistent with the paleodata's implication of about 4 K warming under 2–6 times preindustrial CO_2. Figure 3 shows the annual mean, longitude-averaged surface temperature increase over the present day for the $2 \times CO_2$ and $6 \times CO_2$ Eocene simulations. Also shown are the data points (same points as in Figure 1). It is evident from the figure that although the model's change in tem-

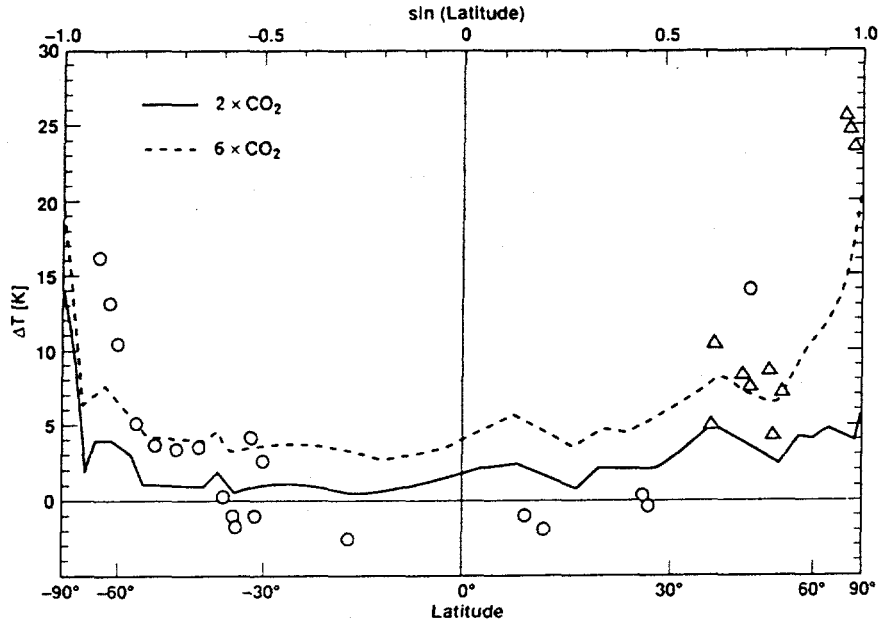


Fig. 3. Eocene paleodata (as in Figure 1) compared with two GCM simulations of the Eocene which assumed 2 and 6 times present levels of atmospheric CO_2.

perature approximately agrees with the data in the global mean, the distribution of ΔT with latitude obtained by the model is far too uniform. Put another way, the model obtains nearly the same sharp Equator-to-Pole contrast in absolute T for the Eocene as for the present day, whereas the geologic data imply that this temperature contrast was greatly reduced. For example, in mid-continental winters the model obtains below-freezing temperatures while data such as alligator and crocodile fossils clearly show that these areas did not undergo seasonal extremes during the Eocene (Sloan and Barron, 1992; Markwick, 1994; Sloan, in press). Similar problems were encountered in simulations of Cretaceous (Barron and Washington, 1984; Barron *et al.*, 1993) and LGM (Manabe and Broccoli, 1985) climates. The tendency of current GCMs to predict relatively uniform global warming, including substantial tropical warming, in the face of observations that indicate that tropical surface temperatures have changed little in the geologic past is a key criticism of the reliability of these models (e.g., Horrell, 1990; Crowley, 1991; Lindzen, 1993).

We are left with a great irony in our effort to understand climatic change. When compared with paleodata, general circulation models show fair performance in the global mean but poor performance at the next levels of approximation, i.e., Equator-Pole and land-sea temperature contrasts. In fact the GCMs agree with each other less and less as the spatial scale of comparison is decreased (Grotch and MacCracken, 1991). There are important exceptions to this discouraging trend

(COHMAP, 1988), but in general the predictions of GCMs, which give the most detailed simulations of climate available, are not reliable unless they are averaged to a global mean. Of course it is the regional details that matter to humans and to natural ecosystems.

A pessimist might conclude that for purposes of explaining past changes in climate or forecasting the future, GCMs are not any better than drastically simpler models, such as radiative-convective models that average over latitude and longitude. GCMs, however, have at least the potential for improvement. More rigorous examination of their inner workings (e.g., Gates, 1992), coupled with model-observation comparisons such as paleocalibration, may in time result in believable predictions of regional climate changes. Furthermore, GCMs even now provide a three-dimensional framework, constrained by known conservation laws for mass, momentum and energy, in which the effects of various climate-relevant processes may be tested. For example, possible attributes of tropical clouds and convection that would lead to negative water vapor feedback (Lindzen, 1990; Sun and Lindzen, 1993) could be programmed into a GCM, and the resulting effects on $\Delta T_{2\times}$ assessed. Such an exercise would not be possible in a one-dimensional model.

6. The Future of Paleocalibration

A glance at Figure 2 shows that even though there is rough agreement between GCMs and paleodata, both the range of model results and the scatter and error bars in the data are large. Uncertainties in the data must be reduced if the data is to distinguish among differing model results, rather than simply confirm that the climate's sensitivity lies approximately within the range of estimates from different models. Reducing uncertainties in the data would also test the validity of the paleocalibration technique itself. If a version of Figure 2 with more careful placement of data points and smaller error bars shows the points lying securely on a line through the origin, then the case for interpolating to find $\Delta T_{2\times}$ would be compelling. If on the other hand the correlation in $\Delta T - \Delta Q$ space were to disappear, the technique would clearly fail.

There are two ways to narrow uncertainty in the paleodata. First, we can try to refine estimates for the time periods we have already considered. This task would involve examining new data, such as the Russian estimates discussed above. It should also involve treating the data in a consistent way. For example, our error bounds for Ice Age ΔT accounted only for the scatter of points in the CLIMAP sea surface temperature data, not the widely held view that the CLIMAP data set systematically underestimates the globally averaged ΔT. On the other hand, our error bounds for CO_2 contribution to ΔQ during the Eocene and Cretaceous incorporate the full range of published estimates. This probably gives outlying estimates undue weight. Had we given less weight to the extreme range obtained

by Arthur *et al.* (1992), for instance, our Eocene ΔQ estimate would have been reduced, and the Eocene point on Figure 2 would have fallen more within the model-predicted range of climate sensitivity. Analogous comments apply to the Cretaceous upper limit of $11\times$ present-day CO_2 (Berner, 1990) used by Hoffert and Covey (1992). The question of how to treat outlying estimates is equivalent to a question of how to give the best estimate of $\Delta T_{2\times}$ – as a (mean value) $\pm$ (standard error), implying the mean value is most likely, or simply as a range of equally-likely possibilities. Here, too, a consistent treatment would be desirable.

A second way to reduce paleodata uncertainty is to use more data points. Considering additional time periods for paleocalibration is especially attractive to us because it provides a test of the main assumption underlying the technique itself, namely the conventional wisdom that global mean temperature is (to first approximation) a unique function of global mean forcing. For example, in the Pliocene era 3 million years ago, globally averaged surface temperature was perhaps 1–3 K warmer than present, and, until recently, atmospheric CO_2 was assumed to be about twice present levels (Crowley, 1991; Webb *et al.*, 1993). More recent data, however, suggests that CO_2 levels may have been only 30–50% above preindustrial (Raymo and Rau, submitted), corresponding to a ΔQ of 1.7–2.6 W m^{-2}. For $\Delta T = 2$ K the implied climate $\Delta T_{2\times}$ is 3.4–5.3 K, assuming as a crude first approximation that CO_2 is the only factor in Pliocene ΔQ.

Of course even a perfectly reliable estimate of $\Delta T_{2\times}$ would have its limitations in forecasting the future. By definition, measuring sensitivity with a single number, based on global-mean equilibrium temperature change, excludes regional and temporal variations. This procedure can underestimate the magnitude of worldwide climate change if temperature changes of opposite sign in different regions cancel in the global mean. One interpretation of the Eocene paleodata, for example, has cooler tropical temperatures coexisting with warmer temperatures at higher latitudes. Perhaps the anomolously low Eocene $\Delta T_{2\times}$ in Figure 2 should be considered in this light. Another potential problem is that, especially in view of nonlinear feedbacks involving water vapor and snow/ice albedo, $\Delta T_{2\times}$ should not be expected *a priori* to be a constant independent of the base state or the degree of perturbation. From Figure 2 it does not appear that this effect is severe enough to obscure the order-of-magnitude agreement between paleodata and model-based sensitivity calculations. The problem may appear, however, in the future. Using a more refined version of paleocalibration, as suggested above, may reveal nonlinearity in the relationship between ΔQ and ΔT.

In conclusion, despite the limitations of paleocalibration in providing accurate and regionally relevant connections between past and future climates, we believe the technique is promising. One point is clear in any case. Paleoclimates give us the only real-world data that includes global changes of the magnitude predicted to occur as a result of human perturbation of the atmosphere during the next century. Except for waiting for such changes to occur, examination of paleoclimatic data is the only way to directly test the validity of the models that predict such changes.

Acknowledgements

This work was performed under auspices of the U.S. Department of Energy's Environmental Sciences Division by the Lawrence Livermore National Laboratory under Contract W-7405-ENG-48. We thank Eric Barron, Tom Crowley and Dan Kirk-Davidoff for their comments on the manuscript.

References

Arthur, M. A., Hinga, K. R., Pilson, M. E., Whitaker, E., and Allard, D.: 1991, 'Estimates of pCO_2 for the Last 120 Ma Based on the $d^{13}C$ of Marine Phytoplanktic Organic Matter', [abs.], *Eos* (Transactions, American Geophysical Union **72**, no. 17, suppl., 166.

Barron, E. J.: 1987, 'Eocene Equator-to-Pole Surface Ocean Temperatures: A Significant Climate Problem?', *Paleoceanogr.* **2**, 729–739.

Barron, E. J.: 1993, *Paper presented at the American Geophysical Union Spring Meeting, Baltimore, 24–28 May.*

Barron, E. J., Fawcett, P. J., Peterson, W. H., Pollard, D., and Thompson, S. L.: 1995, 'A "Simulation" of Mid-Cretaceous Climate', *Paleoceanogr.*, (in press).

Barron, E. J., Fawcett, P. J., Pollard, D., and Thompson, S. L.: 1993, 'Model Simulations of Cretaceous Climates: The Role of Geography and Carbon Dioxide', *Phil. Trans. Roy. Soc. Lond. B* **341**, 307–316.

Barron, E. J. and Washington, W. M.: 1982, 'Cretaceous Climate: A Comparison of Atmospheric Simulations with the Geologic Record', *Palaeogeog., Palaeoclim., Palaeoecol.* **40**, 103–133.

Barron, E. J. and Washington, W. M.: 1984, 'Warm Cretaceous Climates: High Atmospheric CO_2 as a Plausible Mechanism', in Sundquist, E. T. and Broecker, W. S. (eds.), *The Carbon Cycle and Atmospheric CO_2, Natural Variations, Archean to Present*, American Geophysical Union, Washington, DC, pp. 546–553.

Berner, R.: 1990, 'Atmospheric Carbon Dioxide Levels over Phanerozoic Time', *Science* **249**, 1382–1386.

Berner, R.: 1991, 'A Model for Atmospheric CO_2 over Phanerozoic Time', *Amer. J. Sci.* **291**, 339–376.

Bluth, G. J. S. and Kump, L. R.: 1991, 'Phanerozoic Paleogeology', *Amer. J. Sci.* **291**, 284–308.

Bonan, G. B., Pollard, D., and Thompson, S. L.: 1992, 'Effects of Boreal Forest Vegetation on Global Climate', *Nature* **359**, 716–718.

Budyko, M. and Izrael, Y.: 1987, *Anthropogenic Climate Changes*, Gidrometeoizdat, Leningrad (in Russian; English translation by University of Arizona Press, 1990).

Cerling, T.: 1992, 'Carbon Dioxide in the Atmosphere: Evidence from Cenozoic and Mesozoic Paleosols', *Amer. J. Sci.* **291**, 377–400.

Cess, R. D. and Potter, G. L.: 1988, 'A Methodology for Understanding and Intercomparing Atmospheric Climate Feedback Processes in General Circulation Models', *J. Geophys. Res.* **93**, 8305–8314.

Chou, M.-D.: 1994, 'Coolness in the Tropical Pacific during an El Niño Episode', *J. Clim.* **7**, 1684–1692.

CLIMAP Project members: 1976, 'The Surface of the Ice-Age Earth', *Science* **191**, 1131–1137.

COHMAP Project Members: 1988, 'Climatic Changes of the Last 18,000 Years: Observations and Model Simulations', *Science* **241**, 1043–1052.

Covey, C., Taylor, K. E., and Dickinson, R. E.: 1991, 'Upper Limit for Sea Ice Albedo Feedback Contribution to Global Warming', *J. Geophys. Res.* **96**, 9169–9174.

Crowley, T.: 1990, 'Are There Any Satisfactory Geologic Analogs for Future Greenhouse Warming', *J. Clim.* **3**, 1282–1292.

Crowley, T.: 1991, 'CO_2 Changes and Tropical Sea Surface Temperatures', *Paleoceanogr.* **6**, 387–394.

Crowley, T., Baum, S. K., and Hyde, W. T.: 1991, 'Climate Model Comparison of Gondwanan and Laurentide Glaciations', *J. Geophys. Res.* **96**, 9217–9226.

Crowley, T. and Kim, K.-Y.: 1995, 'Comparison of Longterm Greenhouse Projections with the Geologic Record', *Geophys. Res. Lett.*, (in press).

Crowley, T. and North: 1991, *Paleoclimatology*, Oxford University Press, New York, 339 pp.

Del Genio, A. D., Kovari, W., Jr., and Yao, M.-S.: 1994, 'Climatic Implications of the Seasonal Variation of Upper Tropospheric Water Vapor', *Geophys. Res. Lett.* **21**, 2701–2704.

Freeman, K. and Hayes, P.: 1992, 'Fractionation of Carbon Isotopes by Phytoplankton and Estimates of Ancient CO_2 Levels, *Global Biogeochem. Cycl.* **6**, 185–198.

Gates, W. L.: 1992, 'AMIP: The Atmospheric Model Intercomparison Project', *Bull. Amer. Met. Soc.* **73**, 1962–1970.

Grotch, S. L. and MacCracken, M. C.: 1991, 'The Use of General Circulation Models to Predict Regional Climatic Change', *J. Clim.* **3**, 286–303.

Guilderson, T. P., Fairbanks, R. G., and Rubenstone, J. L.: 1994, 'Tropical Temperature Variations since 20,000 Years Ago: Modulating Interhemispheric Climate Change', *Science* **263**, 663–665.

Hansen, J., Johnson, D., Lacis, A., Lebedeff, S., Lee, P., Rind, D., and Russell, G.: 1981, 'Climatic Impact of Increasing Atmospheric Carbon Dioxide', *Science* **213**, 957–966.

Hansen, J., Lacis, A., Ruedy, R., Sato, M., and Wilson, W.: 1993, 'How Sensitive is the World's Climate?', *National Geogr. Res. Exploration* **9**, 142–158.

Hansen, J., Russell, G., Lacis, A., Fung, I., and Rind, D.: 1985, 'Climate Response Times: Dependence on Climate Sensitivity and Ocean Mixing', *Science* **229**, 857–859.

Hansen, J., Sato, S., and Ruedy, R.: 1995, 'Wonderland Model: Radiative Forcing Experiments', (in preparation for *J. Geophys. Res.*).

Hartmann, D. L. and Michelsen, M. L.: 1993, 'Large-Scale Effects on the Regulation of Tropical Sea Surface Temperature', *J. Clim.* **6**, 2049–2062.

Hecht, A.: 1985, *Paleoclimate Analysis and Modeling*, Wiley-Interscience Publishers, New York, 445 pp.

Hoffert, M. I.: 1993, *Paper presented at the American Geophysical Union Spring Meeting, Baltimore, 24–28 May.*

Hoffert, M. I. and Covey, C.: 1992, 'Deriving Global Climate Sensitivity from Palaeoclimate Reconstructions', *Nature* **360**, 573–576.

Hoffert, M. I., Flannery, B. P., Callegari, A. J., Hsieh, C.-T., and Wiscombe, W.: 1983, 'Evaporation-Limited Tropical Temperatures as a Constraint on Climate Sensitivity', *J. Atmos. Sci.* **40**, 1659–1668.

Horrell, M.: 1990, 'Energy Balance Constraints on ^{18}O Based Paleo-Sea Surface Temperature Estimates', *Paleoceanog.* **5**, 339–348.

Imbrie, J. and Imbrie, K. P.: 1979, *Ice Ages: Solving the Mystery*, Enslow Publishers.

Kellogg, W.: 1977, *Effects of Human Activities on Global Climate*, WMO Report No. 486, World Meteorology Organization, Geneva.

Kerr, R.: 1993, 'Fossils Tell of Mild Winters in an Ancient Hothouse', *Science* **261**, 682.

Kirk-Davidoff, D. B. and Lindzen, R. S.: 1993, *Paper presented at the American Geophysical Union annual Fall Meeting, San Francisco, CA.*

Lindzen, R. S.: 1990, 'Some Coolness Concerning Global Warming', *Bull. Amer. Meteorol. Soc.* **71**, 288–299.

Lindzen, R. S.: 1993, 'Paleoclimate Sensitivity', *Nature* **363**, 25–26.

Lindzen, R. S.: 1995, 'How Cold Would We Get under CO_2-Less Skies?', *Physics Today*, February issue, 78–80.

Lindzen, R. S. and Pan, W.: 1994, 'A Note on Orbital Control of Equator-to-Pole Heat Fluxes', *Clim. Dynam.* **10**, 49–57.

Lorius, C., Jouzel, J., Raynaud, D., Hansen, J., and Le Treut, H.: 1990, 'The Ice-Core Record: Climate Sensitivity and Future Greenhouse Warming', *Nature* **347**, 139–145.

Manabe, S. and Broccoli, A. J.: 1985, 'A Comparison of Climate Model Sensitivity with Data from the Last Glacial Maximum', *J. Atmos. Sci.* **42**, 2643–2651.

Markwick, P. J.: 1994, '"Equability", Continentality, and Tertiary "Climate": The Crocodilian Perspective', *Geology* **22**, 613–616.

Mitchell, J. F. B., Senior, C. A., and Ingram, W. J.: 1989, 'CO_2 and Climate: A Missing Feedback?', *Nature* **341**, 132–134.

Ramanathan, V. and Collins, W.: 1991, 'Thermodynamic Regulation of Ocean Warming by Cirrus Clouds Deduced from Observations of the 1987 El Niño', *Nature* **351**, 27–32.

Ramanathan, V. and Collins, W.: 1993, 'A Thermostat in the Tropics?', *Nature* **361**, 410–411.

Rampino, M. R. and Caldeira, K.: 1994, 'The Goldilocks Problem: Climatic Evolution and Long-Term Habitability of the Terrestrial Planets', *Ann. Rev. Astron. Astrophys.* **32**, 83–114.

Raval, A., Oort, A. H., and Ramaswamy, V.: 1994, 'Observed Dependence of Outgoing Longwave Radiation on Sea Surface Temperature and Moisture', *J. Clim.* **7**, 807–821.

Raval, A. and Ramanathan, V.: 1989, 'Observational Determination of the Greenhouse Effect', *Nature* **342**, 758–762.

Raymo, M. E. and Rau, G. H.: 'Mid-Pliocene Warmth: Stronger Greenhouse and Stronger Conveyor', *Science*, (submitted).

Rind, D., Chiou, E.-W., Chu, W., Larsen, J., Oltmans, S., Lerner, J., McCormick, M. P., and McMaster, L.: 1991, 'Positive Water Vapor Feedback in Climate Models Confirmed by Satellite Data', *Nature* **349**, 500–503.

Rind, D. and Peteet, D.: 1985, 'Terrestrial Conditions at the Last Glacial Maximum and CLIMAP Sea-Surface Temperature Estimates: Are They Consistent?', *Quat. Res.* **24**, 1–22.

Robock, A.: 1978, 'Internally and Externally Caused Climate Change', *J. Atmos. Sci.* **35**, 1111–1122.

Sellwood, B. W., Price, G. D., and Valdes, P. J.: 1994, 'Cooler Estimates of Cretaceous Temperatures', *Nature* **370**, 453–455.

Shackleton, N. J. and Boersma, A.: 1981, 'The Climate of the Eocene Ocean', *Geol. Soc. London J.* **138**, 153–157.

Shine, K. P., Derwent, R. G., Wuebbles, D. J., and Morcrette, J.-J.: 1990, 'Radiative Forcing of Climate', in Houghton, J. T., Jenkins, G. J., and Ephraums, J. J. (eds.), *Climate Change: The IPCC Scientific Assessment*, Cambridge University Press, New York, pp. 41–68.

Sloan, L. Cirbus: 1994, 'Equable Climates During the Early Eocene: Significance of Regional Paleogeography for North American Climate', *Geology* **22**, 881–884.

Sloan, L. Cirbus and Barron, E. J.: 1992, 'Eocene Climate Model Results: Quantitative Comparison to Paleoclimatic Evidence', *Palaeogeog., Palaeoclim., Palaeoecol.* **93**, 183–202.

Sloan, L. Cirbus and Rea, D. K.: 1995, 'Atmospheric CO_2 of the Early Eocene: A General Circulation Modeling Sensitivity Study', *Glob. Plan. Change*, (in press).

Sloan, L. Cirbus, Walker, J. C. G., and Moore, T. C., Jr.: 1995, 'The Role of Oceanic Heat Transport in Early Eocene Climate', *Paleoceanogr.* **10**, 347–356.

Sloan, L. Cirbus, Walker, J. C. G., Moore, T. C., Jr., Rea, D. K., and Zachos, J. C.: 1992, 'Possible Methane-Induced Polar Warming in the Early Eocene', *Nature* **357**, 320–322.

Sun, D.-Z. and Lindzen, R. S.: 1993, 'Distribution of Tropical Tropospheric Water Vapor', *J. Atmos. Sci.* **50**, 1644–1659.

Walker, J. C. G. and Sloan, L. Cirbus: 1992, 'Something is Wrong with Climate Theory', *Geotimes* **37**, 16–18.

Warren, S. G. and Schneider, S. H.: 1979, 'Seasonal Simulation as a Test for Uncertainties in the Parameterizations of a Budyko-Sellers Zonal Climate Model', *J. Atmos. Sci.* **36**, 1377–1391.

Webb III, T., Crowley, T. J., Frenzel, B., Gliemeroth, A.-K., Jouzel, J., Labeyrie, L., Prentice, I. C., Rind, D., Ruddiman, W. F., Sarnthein, M., and Zwick, A.: 1993, 'Group Report: Use of Paleoclimatic Data as Analogs for Understanding Future Global Changes', in *Global Changes in the Perspective of the Past*, John Wiley & Sons, Chichester, England, pp. 50–71.

Wigley, T. M. L. and Schlesinger, M. E.: 1985, 'Analytical Solution for the Effect of Increasing CO_2 on Global Mean Temperature', *Nature* **315**, 649–652.

Wolfe, J.: 1985, 'Distribution of Major Vegetational Types During the Tertiary', in Sundquist, E. T. and Broecker, W. S. (eds.), *The Carbon Cycle and Atmospheric CO_2: Natural Variations Archean to Present*, Geophys. Monogr. **32**, Am. Geophys. Union, Washington, D.C., pp. 357–375.

Zachos, J. C., Stott, L. D., and Lohmann, K. C.: 1994, 'Evolution of Early Cenozoic Marine Temperatures', *Paleoceanog.* **9**, 353–387.

Hoffert, M.I., and S.D. Potter: Energy Supply. In Watts, R. G., ed. *Engineering Response to Global Climate Change.* Lewis Publishers, Boca Raton, Florida, pp. 205-260 (in press).

Chapter 4

ENERGY SUPPLY

4.1 INTRODUCTION

Global warming projected for the next century by current climate models arises mainly from increased radiative heating by carbon dioxide (CO_2) accumulating in the atmosphere from the combustion of fossil fuels (Houghton et al., 1990; MacCracken et al., 1990). Also contributing to the predicted warming is deforestation (which can increase atmospheric CO_2 by transferring carbon from large standing carbon pools in the land biosphere) and from other greenhouse gases whose concentrations have increased along with the recent human population explosion, methane (CH_4), nitrous oxide (N_2O) and chlorofluorocarbons (CFCs). But the core greenhouse problem is carbon dioxide from fossil fuel burning.

The present world consumption rate of commercial energy is (Hammond, 1990, p. 316) 2.95×10^{20} J $y^{-1} = 9.35 \times 10^{12}$ W (9.4 TW). Some 95% of this is the heat of combustion of the fossil fuels (coal, oil and natural gas) burned each year, with the remaining 5% attributed to nonfossil primary electricity production (hydro, nuclear and geothermal — values of nonfossil electrical energy are based on the heat value of their electricity: 1 kW-hr = 3.6×10^6J). Though it is not included in many statistical compilation, some 1.3 TW additional is probably consumed as "noncommercial" energy, mainly wood burned for fuel outside cash economies in developing nations (Goldemberg et al., 1988, p.192). Thus, fossil fuels currently supply some 90% of a total world energy consumption of 10.7 TW.

Our civilization is beneficiary of a massive boon of nature: the fossil fuel reserve. Fossil fuels have accumulated in the earth's crust over hundreds of millions of years from organic carbon produced by photosynthesis — a small fraction of which escaped oxidation by burial in peat bogs (coal) or marine sediments (gas and oil). Some 5.7×10^{12} kilograms of carbon are transferred from the lithosphere to the atmosphere each year (5.7 Gt C y^{-1}) by fossil fuel combustion as CO_2 gas. At this rate, reserves of economically recoverable natural gas and oil

(04) Chapter 4 7/31/96 1:23 AM Page 206

will be gone in the next 50 years or so, though coal could continue to provide energy for several hundred years. Clearly, the convenience, cleanliness and transportability in pipelines of gas and oil favor their use over coal. If depletion of oil and gas were the problem, a synthetic fuel technology in which gaseous and liquid hydrocarbons were derived from more abundant coal (and perhaps shale) would seem a viable solution. Humanity would have more than a century to face the need for nonfossil energy sources. Until recently, such a synfuel strategy was conventional wisdom in the energy industry. Global warming has changed the game by raising the specter of an early transition from fossil fuels, possibly as soon as the beginning of the 21st Century.

How much of this could be produced by nonfossil sources if we really had to? How much of the increased energy demand of the next century could be produced by nonfossil sources? And how rapidly could it be deployed?

4.2 ENERGY AND GLOBAL CHANGE

Global energy supplies are constrained by technological, economic and, increasingly, environmental considerations. In contrast to the technological optimism prevailing in the early part of this century are present-day concerns over undesirable environmental impacts generated by technology — photochemical smog, nuclear reactor accidents, toxic wastes, acid rain, oil spills, deforestation, ozone depletion, and endangered species, to name a few. In addition to engineering feasibility and economic viability, contemporary criteria for evaluating any new technology include its compatibility with natural ecosystems and sustainability over time.

At the United Nations Framework Convention for Climate Change in Rio de Janeiro in June 1992, representatives of the Earth's human populations agreed to "stabilization of greenhouse gas concentrations in the atmosphere that would prevent dangerous anthropogenic interference with the climate system (Bolin. 1994)." There is no consensus yet on what constitutes "dangerous interference." Carbon cycle models indicate that major fossil fuel emission reductions may be needed to stabilize atmospheric CO_2 and climate to acceptable levels (Harvey, 1990; Jain et al., 1994; Wigley, 1995). "Integrated Assessments" of

(04) Chapter 4 7/31/96 1:23 AM Page 207

Global Change — which link the causal chain of emissions to climate to environmental impact and perform cost/benefits analyses for different mitigating strategies — are only now being developed (Naki´cenovi´c et al., 1994). Despite uncertainties, it is within the range of outcomes that fundamental changes in worldwide energy consumption and production will be needed to mitigate adverse impacts of climate change from humankind's greenhouse gases.

4.2.1 Fossil Fuel Reserves and the Carbon Cycle

The viability of energy alternatives depends on the amount, cost and environmental impact of fossil fuel reserves. Although fossil energy resources and reserves are fairly well known, real uncertainties exist. And there is sometimes an apples versus oranges confusion in the literature from imprecise definitions of the terms "energy resources" and "energy reserves."

The *total resource base* of an energy source is defined here as the combination of undiscovered and identified, subeconomic as well as economic, concentrations of naturally occurring solid, liquid and gaseous materials in the Earth's crust; the *accessible resource* is that subset of the resource base that can be captured, mined, or extracted by current technology or technology which may be available in the near future (regardless of economics); and the *proven reserves* is that subset of the accessible resource which is identified and can be economically and legally extracted to yield useful energy (SERI, 1990). Oil and gas fuels, for example, are formed at elevated temperatures in the subsurface by the transformation of the organic remains of marine organisms, and coal from undecayed land plants (Tissot and Welte, 1978). The amount of buried organic carbon (kerogen) is thousands of times greater than oil and gas reserves, but it is too diffusely distributed to be economically viable. Similarly, there is much more uranium in the Earth's crust than what is economically extractable as powerplant fuel.

Table 4.1 summarizes energy and carbon contents of proven reserves of coal, oil and natural gas and uranium-235 (the fuel of conventional fission reactors). The total carbon content of the fossil energy reserve based on these numbers is ~ 2500 Gt C (1 Gt C = 10^{15} g C), excluding shale. Present estimates of oil shale are 4–5 times the crude oil reserve (Tissot and Welte, 1978, p. 235), or ~ 500 Gt C of carbon additional. The proven fossil fuel carbon reserve based on pre-

(04) Chapter 4 7/31/96 1:23 AM Page 208

sent technology is therefore ~ 3000 Gt C. We estimate that the accessible resource could likely double to ~ 6000 Gt C given improvements in mining and extraction technologies available in the 21st Century. Historical energy consumption records indicate ~ 220 Gt C have already been transferred from the lithosphere to the atmosphere and oceans since preindustrial times as fossil fuel CO_2 (Sundquist, 1993).

For conservation of mass, the sources, sinks and accumulation of elemental carbon cycling through the Earth System must balance instantaneously and over the long term. Since each 2.13 Gt C remaining in the atmosphere increases atmospheric CO_2 concentration 1 ppm, the observed increase in atmospheric CO_2 concentration from its preindustrial value ~ 279 ppm (594 Gt C) to its present value ~ 354 ppm (754 Gt C) (Boden et al., 1991) represents ~ 160 Gt C added to the atmosphere. Current carbon cycle models indicate the oceans absorbed an additional ~ 140 Gt C since preindustrial times (Sundquist, 1993). The sum of atmosphere and ocean carbon "sinks" is therefore ~ 300 Gt C. This is not quite enough to balance the estimated ~ 380 Gt C that has entered the atmosphere/ocean system from carbon "sources" from 1750-1990: ~ 220 Gt C from fossil fuel burning plus ~ 160 Gt C from changes in land use (Sundquist, 1993).

If the ~ 80 Gt C excess of sources over sinks is real (larger than

Table 4.1 Proven energy reserves and their carbon content (Hammond, 1990, p. 320. With permission.).

Energy Source [size]	Energy Content [10^{21} J]	Energy Content [GW-yr]	Carbon Content [1 Gt C = 10^{15} g C]
hard coal [2.77 x 10^6 mtce]	77.3	2,450,000	1900
soft coal [1.27 x 10^6 mtce]	17.7	561,000	440
oil [0.124 X10^6 mtoe]	5.2	165,000	104
natural gas [0.109 x 10^6 bcm]	4.2	133,000	57
uranium [2.34 x 10^6 tu] (recoverable at < $130 kg)	1.4	44,000	0
Totals	105.8	3,353,000	2501

[a] Energy conversion factors: 1 million tonnes oil equivalent = 41.87 PJ; 1 billion cubic meters natural gas = 38.84 PJ, 1 million tonnes coal equivalent (hard coal) = 27.91 PJ, 1 million tonnes coal equivalent (soft coal) = 13.96 Pl; 1 tonne uranium metal = 0.58 PJ (based on 0.72% ^{235}U isotope abundance) (10^6 PJ = 10^{21} J = 31,700 GW-yr).

[b] carbon conversion factors: 1 PJ oil = 20.0 Gt C; 1 PJ natural gas = 13.6 Gt C; 1 PJ coal = 24.6 Gt C.

(04) Chapter 4 7/31/96 1:23 AM Page 209

uncertainties in source and sink estimates), then some "missing sink" is needed to balance the carbon budget. The most likely candidate is increased carbon uptake by terrestrial biomass fertilized by higher CO_2 and nitrate in the environment.

The ratio of CO_2 remaining in the atmosphere to the total CO_2 input from 1750-1990 (the mean airborne fraction) was ~ 160/380 ~ 42%. In general, the airborne fraction depends on the time-dependence of the source function. It can go as high as 90% or more if emissions increase faster than the oceans absorb excess CO_2. The asymptotic airborne fraction approached after anthropogenic CO_2 emissions have ceased does not approach zero but a value in the range of 13–30% depending on the amount of carbon injected to the system (Maier-Reimer and Hasselmann, 1987). Still, the 220 Gt C emitted by fossil fuel burning thus far is a small fraction of the carbon in the fossil fuel reserve. Given the pressures of population and economic growth, burning of the global fossil energy reserve to depletion seems a likely outcome in the absence of global energy supply alternatives.

4.2.2 How Much Fossil Fuel Can be Burned?

Scenarios for rapidly burning fossil reserves to depletion typically have CO_2 emission rates increasing from the present ~5.6 Gt C y^{-1} four or five times by the year 2100, and then declining over the next 300 years such that ~ 6000 Gt C are transferred from the lithosphere to the atmosphere. This carbon represents the present fossil reserves plus a fraction of the accessible reserves that might be economically exploited in the next century. According to current carbon cycle models, such a rapid burning scenario will produce ~ 1200 ppm atmospheric CO_2 by the year 2100, with a subsequent peak of ~ 2200 ppm by the year 2200, followed by a slow decline to ~ 800 ppm by the year 5000 (Walker and Kasting, 1992). These values of CO_2 are large enough to trigger a global greenhouse unprecedented since the Cretaceous era 100 million years ago when the Earth may have been ~ 10°C warmer (Hoffert, 1990).

Even "freezing" the global fossil fuel burning emission rate at the present rate ~5.6 Gt C y^{-1} would result in CO_2 levels > 500 ppm by 2100, and this would continue to rise until the fossil fuel supply is depleted (Krause et al., 1989; Wigley, 1995). And for such a constant emission scenario, only ~540 Gt C would be burnt over the next hundred years — only 10% of the 21st century's accessible reserve.

(04) Chapter 4 7/31/96 1:23 AM Page 210

Walker and Kasting (1992) explored as a possible longterm goal whether it might be possible to hold atmospheric CO_2 below 500 ppm while still exhausting fossil fuel reserves. They found that sustainability in this long-term sense would only be possible if emissions were reduced by a factor of 25 — not the factor of two sometimes quoted in the press. The conclusion is inescapable that keeping atmospheric levels below 500 ppm would probably require switching to nonfossil energy supplies long before fossil fuel reserves are depleted.

The environmental impacts of global warming are still controversial, whereas the economic stake in the present energy system is well-established and enormous. Accordingly, energy response strategies proposed thus far have emphasized improved energy efficiency and switching from high CO_2 to lower CO_2 emitting fossil fuels, as opposed to abandoning fossil fuel technology entirely. In the coming decades, as increases in carbon emissions come predominantly from the developing world, the issue of equity is likely to loom importantly.

It is possible that improved end use efficiency and switching from high to low CO_2 emitting fossil fuels could mitigate some of humanity's greenhouse gas emissions and their greenhouse warming. Energy conservation, efficiency increases and fuel switching to lower emitting fossil fuels can also buy time as human populations shift to renewable solar, fission breeder reactors or fusion energy (Harvey, 1990). But it is only prudent to vigorously research nonfossil global energy systems as a hedge against the worst case scenarios.

4.3 ENERGY SUPPLY AND WEALTH

A historic goal of nation-states is to increase, or at least to maintain, their economic productivity, as measured by Gross National Product (GNP) or (more recently) Gross Domestic Product (GDP). GNP is the value of all goods and services produced by a country measured in a convertible currency; GDP is similar, but includes the effect of trade imbalances with other countries. The present Gross World Product (GWP) is ~ \$21 trillion (21×10^{12} \$US); with the United States the largest national economy (GDP ~ \$5.6 trillion) followed by Japan (GDP ~ \$3.3 trillion) (Hammond, 1994).

Although economic activity is increasingly globalized, it remains

(04) Chapter 4 7/31/96 1:23 AM Page 211

useful to consider the existing nation-states as a set of data points from which properties of large-scale (macro) economies can be derived. Among other things, this approach supports macroeconomics theories that take a thermodynamic perspective (emphasizing *production* of goods) over neoclassical economics (emphasizing *exchange* of goods according to subjective human preferences).

Cleveland et al. (1984) rationalize energy-based macroeconomics as follows: "Production is the economic process that upgrades the organizational state of matter into lower entropy goods and services. Those commodities are allocated according to human wants, needs and ability to pay. Upgrading matter during the production process involves a unidirectional, one-time throughput of low entropy fuel that is eventually lost (for economic purposes) as waste heat. Production is explicitly a work process during which materials are concentrated, refined, and otherwise transformed. Like any work process, production uses and depends on the availability of free energy. The laws of energy and matter control the availability, rate, and efficiency of energy and matter use in the economy and therefore are essential to a comprehensive and accurate analysis of economic production." And consequently, the economic productivity of nations should be directly linked to their energy consumption. This is observed (Figure 4.1).

Whereas most economic models are open-ended with regard to economic growth, GDP is clearly constrained by the supply of energy and other resources, and ultimately by environmental degradation (Meadows et al., 1992). The Malthusian question is how these constraints work and whether they can be indefinitely overcome by technology. The limits to growth are not obvious. In a pioneering study, Revelle (1985) estimated that bringing an additional 300 million hectares under cultivation for food production and 1000 million hectares for biomass energy plantations would support global population growth from the present 5.3 billion to an estimated 10 billion by the middle of the next century — citing the main limitation as soil erosion. It is now evident that such expansion of cultivated land into natural ecosystems could result in major loss of biodiversity — particularly since unprecedented species extinctions have already been created by humanity's land use appropriations thus far (Wilson, 1992; Hammond, 1994).

4.3.1 Economic Productivitv of Energy

The developed world has essentially stabilized its population and

(04) Chapter 4 7/31/96 1:23 AM Page 212

enjoys the highest per capita income in human history. But because of the lopsided distribution of wealth between rich and poor nations — exacerbated by rich versus poor disparities in population growth rates — the coming decades may see conflicts between economic growth by the developing world and preserving the environment from adverse effects of fossil-fuel-induced climate change. These objectives might be at least partly reconciled by increasing the economic productivity of energy (Goldemberg et al., 1988; Redi and Goldemberg, 1990).

Mid-range projections by the Intergovernmental Panel on Climate Change based on United Nations data show global population rising from the present 5.3 billion to 8.4 by 2025 and 11.3 by 2100, and CO_2 emissions rising from the present 5.7 billion tonnes carbon per year to 10.7 by 2025 and 19.8 by 2100 (IPCC scenario IS92a; Leggett et al., 1992). In this scenario ~ 95% of population growth and ~ 75% of carbon emissions growth by the year 2025 comes from developing countries. By the middle of the next century the world population will be ~ 10 billion, of which less than 15% — some 1.4 billion people — will live in presently developed countries and 8.7 billion will live in presently less-developed countries. The relationship between energy supply and economic development is critical for developing a global strategy to reduce greenhouse gas emissions.

For the global population as a whole, present per capita GDP, energy consumption and carbon emission rates are

$$\text{GDP} = \frac{21 \times 10^{12}\ \$\text{US/yr}}{5.3 \times 10^{9}\ \text{people}} \sim 4000\ \$\text{US yr}^{-1}\text{-person}^{-1} \qquad \text{(Eq. 4.1)}$$

$$\text{energy consumption} = \frac{10.7 \times 10^{12}\ \text{W}}{5.3 \times 10^{9}\ \text{people}} \sim 2\ \text{kW person}^{-1} \qquad \text{(Eq. 4.2)}$$

$$\text{carbon emissions} = \frac{5.7 \times 10^{9}\ \text{tonne carbon}}{5.3 \times 10^{9}\ \text{people}} \sim 1\ \text{tonne C person}^{-1}\ \text{yr}^{-1} \qquad \text{(Eq. 4.3)}$$

These averages are important reference points. To address the energy supply problem we need to also consider the *distribution* of wealth and energy consumption over the Earth's nation-states.

(04) Chapter 4 7/31/96 1:23 AM Page 213

Figure 4.1 shows per capita GDP (GDP/N) versus per capita energy consumption (E/N) of United Nations member states plotted on a log-log scale.

That logarithmic scales spanning more than two orders of magnitude are needed to display the data is a measure of the inequity in the distribution of wealth of nations. For example, the mean energy consumption rate of the United States is about 10,000 W (~ 10 kW) while that of Bangladesh is only 60 W (~ 0.06 kW). For scenario analysis, countries are often aggregated into classes based on regions or economic groups; or simply into "developed" (primarily, the U.S., Japan and Western Europe plus the Former Soviet Union) and "developing" (everyone else). Goldemberg et al. (1988, p. 302) cite the present E/N of developed and developing nations as 6.3 and 0.54 kW, respectively.

Whereas global mean CO_2 carbon emissions are ~ 1 tC per person, the rich, industrial nations presently produce much more than their share of this greenhouse gas (Figure 4.2). Inhabitants of the United States emit 5 times as much carbon as the global mean. Part of the differences in per capita emission is due to differences in wealth, part is associated with differences in land use (deforestation, agriculture, etc.), and part can be attributed to differences in energy efficiency.

The correlation between per capita Gross Domestic Product (GDP/N) and per capita energy consumption (E/N) on an individual nation basis is evident in Figure 4.1. Let the ratio P = GDP/E = (GDP/N)/(E/N) be the *economic productivity of energy.* P is the wealth created by a unit of raw energy consumption as hypothesized by energy theories of macroeconomics, and is influenced by the end-use energy conversion efficiency. It applies to renewables as well as fossil-fuel-based energy sources. The economic productivity of energy for the world as a whole based on present numbers is

$$P \sim \frac{\frac{4000\ \$US}{\text{person-yr}}}{\frac{2\ kW}{\text{person}}} \times \frac{1\ yr}{8760\ hr} \times \frac{100\ ¢}{1\ \$US} \sim 23\ ¢/\text{kW-hr.} \qquad \text{(Eq. 4.4)}$$

The solid line in Figure 4.1 is the linear equation (GDP/E) = P x (E/N) for P = 23¢ GDP/kW-hr. Countries above this line are more efficient at converting energy to GDP than the world average (France,

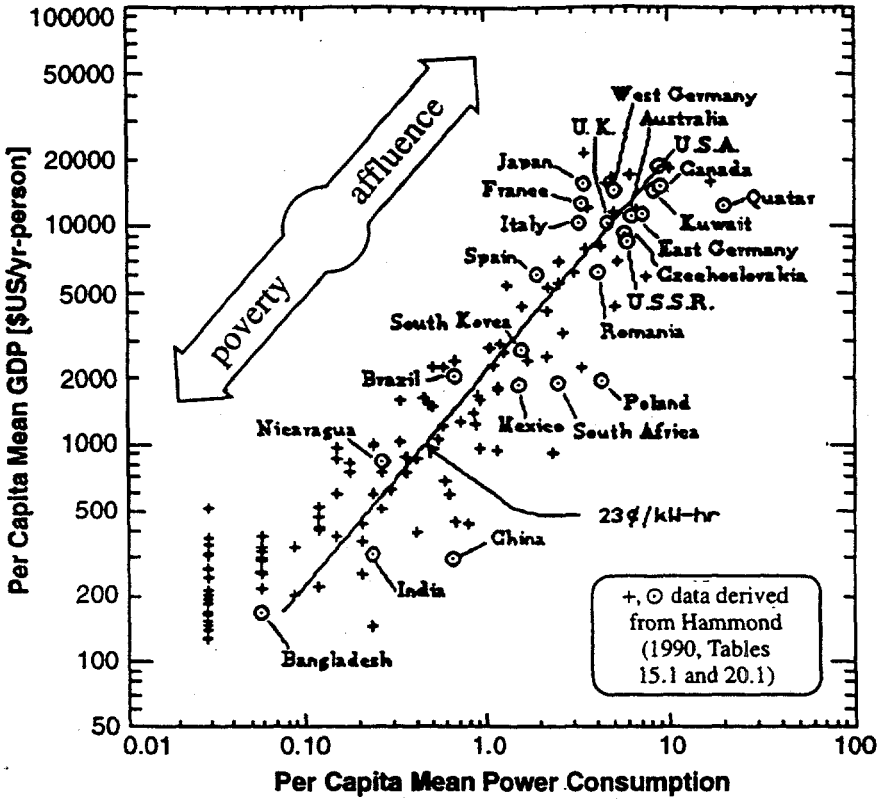


Figure 4.1 Distribution of per capita GDP (GDP/N) versus per capita raw commercial energy consumption (E/N) for members of the United Nations. Named countries are denoted by circles, unnamed by crosses.

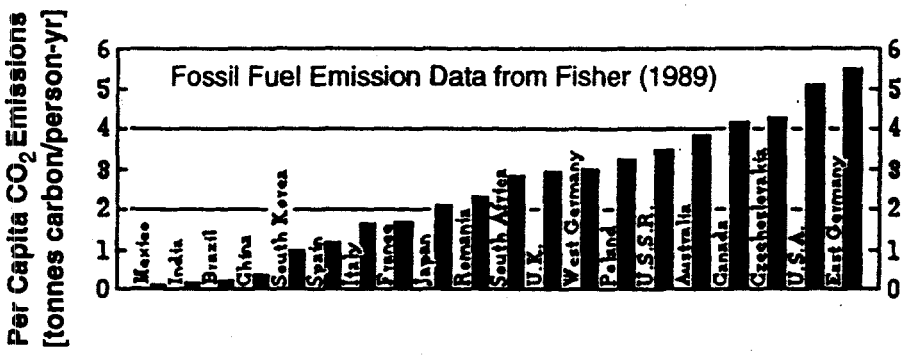


Figure 4.2 Per capita CO_2 emissions in tC person from fossil fuel burning for some nations shown in Figure 4.1.

(04) Chapter 4 7/31/96 1:23 AM Page 215

Japan, West Germany), countries below the line less efficient (USSR, Poland, China). Most of the variance over three log cycles (factor of a thousand) is accounted for by the GDP-energy correlation with a residual factor-of-two scatter associated with nation-to-nation differences in energy-to-GDP efficiency.

A legacy of our hydrocarbon economy is that energy (and energy per unit GDP) is often expressed as equivalent oil amounts[1]. Crude oil is a mixture of petroleum liquids and gases in various combinations. However, conversion factors between barrels or metric tons of "oil" and energy have become standardized (Häfele et al., 1981):

1 barrel oil equivalent (42 gallons) =
1 boe = 6.12×10^9 J = 1700 kW-hr

1 metric ton oil equivalent =
1 toe = 44.76×10^9 J = 12,430 kW-hr

The energy consumed in producing a unit of GDP — the inverse of the economic productivity of energy P — is the *energy intensity* I = E/GDP = (E/N)/(GDP/N) = P^{-1}. The global mean energy intensity in metric tons of oil equivalent per thousand US dollars for the world economy is

$$I \sim \frac{1\ \text{kW-hr}}{23\,¢} \times \frac{1\ \text{mtoe}}{1.243 \times 10^4\ \text{kW-hr}} \times \frac{100\,¢}{1\ \$US} \times \frac{1000\ \$US}{\$1000\ US} \sim 0.35 \frac{\text{mtoe}}{\$1000\ US}$$

(Eq. 4.5)

The value of I differs for individual nations as they move through different phases of economic development, and differs among nations at any given time. Such I-variations are often displayed to analyze effects of end-use efficiency. (Note that economies which are more efficient at converting raw energy use to GDP have *lower* values of I.)

Figure 4.3 shows that as Western Europe and North America developed, their energy consumptions grew faster than their GDPs. This is normally attributed to the building of industrial infrastructures. Roads, bridges, homes and heavy industry involved capital investment

[1]In Europe, oil amounts are reported in metric tons (tonnes); in Japan in kiloliters (kl). But in the United States and Canada, and colloquially throughout the world, the basic unit is the "barrel" (bbl). In his historical account, Yergin (1991) relates that when oil began flowing from the wells of Western Pennsylvania in the 1860s, "...desperate oil men ransacked farmhouses, barns, cellars, stores and trashyards for

(04) Chapter 4 7/31/96 1:23 AM Page 216

in materials that did not yield immediate GDP increases. The variation since 1840 of both I and P is shown (on vertical log scales) for the U.S., U.K., West Germany, France and Japan. The present global mean P = 23 ¢/kW-hr is also shown for reference. During their infrastructure-building periods, P dipped as low as 8–10 ¢/kW-hr in England and the U.S., although France and Japan were able to avoid P < 15 ¢/kW-hr. (Note that peaks in I are valleys in P.)

Advocates of efficient energy end-use technology to reduce CO_2 emissions from developing nations observe that materials can now be produced with less energy, and that smaller quantities of modern materials can replace larger amounts of older ones (Reddy and Goldemberg, 1990). Consequently, developing nations may be capable of achieving comparable levels of industrialization at a lower ratio of consumed energy to GDP growth. This is represented by the shaded area projections for developing nations. The critical questions regarding end-use efficiency are quantitative. Consider the optimistic scenario of Figure 4.3 in which the global mean P increases by a factor of two by 2040 to, say, 46¢/kW-hr. This could cut per capita emissions by 50% at constant GDP even with continued (but more efficient) fossil fuel energy technology.

But halving per capita emissions would essentially be canceled by the projected doubling of world population fifty years hence. And this does not address economic growth or "equity" issues. A factor of two from energy efficiency cannot by itself bridge the gap between rich (developed) and poor (developing) nations. To do this new global scale nonfossil fuel energy supplies will be needed.

To address equity issues, some redistribution of per capita energy from rich to poor countries is implied. Presumably, GDP could be maintained in rich nations with declining per capita energy consumption by improving enduse efficiency. Goldemberg et al. (1988) advocate a "no-global-energy-growth" scenario in which per capita energy consumption falls in developed countries from 6.3 to 3.5 kW by the

any kind of barrel — molasses, beer, whiskey, cider, turpentine, salt, fish, and whatever else was handy. But as coopers began to make barrels specially for the oil trade, one standard size emerged, and that size continues to be the norm to the present. It is 42 gallons. The number was borrowed from England, where a statute in 1482 under King Edward IV established 42 gallons as the standard size barrel for herring in order to end skullduggery and divers deceits in the packing of fish... By 1866, seven years after Colonel [Edwin L.] Drake drilled his well, Pennsylvania producers confirmed the 42-gallon barrel as their standard..." Few petroleum workers have ever seen a "barrel," except in a museum.

(04) Chapter 4 7/31/96 1:23 AM Page 217

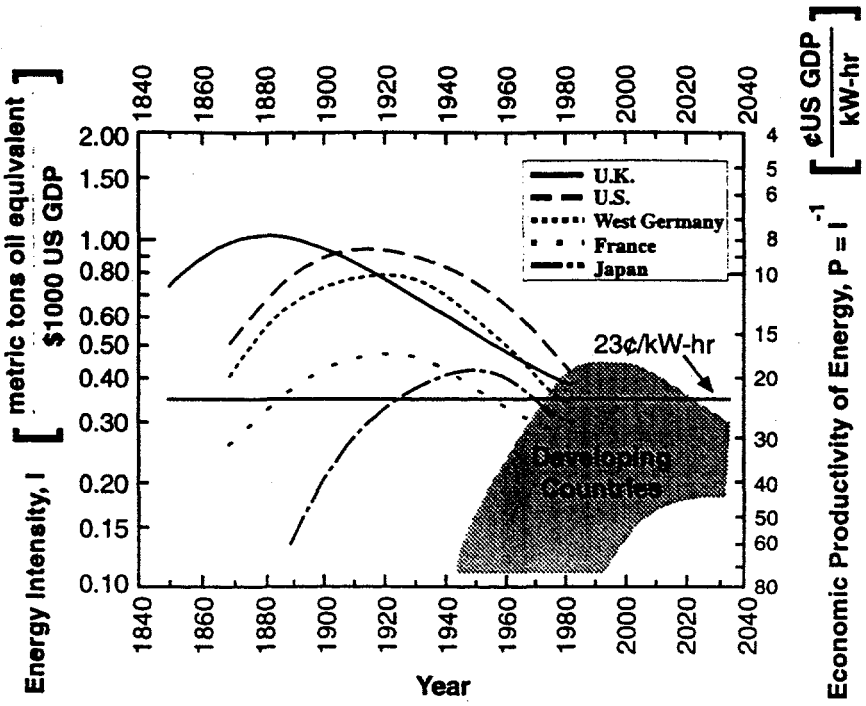


Figure 4.3 Evolution of Energy Intensity, I, (left hand scale) and Economic Productivity of Energy, P, (right hand scale) for selected nations. Because of improvements in materials science and energy efficiency, the maxima in I (minima in P) reached by developed countries during industrialization have decreased over time. An objective of energy end-use strategies for developing countries is to avoid repeating the history of the industrialized world by using energy more efficiently (replotted from Reddy and Goldemberg, 1990. With permission.).

year 2020, and rises in developing countries from 0.54 to 1.2 kW. These authors also consider — and discount, as requiring unavailable capital investment funds — a scenario with global energy growth of 2% per year at constant per capita energy use by developed nations. At a 2% overall energy growth rate, energy consumed per capita in the developed world can remain at 6.3 kW and rise in the developing world to 2.2 kW. But global energy consumption must increase to 19.4 TW by the year 2020.

There is no way to set goals for the future economic and environmental state of the Earth without invoking human values. The morality of a world in which energy resources are shared more equitably by nations seems positive enough to move in that direction. But

(04) Chapter 4 7/31/96 1:23 AM Page 218

a zero sum game in which the poor get richer at the expense of the rich getting poorer is politically naive and may be akin to re-shuffling deck chairs on the Titanic. A better strategy is to combine increases in the economic productivity of energy with new energy supplies and modest per capita energy redistributions. Still, there is no optimum path independent of subjective human values.

We therefore propose the goal of a global economy by the year 2040 (20 years after Goldemberg et al.'s projections) with a population of 10 billion consuming 20 TW in which CO_2 emissions are cut 50% relative to the present. Global per capita energy consumption would stay the same as today's 2 kW. If P rose by a factor of two or more from improvements in energy efficiency, the Gross World Product could more than double, most of which could be distributed to less-developed nations. Five TW could still be generated by fossil fuel combustion; but global nonfossil energy technologies would have to be in place fifty years from now capable of producing 15 TW of thermal power free of greenhouse gas emissions — 150% the present global consumption rate.

4.3.2 Fossil Fuel Costs, Carbon Taxes and the Cost of Electricity

To assess the global scale power supply it is necessary to consider engineering feasibility and cost. We can identify at this time only a few energy systems with the capacity to power human civilization. Will economics determine whether any of these will replace fossil fuels?

Classical economics holds that technologies are in a kind of Darwinian competition with each other in which the "invisible hand" of the market acts as natural selection. But the winners are not preordained. A particular energy technology like the light water nuclear reactor is no more inevitable than the elephant or hummingbird. More recent theories view the global economy as a complex adaptive system which is constantly making predictions based on its internal models of the world. If the greenhouse model of climate change is accepted by governments, that model would itself influence the development of energy supply technologies, regardless of whether the climatic changes predicted for fossil fuel burning would have occurred. This is no different from the way the economy may work in other regards.

John Holland of the Santa Fe Institute put it this way (Waldrop, 1992, p. 147): "...Complex adaptive systems have many *niches*, each

(04) Chapter 4 7/31/96 1:23 AM Page 219

one of which can be exploited by an agent adapted to fill that niche. Thus, the economic world has a place for computer programmers, plumbers, steel mills, and pet stores, just as the rain forest has a place for tree sloths and butterflies. Moreover, the very act of filling one niche opens up more niches — for new parasites, for new predators and prey, for new symbiotic partners. So new opportunities are always being created by the system. And that, in turn, means that it's essentially meaningless to talk about a complex adaptive system being in equilibrium: the system can never get there. It is always unfolding, always in transition... And by the same token, there's no point in imagining that the agents in the system can ever 'optimize' their fitness, or their utility, or whatever. The space of possibilities is too vast; they have no practical way of finding the optimum. The most they can ever do is change and improve themselves relative to what the other agents are doing."

Practically speaking, an optimum energy strategy for the world may not exist. Even if one could define a theoretical optimum utility function it would be impossible to predict. But there are better and worse solutions. And just as in evolutionary biology, where not all organisms are possible, there are physical and economic constraints on energy supplies. A practical upper limit on the cost of a global-scale energy source is when it becomes more expensive than the wealth it produces. Below this limit — and one seeks to stay as much below this limit as possible — incentives like carbon taxes can make nonfossil energy sources competitive earlier and thereby help limit CO_2 emissions.

Convenience, high energy density, relative cleanliness and ease of transportation have led oil to dominate as the leading fossil energy source in recent years. Since production began in 1861, crude oil prices have fluctuated with market pressures from < $5/bbl in the mid-1960s to ~ $20/bbl after the Arab Oil embargo of 1973 to as much as ~$40/bbl in the 1979-1981 Oil Panic (inflation-corrected prices in 1990 $US/bbl from Yergin, 1991, p. 785). The price of oil also varies day-to-day, and depends on the source, which is why it is traded as a commodity. At present, oil is relatively inexpensive. Crude prices quoted in the *New York Times* of Thursday, June 13, 1996 are:

	Wed.	Tue.
Saudi Arabian light $ per bbl fob	16.60	16.57

(04) Chapter 4 7/31/96 1:23 AM Page 220

North Sea Brent $ per bbl fob	18.25	18.22
West Texas Intermed $ per bbl	20.08	20.13
Alaska No. Slope del. US Gulf Coast	18.60	18.71

For illustrative purposes, assume the cost of unprocessed fossil fuel is CF ~ 20 $US/boe. In cents per kilowatt-hour of chemical (or thermal) energy this representative fossil fuel cost works out to

$$CF \sim \frac{20\ \$US}{1\ boe} \times \frac{1\ boe}{1700\ kW\text{-}hr} \times \frac{100\ ¢}{1\ \$US} \sim 1.18\ ¢/kW\text{-}hr\ .$$

(Eq. 4.6)

The ratio of the global mean productivity of energy to the cost of fossil fuel energy is then P/CF ~ (23 ¢/kW-hr)/(1.18 ¢/kW-hr) ~ 20. In other words, the present productivity of energy implies a twenty fold return on investment to the global economy from fossil fuel consumption. Conversely, the world's consumption of fossil fuels at the rate of ~ 5% of the Gross World Product (GWP) is > $1 trillion/yr — a prodigious spending rate. It is not surprising that the relatively few multinational corporations dealing in oil (and often other forms of fossil energy) have annual incomes exceeding the GDPs of most developing nations.

Fossil fuels are the most cost-effective global-scale energy sources available at present if "externalities" like the adverse environmental impact of global warming are ignored. The science of climate impact analysis is too uncertain at this point to provide reliable environmental costs as a basis for carbon taxes. Indeed, some argue that higher CO_2 levels should be assigned a credit associated with enhanced agricultural yields. But as we have discussed previously, it is within the range of outcomes that serious disruption to global ecosystems could occur if more than a small fraction of the fossil fuel reserve is burned.

In any event, nonfossil fuel energy technologies will significantly penetrate global energy markets only if their costs come down and/or fossil fuel costs go up. A policy to make alternative energy technologies more competitive is the imposition of carbon taxes (Edmonds and Reilly, 1985, 1986; Nordhous, 1991; National Academy of Sciences, 1992). Taking oil as a representative fossil fuel (1 mtoe = 0.89 tC = 12430 kW-hr), each $10/t$CO_2$ carbon tax increases the fuel cost per

(04) Chapter 4 7/31/96 1:23 AM Page 221

unit thermal energy by

$$\$10/\,tCO_2 \sim \frac{10\,\$US}{tCO_2} \times \frac{44\,tCO_2}{12\,tC} \times \frac{0.89\,tC}{12430\,kW\text{-}hr} \times \frac{100\,¢}{1\,\$US} \sim 0.26\,¢/kW\text{-}hr. \quad \text{(Eq. 4.7)}$$

Carbon taxes would fall differently on different fossil fuels. Relative to oil (= 1.00), carbon emissions per unit energy are 1.23 for coal and 0.68 for natural gas. A tax of $10 per tonne of CO_2 emitted would therefore be ~ 0.32 ¢ for coal, ~ 0.26¢ for oil and ~ 0.18¢ for gas for the same kilowatt-hour of thermal energy. Among other things, carbon taxes tend to deplete oil and gas reserves earlier than a tax-free policy.

Based on the oil standard, a tax of ~ $10/t$CO_2$ would raise the "fossil fuel" price ~ 20% (from ~ 1.2 to ~ 1.4¢/kW-hr) whereas ~ $100/t$CO_2$ would more than triple it (from 1.2 to 3.8¢/kW-hr). Of course, these costs are representative, and a comprehensive analyses should include additional market variables; e.g., coal is cheaper than oil, but has a higher emission factor; synthetic fuels derived from coal are more expensive; and so on. Econometric models including these details indicate that a 50% reduction in fossil fuel carbon emissions would require taxes at the margin in the range 27–164 $/t$CO_2$ (100-600 $/tC) (Nordhous, 1991; Manne and Richels, 1992; Scheraga and Leary, 1992). An average carbon tax of $46/t$CO_2$ would double fossil fuel costs (CF + CTAX ~ 2.4¢/kW-hr), half the ratio of the productivity of energy to-the fuel cost (P/(CF + CTAX) ~ 10) and provide revenue of over a trillion $US worldwide (~ 5% of GWP).

It is not clear who would collect carbon taxes or to what use they would be put. If governments were the agents, the balancing of national budgets, health care, etc., are likely candidates. Alternately, multinational energy corporations might be more than willing to act as *de facto* tax collectors by raising prices, as OPEC did in 1973 — particularly if antitrust litigation could be avoided. A more rational policy from our point of view would be to assign carbon tax revenue to research, development and demonstration of nonfossil fuel energy sources.

There are precedents for subsidizing alternative energy technologies. Grubb and Meyer (1993) analyzed some lessons from the last energy crisis: "The rapid development of wind power in California in

(04) Chapter 4 7/31/96 1:23 AM Page 222

the early 1980s was a product of 'sledgehammer' promotion. Generous federal tax credits, combined with state incentives, produced large subsidies for wind installations. In addition, the Public Utilities Policy Act mandated that utilities buy energy at its full avoided cost from independent generators, thus ensuring a market for wind power... Installation rates in California increased six-fold in one year, rising from 10 MW in 1981 to 60 MW in 1982, and by 1984 had reached 400 MW per year. By 1986, the cumulative investment in wind energy totaled about $2 billion, with the value of energy generated put at $100 million per year."

But rapid development of wind energy has its cautionary tales. Grubb and Meyer relate that, "...Many of the machines were of poor quality and broke down during the first season of operation. As one manufacturer complained, some of the early companies knew more about tax minimization than they did about engineering. Machines were often sited carelessly, and some were sold based on blatantly fraudulent promises... Such slipshod technology, as well as visions of thousands of motionless machines threatened to destroy wind power in the United States. However, several companies, aided by favorable state-financing packages, invested heavily in wind energy technology and helped advance the field significantly. The total net cost of the program to the California economy (minus fuel savings) is estimated to be about $500 million, but without such a program, it is unlikely that wind energy would have reached its current stage." The message to anyone viewing cost-effective wind farms today — for example, at the Altamont Pass near the Lawrence Livermore National Laboratory in California — is that tax incentives to alternative energy technologies can work if proper safeguards are in place.

A recent, if modest, threat to the fossil fuel infrastructure is the requirement in certain states that automobile manufacturers begin selling zero emission vehicles (ZEVs). Again, California is taking the lead. It has mandated that by 1998, 2 percent of all vehicles sold be ZEVs, rising to 10 percent by 2003. ZEVs are being promoted not to reduce CO_2 emissions, but to reduce photochemical smog precursors (hydrocarbons and nitrogen oxides). If electricity used to charge an electric car's batteries were produced by fossil fuel burning, the net effect would be an *increase* in CO_2 emissions. It is normally more efficient to use internal combustion engines in motor vehicles than to drive electric generators in remote fossil fuel-fired powerplants, trans-

(04) Chapter 4 7/31/96 1:23 AM Page 223

mit the electricity over power lines, charge onboard vehicle batteries, and deliver rotational energy to wheels via electric motors.

If the goal is smog reduction with reasonable performance, a compromise could be hybrid electric-internal combustion. Such cars would generate electricity onboard from a combination of internal combustion and regenerative braking. They require smaller batteries than electric cars while eliminating transmission line losses. In the future, hydrogen-powered vehicles might do even better. But with today's technology, "zero emissions" means battery-powered vehicles with low stored energy-to-mass ratios and high costs. To meet ZEV quotas, manufacturers would presumably have to spread these costs over their nonelectric fleets. This has fostered protests by auto makers and oil companies that electric cars are being subsidized contrary to normal marketplace competition.

Historically, many key technologies in the commercial sector have been stimulated by government policies. The Interstate Highway System initiated by President Dwight D. Eisenhower is a continuing federally-supported investment in the internal combustion engine. Commercial jet aircraft, computer chips and the Internet are spin-offs of military research. The tax-subsidized development of wind energy in California begun in the Energy Crisis years of the late 1970s helped to make wind the most cost-effective renewable energy source today. More recently, the Strategic Defense Initiative (SDI) of the Reagan and Bush administrations fostered remote sensing and space technologies, which has been tested in a joint DoD/NASA mission to orbit the Moon by the Clementine space probe. It has been argued that a Strategic Environment Initiative (SEI) including tax incentives for new technologies and disincentives for old could likewise focus technological talent on environmentally appropriate energy technologies (Gore, 1992).

In any event, a transportation system aimed at slowing the growth of atmospheric CO_2 should employ a primary power source that minimizes CO_2 emissions. It is better, for example, to make H_2 for hydrogen-powered cars by water electrolysis using nonfossil fuel electricity than to use H_2 from stream-reformed methane — 99% of present-day industrial hydrogen is derived from natural gas, essentially methane. From a carbon emissions point of view, the goal of ZEVs should be to build an energy infrastructure in which nonfossil electricity (and/or hydrogen) is a universal energy carrier.

(04) Chapter 4 7/31/96 1:23 AM Page 224

Methodologically, cost comparisons of fossil fuel energy with or without carbon taxes versus renewables, fission, fusion and solar power satellites imply such a universal energy carrier.

Electricity is a thermodynamically higher quality of energy than heat. Most of the world's electricity is produced by fossil or nuclear heat engines with H_2O (water and steam) as the working fluid. In the classical Rankine cycle, superheated steam generated in a boiler passes through a turbine (driving an electrical generator) to a condensor, and is returned to the boiler as water. The advantage relative to a gas turbine is that it is cheaper to pump liquid from a low-pressure condensor to a high-pressure boiler than to compress a gas. However, the efficiency of a vapor/liquid cycle is lower than Carnot because cold fluid plunged into the hot boiler vapor mixes irreversibly.

Consider a typical boiler providing steam to a turbine at ~ 200 bars and T_{max} ~ 540 °C (813 K), after which it enters a condenser at ~ 2 bars and T_{min} ~ 60 °C (333 K). The Carnot efficiency is $(T_{max} - T_{min})/T_{max}$ ~ 59%. But even with regenerative heaters to improve performance, the theoretical efficiency of this steam turbine cycle is ~ 45% (Faires and Simmang, 1978, p. 258). Assuming efficiencies of 95% each for combustion-to-thermal and turbine-to-electrical power, the overall electrical conversion efficiency of a fossil fuel powerplant is thus only η_e ~ 40%. In that case, 1 kW-hr (thermal) ~ 0.4 kWe-hr (electrical). A typical generating plant fired by oil at \$20/boe (1.18 ¢/kW-hr) has a fuel cost per unit electrical power of (1.18/0.4) ~ 3.0 ¢/kWe-hr exclusive of capital costs.

The levelized cost of electricity (COE) is the weighted average cost of energy production in ¢ per kWe-hr including capital investment recovery, fuel and nonfuel operating costs projected over the facility's economic lifetime,

$$COE = IC\left(\frac{FCR + OMR}{DUTY}\right) + \left(\frac{CF + CTAX}{\eta_e}\right). \qquad \text{(Eq. 4.8)}$$

Here IC is the initial capital cost of the plant (\$ per kWe installed x 100 ¢/\$), FCR is the fixed charge rate and OMR the operation and maintenance rate (% of IC per yr), DUTY is hrs per year on-line (% time operational x 8760 hr.), CF is the cost of fuel and CTAX the carbon tax (¢ per kW-hr), and η_e the electrical conversion efficiency (the

(04) Chapter 4 7/31/96 1:23 AM Page 225

dimensionless ratio). The first term contains all plant costs except fuel and carbon taxes. We chose values for a reference oilfired powerplant such that the first term is ~ 3.6 ¢/kWe-hr and COE (no carbon taxes) ~ 3.6 + 3.0 ~ 6.6 ¢/kWe-hr — close to the 7¢/kWe-hr cited as the average cost of electricity in the U.S.

Table 4.2 shows the effect of carbon taxes on the cost of electricity computed for our hypothetical average fossil-fuel powerplant. The Low (0–10 $/tCO$_2$), Med (10-100 $/tCO$_2$) and High (100-200 $/tCO$_2$) categories are similar to those employed by the National Academy of Sciences (1992) in their study of mitigation of greenhouse warming. Apart from their effect on reducing CO_2 emissions by encouraging energy conservation, carbon taxes at the 100 $/tCO$_2$ level would essentially double the cost of electricity to ~ 13¢/kWe-hr. Compare, for example, the cost of fossil fuel electricity in our Table 4.2 example with the cost of energy from renewable sources estimated by the Department of Energy in their report to the U.S. Congress in Table 4.3.

Based on these numbers, geothermal, wood-fired electric, hydroelectric and wind are all nearly cost-effective today, and would be significantly cheaper than fossil fuel electricity at 13¢/kWe-hr. Notice in Table 4.3 that terrestrial solar thermal and photovoltaic electricity cannot provide cost-effective electricity today even with High carbon taxes, though they might be by the year 2030. However, the potential magnitude of an energy source is important. Even carbon taxes in the Med to High range would be insufficient to drive fossil fuels from the market unless their replacements could collectively supply the ~ 10

Table 4.2 Effect of carbon tax on cost of electricity (COE). Calculations assume an oil-fired powerplant @ $20/boe, initial capital cost = $1500/kWe, fixed charge rate = 15%/yr, operation & maint. = 3%/yr, duty cycle = 85%; hence, generating cost per electrical energy output (less fuel cost) = 3.6¢/kWe-hr. Also assumed are electricity/fuel conversion efficiency = 40% and carbon emission factor = 0.122 tC/boe = 0.89 tC/toe.

	Carbon Tax		Energy Costs fuel+Ctax		electricity
Tax Category	$/tCO$_2$	$/tC	¢/kW-hr	¢/kWe-hr	¢/kWe-hr
None	0	0	1.2	3.0	6.6
Low	0–10	0–37	1.2–1.4	3.0–3.5	6.6–7.2
Med	10–100	37–370	1.4–3.8	3.5–9.5	7.2–13.1
High	100–200	370–730	3.8–6.4	9.5–16.0	13.1–19.6

(04) Chapter 4 7/31/96 1:23 AM Page 226

Table 4.3 Projected levelized cost of electricity (COE) for terrestrial renewable energy technologies in current US cents/kWe-hr (after Bradley et al., 1991, Appendix E. With permission).

Technology	1990	2030
Geothermal	5.5	4.4
Wood-fired electric	7.2	7.2
Hydroelectric	8.3	8.3
Wind	9.5	5.0
Solar thermal electric	18.7	9.4
Photovoltaics	54.3	5.1

TW of global demand.

To summarize: Low cost, high energy density, and availability will tend to continue humanity's reliance on fossil fuels in the near term unless there are incentives to change. Although there is no "optimum" energy supply, a nonfossil energy infrastructure capable of powering human civilization at > 20 TW will be needed sooner or later — perhaps as early as the middle of the next century.

Displacing fossil fuels by alternate energy sources prior to depletion of recoverable oil, gas and coal would be fostered by carbon taxes. A worldwide investment as large as ~ 5% of GWP might be needed to cut CO_2 emissions 50% — on the order of $1 trillion per year. Carbon taxes could work only if they were international — particularly if revenue were targeted to research, development and demonstration of new technologies. Developed nations could probably reduce emissions by a factor of two by well-understood improvements in energy conversion. But new ideas, technology and financing are needed to provide greenhouse-free energy to the developing world where 95% of global population growth is projected by the year 2040. An appealing vision is the implementation of a "green energy" infrastructure compatible with the biological species diversity of the planet, and sustainable over time scales comparable to recorded human history.

4.4 TOWARD A NONFOSSIL FUEL CIVILIZATION

(04) Chapter 4 7/31/96 1:23 AM Page 227

It is well-known that cost estimates of innovative energy technologies (for example, those in Table 4.3) are problematical because of the difficulty of estimating economies of scale and future manufacturing efficiencies. A technological optimist could, for example, cite dramatic decreases in unit costs of very large scale integrated circuits (VLSI) in personal computers and consumer electronics. Few would have predicted in 1970, for example, that transistors would be virtually free today — you can now buy a transistor for 4,000-millionths of a cent (Gilder, 1993). A pessimist could cite erroneous predictions by the Atomic Energy Commission of the 1950s that nuclear power would be "too cheap to meter." Is there a methodology that can help policy makers make choices regarding alternate energy technologies?

Eminent scientists have made notoriously wrong predictions of technological change. John von Neumann, father of the modern electronic computer and one of the 20th Century's great mathematicians made two such predictions in the early 1950s: first, that in the future computers would become so complex that only governments could afford them, and second, that computers would be able to make accurate long-range weather forecasts (Kaku, 1994). He based his predictions on linearly extrapolating trends at the time, and he was wrong on both counts: Computers are now inexpensive enough to have become home appliances, whereas the weather is known to be statistically unpredictable more than three to five days in the future. In the first instance, von Neumann missed the advent of quantum electronic devices like VLSI; in the second, he missed the significance of nonlinearities on the predictability of the fluid dynamical equations governing atmospheric motion (chaos).

In the short run — a few years — it is reasonable to quantitatively extrapolate existing technologies linearly. On the decadal to century time scales of global warming it is more likely that qualitative breakthroughs will become dominant, and that new industries will open up in unexpected places. One reason why the system behaves nonlinearly is that technologies form "clusters" that reinforce each other and create whole new capabilities. Jesse Ausubel (1994), who studies the impact of technology change on global warming, realizes this: "...Imagining how the clusters will affect lifestyles and restructure the economy, and thus affect emissions and vulnerability to climate, is a tremendous intellectual challenge."

Science fiction writers may have more intuitive feeling for nonlin-

(04) Chapter 4 7/31/96 1:23 AM Page 228

ear technology forecasting than establishment scientists. A recently discovered manuscript by Jules Verne written in the 1860s, *Paris in the 20th Century*, was rejected for publication in its time because it pictured a future too strange to be credible (Riding, 1994): In this work, Verne imagined a future in the 1960s where people traveled by subway and in gas-driven cars, where they communicated by fax and telephone, where they used computers, and where "electric concerts" provided entertainment. In this world, everyone could read but no one read books. It was a society dominated by money where destitute homeless people roamed the streets. Strange indeed.

In contrast to what is often assumed in studies of global warming mitigation, energy technologies of the 21st century are likely to differ from those of the 20th century as much as 20th century technology differed from 19th. If the past is any guide, global warming mitigation studies that assume progressive decarbonization of fossil fuels by switching from coal to oil to methane, increasing end use efficiency of energy conversion devices, and even the more widespread use of existing renewable energy technologies (wind, terrestrial solar and biomass) will prove too conservative. Polls taken in the 1920s predicted that we would have, within a few decades, huge fleets of blimps taking passengers across the Atlantic (Kaku, 1994). The reality was that Lindbergh crossed the Atlantic in 1929 in the heavier-than-air *Spirit of St. Louis,* while the *Hindenberg* explosion of 1938 effectively ended air travel by blimps. The progress of theoretical aerodynamics and propulsion technology during World War II, particularly the invention of the aviation gas turbine, led to the present era of commercial jet aviation as the dominant long-distance transport mode. Given their high frontal area and drag, there is no way that lighter-than-air craft can compete for high-speed passenger service. Again, the technology of transportation turned on the physics.

What we have seen in the past century is that technological progress was determined more in the long run by limitations and opportunities stemming from fundamental scientific principles than by naive extrapolations of existing technology. It seems fruitful, therefore, to assess 21st Century nonfossil fuel energy technologies as imaginatively as possible, but always constrained by our understanding of the physics. The laws of physics determine what *cannot* be achieved — nonconservative energy sources, perpetual motion machines, efficiencies higher than those allowed by thermodynamics,

and (probably) "cold fusion" — but they do not guarantee what *can*. They are necessary but not sufficient conditions. What can be done is likely to be constrained by available energy fluxes and reservoirs, innovations in energy transmission and storage, and the impact of harvesting and redistributing energy flows on the Earth's environment and ecosystems.

4.4.1 Renewable Energy

Many (but not all) sustainable energy sources are forms of solar energy. The Sun is a normal main sequence star (spectral class G2) powering the Earth's atmospheric and oceanic circulations, hydrological cycle and biosphere. It radiates energy derived from hydrogen to helium fusion in its core at a slowly increasing rate — the Sun is more than 25% brighter now than when the solar system formed ~ 4.6×10^9 yr ago. Its present radius, $R_{sun} \approx 6.96 \times 10^8$ m, and luminosity, $L_{sun} = 3.90 \times 10^{26}$ W, correspond to an energy flux at the solar surface of $F_{sun} = L/(4\pi R_{sun}^2) = 64.1 \times 10^6$ W m^{-2} (Allen, 1976). It follows from the blackbody radiation law ($F = \sigma T_{eff}^4$ where $\sigma \approx 5.67 \times 10^{-8}$ Wm^{-2} K^{-4} is the Stefan-Boltzmann constant) that the effective temperature of the Sun's visible surface (the photosphere) is $T_{sun} = (F_{sun}/\sigma)^{1/4} \approx$ 5800 K.

Most of the energy radiated from Sun to the universe as a whole is, from our point of view, wasted. Solar flux on a surface perpendicular to the Sun's rays decreases inversely with the square of the distance from its center. By the time solar photons have spread to the Earth's mean orbital distance, a = 1 A.U. = 1 astronomical unit $\approx 1.5 \times 10^{11}$m, the flux per unit area (solar constant) has dropped to $S_0 = F_{sun} \times (R_{sun}/a)^2 \approx 1380$ W m^{-2}.

4.4.1.1 Terrestrial Solar Energy

We are mainly concerned in this section with energy fluxes available at the surface of the Earth, of area $4\pi R_E^2 = 5.1 \times 10^{14}$ m^2, where $R_E \approx 6.37 \times 10^6$ m is the Earth's radius. The *hectare* is a convenient SI unit employed in scientific literature for areal measure (1 ha = 10^4 m^2). Land, ocean and total surface areas in 10^6 ha are (Sverdrup et al., 1970):

$$A_l = 15{,}000 \times 10^6 \text{ ha}$$
$$A_w = 36{,}000 \times 10^6 \text{ ha}$$

(04) Chapter 4 7/31/96 1:23 AM Page 230

$$A_E = 51{,}000 \times 10^6 \text{ ha.}$$

Changing land use patterns, as the human population grew from 600 million in 1700 to 4,430 million in 1980, based on recent estimates, and assuming that "badlands" (deserts, ice sheets, etc.) remained constant at 11% of A_1, are shown in Table 4.4.

Despite a "Green Revolution" in crop yields subsidized by energy in the form of fertilizers and farm machinery, humankind's agriculture has grown from a few percent of global land use preindustrially to ~

Table 4.4 Global land use in 10^6 ha from 1700 to 1980 (adapted from Richards, 1990, Table 10-1. With permission.).

Year	Forests & Woodlands	Grassland & Pasture	Deserts & Icesheets	Agricultural Croplands
1700	6,215	6,860	1,661	264
1850	5,965	6,837	1,661	537
1920	5,678	6,748	1,661	913
1950	5,389	6,780	1,661	1,170
1980	5,050	6,788	1,661	1,501

10% today, primarily at the expense of temperate and tropical forests.

Per unit surface area of the spinning Earth, solar flux incident at the top of the atmosphere is $S_o/4 = 345$ W m^{-2}, of which 70% is absorbed by the atmosphere, clouds and surface: That is, $(S_o/4)(1-\alpha) \approx 242$ W m^{-2} is absorbed by the atmosphere/earth system. The factor of 4 is the ratio of the Earth's surface to disk area, and $\alpha = 0.30$ is the Earth's visible reflectivity (albedo). Energy conservation requires solar energy absorbed to be balanced in the steady state by radiative cooling to space at the blackbody temperature, $T_{eff} = [(S_o(1-\alpha)/(4\sigma)]^{1/4} \approx 255$ K. Radiative cooling is affected by infrared absorbing atmospheric gases (H_2O, CO_2, O_3,...). These warm the surface ~ 33 K above T_{eff} (the greenhouse effect). For the Earth as a whole, $T_s \approx 288$ K. In addition, differential absorption of solar energy at high and low latitudes creates an equator-to-pole surface temperature gradient which drives atmosphere and ocean circulations (Peixoto and Oort, 1992).

Of the 345 W m^{-2} solar flux at the top of the atmosphere, only 51% reaches the surface. The rest is absorbed by the clear sky and clouds or reflected back to space (Wallace and Hobbes, 1977, p. 321). Thus $0.51 \times (S_o/4) = 180$ W m^{-2} is a typical solar flux at the surface. Because of cloudiness, seasonality and the rotation of the Earth, there

(04) Chapter 4 7/31/96 1:23 AM Page 231

is substantial intermittency in the local values of surface solar insolation. $F_s \approx 180$ W m^{-2} should be considered a long-term average over diurnal and seasonal cycles and over the Earth's entire surface area. A fraction of this flux is available, in principle, for humankind's energy end uses: space heating, transportation, industrial processes and electricity.

Apart from land, the direct conversion of solar flux to electricity or fuel places stringent requirements on materials. These must absorb the solar flux to produce electronic excitations, separate the electron-hole pairs making up the exciton, and convert the separated charge either into electrical energy in an external circuit or into chemical fuel. These steps occur in nature (in photosynthesis) as well as human technology (in photovoltaic, or PV, solar cells). There is a common body of theory linking these processes (APS, 1979).

A key material property for solid state solar energy conversion is the *bandgap energy*, $\varepsilon_o = h\nu_o$, where $h \sim 6.63 \times 10^{-34}$ J-s $\sim 4.14 \times 10^{-15}$ eV-s is Planck's constant and ν_o the bandgap frequency. The monochromatic efficiency of a single bandgap absorber is (Schockley and Queisser, 1961; Thorndike, 1976):

$$\eta(\nu,\nu_o) = \left\{ \begin{array}{ll} (\nu_o/\nu); & \nu \geq \nu_o \\ 0 & \nu < \nu_o \end{array} \right\},$$

(Eq 4.9)

where ν is the frequency of light. Light from a monochromatic laser at the bandgap frequency ν_o is absorbed with an efficiency of 100%. But more energetic photons can only transfer a fraction of their energy, and less energetic photons are not absorbed at all.

The high efficiency of monochromatic light absorption by PV cells might be exploited in special applications where it is desirable to transmit a narrow beam of energy over long line-of-sight distances. For example, a recent proposal to power a lunar outpost during exploratory missions to the Moon is to beam the energy from Earth-based sites using high-powered lasers tuned to the bandgap frequency of a PV receiver on the lunar surface (Matthews et al., 1994).

For the (multicolored) spectral distribution of sunlight per unit frequency $I_\nu(\nu)$ available in the natural environment, the efficiency of photon energy transfer is

(04) Chapter 4 7/31/96 1:23 AM Page 232

$$\bar{\eta}(\nu_o) = \frac{\int_0^\infty \eta(\nu,\nu_o) I_\nu(\nu) d\nu}{I},$$

(Eq. 4.10)

where $I = \int_0^\infty I_\nu(\nu) d\nu$ is the frequency-integrated sunlight.

Power engineers (and natural selection) favor materials with bandgap energies (ε_o or ν_o) that optimize the integrated energy absorption efficiency, $\bar{\eta}(\varepsilon_o)$.Typical bandgap energies are $\varepsilon_o \sim 1.1$–1.6 eV for single and polycrystalline silicon (Si), gallium arsenide (GaAs-GaAlAs) and related PV materials, and $\varepsilon_o \sim 1.8$ eV for green plant chlorophyll photosynthesis (APS, 1979). The theoretical peak efficiency of photon absorption for normal sunlight at the Earth's surface is ~ 35% (Thorndike, 1976, pp. 33-35), while the realized peak efficiencies of PV cells is ~ 17% for Si crystals and ~ 22% for GaAs-GaAlAs crystals (APS, 1979). In principle, more efficient cells could be created using multiple light-absorbing layers of different bandgap energies (Vos, 1980).

In practice, commercialization of PV power is limited more by manufacturing costs than by energy efficiency, as such. In the case of silicon, the raw material (sand) is virtually free and yet the current cost of PV electricity is > 50 ¢/kWe-hr (Table 4.3). Recent efforts to lower PV costs emphasize amorphous (as opposed to crystalline) semiconductors, which have slightly lower efficiencies but are cheaper to fabricate (Ovshinsky and Madan, 1983).

Liquid hydrocarbon fuels are perhaps the greatest challenge to replace with solar energy. Because of its high energy per unit mass, hydrogen derived from water electrolysis has many advantages as an energy carrier, although a prime power source is, of course, needed, and problems of onboard vehicle storage remain (Winter and Nitsch, 1988). As early as 1923, J.B.S. Haldane suggested electrolytic hydrogen production from wind power; in 1927 A.J. Stuart proposed hydrogen from hydroelectricity; and in 1960, J. O'M. Bockris proposed solar power as an energy source for hydrogen electrolysis (Bockris, 1980). Advantages of solar hydrogen are that it buffers intermittent and spatially nonuniform sunlight. Its main disadvantage is high cost

(04) Chapter 4 7/31/96 1:23 AM Page 233

(on top of high PV costs).

Ogden and Williams (1989) advance an ingenious argument for "mining" the Earth's crust for amorphous silicon solar power versus mining it for uranium and thorium to fuel breeder reactors: "Even though the energy released in the fissioning of a single uranium nucleus is 100 million times greater than the 'energy released' when a photon is absorbed in amorphous silicon, a uranium atom can fission only once, whereas a silicon solar cell can repeatedly absorb photons and convert solar energy into electricity. An amorphous solar cell contains an amorphous silicon layer about 1 micron thick, amounting to some 3 grams of silicon per square meter of cell area. A 15-percent efficient PV system operated in the southwestern United States, where the insolation averages about 250 watts per square meter, would thus produce about 3300 kW-hr per gram of silicon over the expected thirty-year PV system life — about the same as the amount of electricity from a gram of nuclear feedstock using breeder reactors."

Assuming cost and materials problems can be solved, the fundamental problem of terrestrial solar energy is low power density. If we assume that it will eventually be possible to divert 5% of solar flux at the surface without adversely affecting the global environment, then ~ 10 W m^{-2} could be available for human use by solar thermal powerplants and PV cells. The heat rejected would in any event be radiated to space, but could be redistributed by global scale solar energy systems. Radiative forcing by greenhouse gases emitted by humankind is already ~ 2.5 W m^{-2} (Shine et al., 1990, Table 2.6). Note also that solar flux available for human use as electricity is further reduced by low efficiencies of solar collectors and converters and by the area devoted to storage and transmission of this intermittent resource. In the case of PV cells, F_{pv} ~ 1 W m^{-2} is a reasonable electricity output flux, not counting storage.

4.4.1.2 Biomass Energy

Ecosystems of the terrestrial and marine biospheres are energized by solar photons (hν) converted to chemical energy in green plants, algae and cyanobacteria. The photosynthesis reaction is $h\nu + CO_2 + H_2O \rightarrow CH_2O + O_2$, where CH_2O is organic carbon (shorthand for 1/6 of a glucose molecule, $C_6H_{12}O_6$) and $\Delta H = 467$ kJ $mol^{-1} = 3.89 \times 10^4$ J gc^{-1} is the heat of reaction (Bolton and Hall, 1991). Net primary productivities (NPP, photosynthesis less respiration) per unit

(04) Chapter 4 7/31/96 1:23 AM Page 234

area of real ecosystems are in the range (NC/A) = 200–2000 gC m^{-2} y^{-1} (Schlesinger, 1991, Table 5.2) corresponding to NPP efficiencies in the range $\eta_p = [(N_C/A)\Delta H]/F_s \sim 0.1–1.3\%$.

The theoretical maximum efficiency of photosynthesis has been estimated by Bolton and Hall (1991) at $\eta_{p,max} \approx 9\%$ (independent of the cell or ecosystem in which it occurs). This is virtually impossible to achieve in real organisms because it ignores plant respiration, reflected light, limitations due to inadequate water and nutrients and functional requirements of plant growth and physiology.

What is impressive is that ecosystems with sufficient water and nutrients like swamps and rainforests have evolved NPPs within an order-of-magnitude of the theoretical peak. Tropical rainforests are vertically organized to waste as few photons as possible. Broad-crowned, widely scattered emergent trees tower 50 m or more above the main canopy. The densest foliage, containing many coexisting tree species, is frequently at heights of 20-30 m. Below this second tier, small plants energized by scattered, diffuse sunlight take their places in vertical sequence — treelets, shrubs, and finally, herbs (Terborgh, 1992). These are complex and fragile ecosystems, home to a great diversity of biological species — how many is unknown, but it is of the order of millions. Tropical rainforests can impact nontropical environments through the hydrological cycle, atmospheric chemistry and in other important but poorly understood ways (Dickinson, 1987). But unchecked, the expansion of human populations through land use conversion could effectively destroy the tropical rainforests and make countless species extinct in the next century (Table 4.4).

In the process of photosynthetic organic carbon production, trees remove CO_2 from the atmosphere, virtually all of which is regenerated by respiration when the trees decay. The primary productivity of the terrestrial biosphere collectively is of the order of 100 Gt C y^{-1} (Schlesinger, 1991). Global biomass is close to a steady state over the long term, but if the standing carbon crop in trees decreases or increases, the result is a transient carbon source or sink to the air/sea system. Current rates of tropical deforestation are believed to be releasing ~ 2.4 Gt C y^{-1} on top of ~ 5.6 Gt C y^{-1} from fossil fuel burning (Houghton et al., 1990).

Meyers and Goreau (1991), and others, have proposed CO_2 absorption by rapidly growing tropical tree plantations to mitigate carbon dioxide buildup by fossil fuel combustion and deforestation. But car-

(04) Chapter 4 7/31/96 1:23 AM Page 235

bon sequestration is a one-shot use of land whose scarcity is likely to increase in the 21st century. Such land use also runs counter to perceived economic aspirations of developing nations in the tropics (Marland, 1991) — not to mention the threat to the biodiversity of indigenous ecosystems by monoculture farms.

Hall et al. (1991) and Marland (1991) observed that biomass farms are preferable to CO_2 sequestration by unharvested trees because the land use is renewable. This is a valid argument up to a point. Although the combustion of wood (and its alcohol fuel derivatives) produces CO_2, biomass farms are an essentially greenhouse-free form of solar energy because virtually all the CO_2 recycles rapidly regardless of whether the tree is burned for fuel or decays naturally. But the threat to biodiversity and competition with agriculture remains; and biomass energy plantations have the additional need to restore soil nutrients depleted by the harvest. For a realistic 1% energy efficiency by biomass plantations, we estimate $F_{BIO} \sim 2$ W m^{-2} which drops to ~ 1 W m^{-2} after conversion to electricity in (say) wood-fired power plants. This does not include the energy needed to farm the trees (nutrients, extraction, transportation, etc.).

4.4.1.3 Windpower

Atmospheric winds driven by temperature gradients from differential heating by the Sun are another form of solar energy. The kinetic energy flux per unit *vertical* cross sectional area is $(1/2)\rho_a U^3$, where ρ_a is air density and U horizontal wind speed. The large-scale circulation is driven by the equator-to-pole temperature gradient $\Delta T \approx 45$ K interacting with Coriolis forces produced by the Earth's spin ($\Omega = 2\pi$ radians/day $\approx 7.3 \times 10^{-5}$ s^{-1}) embodied in the Coriolis parameter, $f = 2\Omega\sin(\text{latitude}) \sim 10^{-4}$ s^{-1}. The result is strong high-altitude winds in the East-West (zonal) direction at midlatitudes (the jet stream).

The vertical gradient of the zonal wind is (Peixoto and Oort, 1992, p. 156), $\partial\langle u\rangle/\partial z \sim (g/fT_s)\,\partial T/\partial y \sim g\Delta T/(fT_s R_E)$, where z is altitude, y the meridional coordinate and $g \approx 10$ m s^{-2} the gravitational acceleration. Zonal winds in the jet stream near the tropopause altitude, $H_t \sim 8 \times 10^3$ m, are of order $\langle u_{jet}\rangle \sim gH_t\Delta T/(fT_s R_E) \sim 20$ m s^{-1}. Unfortunately, such high wind speeds (and very high kinetic energy fluxes) eight kilometers up are inaccessible to current energy technology, whereas mean surface winds are much less intense (Oort, 1983), $\langle u_s\rangle \sim 3$ m s^{-1}. Surface air density, $\rho_a \approx 1.2$ kg m^{-3} and $\langle u_s\rangle \sim 3$ m s^{-1}

(04) Chapter 4 7/31/96 1:23 AM Page 236

give a kinetic energy flux $F_w \approx (^1/_2)(1.2)(3)^3 \sim 16$ W m^{-2}. But surface windspeed is highly variable, and the mean kinetic energy flux (proportional to $\langle u_s^3 \rangle$) at a site is usually much higher than the flux at the mean windspeed $\langle u_s \rangle$. Typically (Grubb and Meyer, 1993), $\langle u_s^3 \rangle \sim 2\langle u_s \rangle^3$. Also, current wind farm practice suggests a ratio of wind turbine disk area to horizontal area $A_F/A_H \sim {}^1/_{10}$.

Combining these estimates gives a mean wind energy flux in the atmospheric boundary layer per unit *surface* area $F_w \sim (A_F/A_H)\rho_a \langle u_s \rangle^3 \sim 3$ W m^{-2}. The major environmental impacts of wind turbines are their land and materials requirements and their need to be integrated into the power production infrastructure.

4.4.1.4 Ocean Thermal Energy

The oceans are a source of renewable energy. A most promising technology is Ocean Thermal Energy Conversion (OTEC) — floating heat engines grazing the surface-to-deep temperature gradient. The vertical mean ocean temperature gradient occurs mainly in the first kilometer of ocean depth (the thermocline). A typical horizontally-averaged temperature difference is $\Delta T_w = T_{surface} - T_{deep} \sim 15$ K. OTECs date to 1881, when d'Arsonval first proposed a closed-Rankine-cycle device in which a working fluid is vaporized by heat exchange with cold seawater drawn through a vertical pipe from 700-1200 meter depth (Dugger et al., 1983).

Because of the small temperature gradient, and the need to divert power to the cold water pump, the energy conversion efficiency of OTECs is only a few percent. On the positive side, OTECs bring nutrients, as well as cold water, to the surface, offering along with energy the possibility of aquiculture (food) (Savage, 1994). An upper limit on power available is the rate heat is transferred to the surface from deep-water by upwelling. Nature pumps cold high-latitude seawater to the deep ocean by polar downwelling at a volumetric rate of (Tolmazin, 1985, Table 7.2) $Q_w \sim 34 \times 10^6$ m^3 s^{-1}, after which the cold water upwells into the world's oceans. The thermal power of this flow is $\rho_w C_p Q_w \Delta T_w$, where $\rho_w \sim 1030$ kg m^{-3} and $C_p \sim 4000$ J kg^{-1} K^{-1} are the density and heat capacity of seawater, so the upward power flux per unit ocean area is $(\rho_w C_p Q_w \Delta T_w)/A_w \sim (1030)(4000)(34 \times 10^6)(15)/(3.6 \times 10^{14}) \sim 5.8$ W m^{-2}. This upward flux of cold water is balanced in the steady state by downward flux of warm surface water by vertical mixing. The balance determines thermocline structure.

(04) Chapter 4 7/31/96 1:23 AM Page 237

Were the entire 5.8 W m^{-2} thermal flux exploited by OTECs, it could destroy (make isothermal) the thermocline structure. Marine ecosystems might not survive. But diverting a fraction, say 10%, of the thermocline energy flux might be sustainable. A 10% diversion would make ~ 0.6 W m^{-2} available as OTEC input power, with fertilization of surface ocean plankton and fisheries a bonus. For a realistic energy conversion efficiency of 3%, electric power output flux harvestable over the oceans is F_{OTEC} ~ 0.02 W m^{-2}.

4.4.1.5 Hydroelectricity

The largest present-day form of commercial renewable energy is hydroelectricity. Driven by differential solar heating, water in liquid, vapor and solid phases cycles through clear air, clouds, rain, snow, rivers, seawater, sea ice and glaciers. The hydrological cycle, among other things, produces an excess of precipitation over evaporation over elevated land surfaces, which runs off to the seas. Volumetric runoff by all by the world's rivers is (Budyko, 1974, p. 228; Tolmazin, 1985, Table 7.3) $Q_r \sim 1.3 \times 10^6$ m^3 s^{-1}.

Apart from flowrate, hydropower depends on the "total head," $H = h + V^2/(2g)$, where the first term on the right is gravitational potential energy and the second kinetic energy (both in units of elevation). Mean elevation above sea level of the Earth's land surfaces is (Sverdrup et al., 1970, p. 19) $H_1 \approx 840$ m. An approximate first-principles estimate of global hydropower per unit land area is therefore $(\rho_w g H_1 Q_r/A_1) \sim (1000)(10)(840)(1.3 \times 10^6)/(1.5 \times 10^{14}) \sim 0.07$ W m^{-2}. Assuming 20% can actually be converted to electricity gives $F_{HDRO} \sim 0.014$ W m^{-2}, smaller in potential than most renewables but the one most exploited today. Although hydropower is concentrated in rivers (ideally with large changes in elevation, or head), the proper conditions do not often occur naturally. Niagara Falls is an exception.

It is typically necessary to create artificial reservoirs with dams by diverting rivers and flooding existing lands to tap hydropower, with potentially adverse environmental impacts. A recent example is the Three Gorges hydroelectric project on the Chang Jiang (Yangtze) River in China, approved by the National People's Congress in April, 1992. Chinese officials estimate that the reservoir will partially or completely inundate 2 cities, 11 counties, 326 townships, and 1,351 villages. About 23,800 hectares of cultivated land will be submerged, and 1.1 million people will have to be resettled. The Yangtze dolphin,

(04) Chapter 4 7/31/96 1:23 AM Page 238

one of the world's most endangered species, would be further threatened, as would other aquatic life — not to mention the destruction by the dam and reservoir of some of China's finest scenery. On the positive side, the project's 9.6 GW output is expected to displace 40-50 million tons of coal per year (Hammond, 1994, p. 68).

4.4.1.6 Geothermal Energy

It is now feasible to tap some of the earth's interior heat with steam turbines at "hot spots" where underground steam is produced by the proper geologic conditions. Steam geysers like "Old Faithful" in Yellowstone Natural Park are indicative of such locales. In general, upwelling geothermal heat comes mainly from radioactive decay of uranium, thorium and potassium isotopes (^{238}U, ^{235}U, ^{232}Th, ^{40}K) in the Earth's mantle — 80% of the heat flux is from radioactive decay, 20% from cooling of the primordial Earth. On a global mean basis, the geothermal flux is substantially lower than that associated with solar energy. Measured fluxes are .0505 W m^{-2} in the continental crust and 0.0782 W m^{-2} in the (thinner) oceanic crust, with a global mean of ~ 0.07 W m^{-2} (Turcotte and Schubert, 1982, Table 4.1).

Geothermal heat is accessible to present-day technology only when there is an existing hydothermal-convection system that transfers heat to the surface by underground circulating water and steam. Ideally, one seeks geothermal steam fields such as those in New Zealand associated with the Tonga deep ocean trench in the western Pacific. These are normally found in volcanic zones near geologic faults occupying a few percent at of the Earth's area (Shepard et al., 1977). Moreover, some 80% of underground hydrothermal reservoirs are not hot enough to drive a turbine.

Our estimate for accessible geothermal energy is ~ 2% of the global mean heat flux; or F_{GEO} ~ 0.0014 W m^{-2}. This could conceivably increase by drilling deeply into dry rock. In the most optimistic case one could recover only < 0.1% of the solar flux at the surface (averaged over large areas). Apart from the technological challenge of deep drilling through the crust to underlying magma, the impact on plate tectonics of large-scale mining of the Earth's interior heat by humans is unknown.

4.4.1.7 Can Renewable Energy Fluxes be Captured?

Advocacy of renewable energy as the goal of human energy supply

(04) Chapter 4 7/31/96 1:23 AM Page 239

can be found in many articulate analyses (Lovins, 1977; Williams, 1979; Goldemberg et al., 1988; Ogden and Williams, 1989; Brower, 1990). Why have these technologies had so little impact? The latest global commercial energy consumption statistics show fossil fuels remain, by far, the dominant source (90.4%), followed by nuclear (6.8%), hydro (2.4%) and geothermal plus wind (0.4%) (Hammond, 1994, Table 21.1). Biomass burning, primarily by developing countries off the cash economy, does not even show up as commercial energy.

Undoubtedly, part of the resistance comes from the entrenched infrastructure and investment in fossil energy systems. But a more fundamental problem, and one following from the underlying physics, is low power density. Naturally available energy flows are thin gruel. They impose severe constraints on land use, materials, energy transmission and energy storage, all of which have to be addressed before they can become practical. Our analyses indicate energy fluxes potentially exploitable by renewable energy technology in the range of 0.01 to 10 W m^{-2}. These are orders of magnitude below power extraction rates per unit area from coal and oil fields, and fossil and nuclear thermal power plant heat fluxes. Although significant variations can be found in estimates of harvestable renewable energy fluxes, our numbers are generally comparable in order-of-magnitude to those derived by others.

Figure 4.4, for example, shows power flux versus area regimes for various renewable and fossil energy sources estimated by Smil (1991). As he observes, "Graphic comparison of power densities of various energy supply modes with those of final uses reveals a wealth of spatial implications... The most important conclusions concern the spatial consequences of the eventual transition to a solar civilization. *These contrasts bypass the dubious cost comparisons to demonstrate the fundamental physical difficulties of such a transformation* (our italics)." Large power densities of present-day fossil and nuclear steam turbine powerplants derive from the concentrated nature of their energy sources and the consequently large heat fluxes attainable in the boilers. For example, a typical heat absorption rate per unit area of the walls of a pulverized coal furnace is 160,000 W m^{-2} (~ 50,000 BTU hr^{-1} ft^{-2}, from Kessler, 1978, Fig. 31, p. 9.21). Present-day thermal engineering can handle hundreds of kW m^{-2} to as much as 10 MW m^{-2} in blast furnaces, Such energy fluxes are available primarily

(04) Chapter 4 7/31/96 1:23 AM Page 240

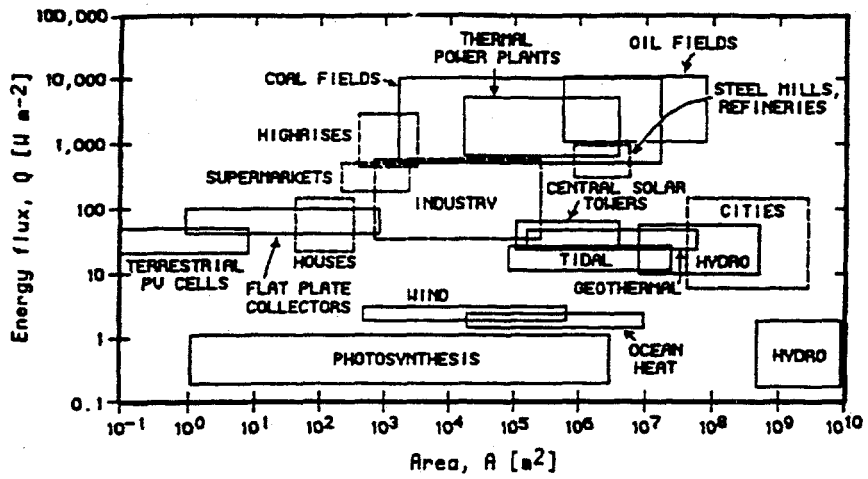


Figure 4.4 Flux-area domains of energy sources and sinks. This comparison highlights differences — often several orders of magnitude— and overlaps between the harnessing of renewable energy fluxes and extraction of fossil fuels on the one hand (solid rectangles), and energy end use fluxes on the other (dashed rectangles). Since the present global energy infrastructure is based on high energy density fossil fuels, a transition to low power density renewable energy is likely to require socio-economic, as well as technological, transformations. (Redrawn from Smil, 1991, p. 243. With permission.)

from fossil, nuclear and (presumably) fusion power sources. These inherently require less materials and land than terrestrial renewable energy.

The flux of solar energy on a given spot can be increased with optical concentrators, where the area ratio of the concentrator aperture to the solar image spot size is the concentration ratio, CR. However, the collection area for a given radiant power to an energy conversion device is as large as without concentration, and the materials and land use implications are similar. Solar concentrators are worth pursuing — as in "power tower" or heliostat configurations — because higher temperatures in a working fluid near the focal point can drive more thermodynamically efficient heat engines. Two-axis mirror tracking is normally required to keep axisymmetric concentrators pointed at a fixed focal spot on a heat exchanger.

The thermodynamic limit on the concentration ratio of axisymmetric solar concentrators is (Welford and Winston, 1981, p. 3.11) $CR_{max} = 1/(\sin\theta)^2$, where θ is the semi-angle subtended by the sun ~

(04) Chapter 4 7/31/96 1:23 AM Page 241

(R_{sun}/a) ~ 4.64 x 10^{-3} radians ~ 0.27°. Hence, $CR_{max} \sim \theta^{-2} \sim 46{,}000$. In principle the solar constant flux, 1380 W m^{-2}, could be focused on a spot with a flux of $(1380)/(4.64 \times 10^{-3})^2 \sim 64$ MW m^{-2} — the radiative intensity at the solar photosphere. But because of the low mean power flux available, and the intermittent nature of the source, even if fluxes of order ~ 1–10 MW m^{-2} could be produced by the Sun at the focal spot of a central power station, land and material requirements would be orders of magnitude larger than fossil or nuclear plants of the same capacity.

The land use problem is fundamental. A renewable energy yield of 1 W m^{-2} requires 10% of the earth's land area (1.5 x 10^{13} m^2) to produce 15 TW (15 x 10^{12} W), the level of nonfossil energy we estimated earlier might be required in the next 50 years to mitigate global warming by greenhouse gases.

Boyle (1994) has reported an assessment by Greenpeace International and the Stockholm Environmental Institute–Boston (SEI–B) of a Fossil Free Energy Scenario (FFES) in which energy-efficiency is combined with the phaseout of fossil fuel by the year 2100 and the phasein of renewable technologies. Their scenario assumes that biomass technologies, windpower and a limited expansion of hydropower provide the bulk of the renewables in the short run, while over the long term a solar energy system (solar PV, solar thermal power and solar water and space heat) linked to hydrogen predominates, providing 80% of global energy by 2100. Boyle's (1994) FFES scenario requires 8% of global arable/pasture/forest and woodland (6% of total global land area). These land use results are comparable to ours. Boyle argues that the use of non-productive land (deserts and degraded areas) plus roof tops can alleviate potential problems. Note that the technology employed in the FFES scenario assumes renewable energy technologies of the 21st Century are those already "on the shelf" in the latter part of the 20th Century.

As for land requirements, rooftops are convenient sites for solar collectors, particular for solar space heating of well-insulated structures, but are inadequate to power a solar civilization. Moreover, the use of remote areas for renewable energy poses special problems of long-distance energy transport. Deserts are well-endowed with sunlight, but usually distant from the intended points of energy use. Hydrogen has been proposed as a potentially important energy carrier (Winter and Nitsch, 1988), but needs to be considered in a total

(04) Chapter 4 7/31/96 1:23 AM Page 242

energy analysis as many energy conversion steps are needed (water electrolysis, transmission of H_2 by pipeline, combustion or electric power generation in a fuel cell).

Given the low power density and intermittency of natural energy fluxes, transmission and storage are critical, and relatively unexplored, components of sustainable global renewable energy systems. We recommend that innovative technologies be explored which can qualitatively improve global energy infrastructure through new modes of transmission and storage.

4.4.1.8 Electric Power Transmission

Electric power cannot efficiently be transmitted today more distantly than several hundreds of kilometers owing to the finite resistance of ambient temperature conductors (aluminum and copper). This reach is insufficient for a global solar energy system to bridge day/night cycles or to bypass overcast sunless regions. A major problem for any global-scale renewable energy system is "load matching" — connecting (or wheeling) supply and demand over long distances between consumers and producers.

Ohm's law ($V = IR$) indicates that power losses from the resistance, R, of an electrical transmission line are minimized when electricity is transmitted at the highest possible voltage, V, for a given transmitted power ($P = IV$), $P_{loss} = I^2R = P^2R/V^2$. The volumetric resistivity of transmission lines is a material property, $\rho = RA/l$ (~ 8.4 $\mu\Omega$-cm for aluminum-clad steel), where l is the length of the conductor and A its cross sectional area; so power loss per unit length is $P_{loss}/l = P^2\rho/(AV^2)$. Per unit length of wire, the material cost increases nearly linearly with A, whereas the cost of the power loss decreases inversely with A. The sum of these (the total cost) versus A thus has a minimum which determines the most cost-effective wire cross section. Maximum line voltage is normally limited by the breakdown of insulation or air separating the conductors (~ 1000 kV in current practice).

Nikola Tesla reasoned in the early years of this century that the highest possible voltages should be used to minimize I^2R losses in electric power transmission. He invented a series of alternating current (AC) generators and transformers to accomplish this. Undaunted by Ohm's law, Telsa's former employer (and later commercial opponent), Thomas Edison, backed low voltage direct current (DC). Edison and Tesla fought a bitter battle for public acceptance called the

(04) Chapter 4 7/31/96 1:23 AM Page 243

"war of the currents" (Hackmann, 1993). Tesla had to win (and did), because physics favors high voltage, and because the technology of the time for controlling voltage was AC. Three-phase AC power lines became standard in much of the world for transmitting electricity the few hundred kilometers or less from fossil and nuclear central power stations to consumers. But existing grids are inadequate for global renewable energy systems where electrical sources and sinks would necessarily be much more widely dispersed.

Technological breakthroughs could make long-distance electricity transmission a reality in the 21st century. Remarkably, the resistivity of certain substances vanishes entirely ($\rho \rightarrow 0$) at temperatures close to absolute zero. Typically, these "Type I" superconductors require cooling by liquid helium (He) which, at atmospheric pressure, boils at 4.2 K — an expensive proposition because of the large refrigeration power and the scarcity of He gas. In a pivotal discovery, Bednorz and Muller (1986) found ceramic oxides exhibiting a new form of superconductivity ("Type II") at temperatures as high as 30 K. Chu and Wu, and others, later found materials exhibiting Type II superconductivity above the 77 K boiling point of liquid nitrogen (Cambel and Koomanoff, 1989). This development opens up the possibility of cost-effective superconducting power transmission because N_2 is an abundant atmospheric constituent, and the power needed to liquefy it is orders of magnitude less than for He.

Superconducting transmission lines would most likely be buried cryogenically cooled cables (Bockris, 1980). Preliminary studies emphasize AC lines compatible with electric utility practice in the U.S. (Giese et al., 1988). However, a potentially fatal power loss of AC superconductors is the "skin effect" caused by vortex motion into and out of the cable (Dale et al., 1990).

A solution to the AC loss problem may simply be to use DC in the superconducting link, with state-of-the art semiconductor AC alternators and inserters at the ends (Engström, 1978). Edison did not have these. But there is no theoretical reason why they should not work. A superconducting power line 100 km long could be > 99.8% efficient if the only power loss were that needed to replenish the liquid N_2 (Giese et al., 1988). Maintaining this power loss per unit length over a 10,000 km transmission link gives a point-to-point transmission efficiency of $\eta_T = (.998)^{100} \times 100\% \sim 81\%$, quite acceptable for wheeling renewable energy planetwide.

(04) Chapter 4 7/31/96 1:24 AM Page 244

The prospect of global superconducting transmission lines was raised by Buckminster Fuller even before the discovery of Type II superconductivity (Fuller, 1981). His interest was motivated partly by geometric thinking and partly by the realization that population growth is inversely correlated with per capita GDP and energy consumption. The Fullerian approach to the latter problem was to foster economic development worldwide by a planetwide electricity grid (Figure 4.5).

Fuller's thinking is best expressed in his own inimitable language (Fuller, 1991): "...We must be able to continually integrate the progressive night-into-day and day-into-night hemispheres of our revolving planet. With all the world's electric energy needs being supplied by a twenty-four-hour around, omni-integrated network, all of yesterday's, one-half-the-time-unemployed, standby generators will be usable all the time, thus swiftly doubling the operating capacity of the world's electrical energy grid." Regarding his icosahedral map projection, he explains, "... A half century ago I discovered with my non-visibly distorted, one-world-island-in-one-world-ocean, 90° longitude-meridian-backbone, north-south-oriented, sky-ocean world map that a world energy network grid would be possible if we could devel-

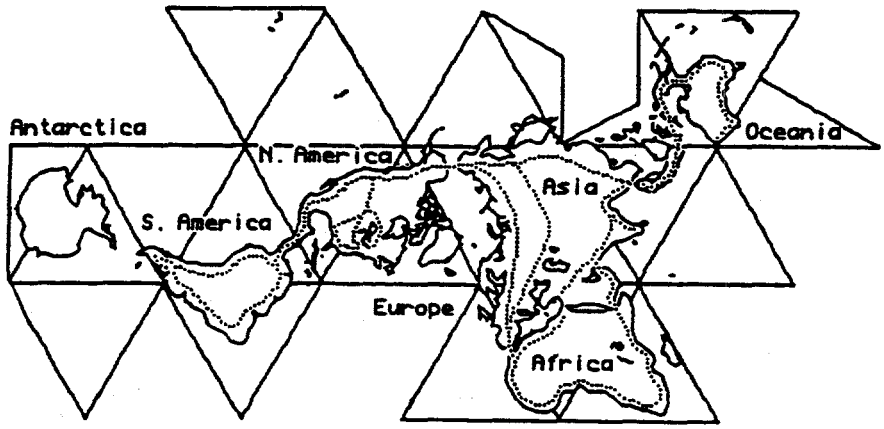


Figure 4.5 A proposed global superconducting electric grid (dotted lines). An icosahedron is a polyhedron of twenty equilateral triangles approximating a sphere. The world's land masses and oceans are shown in an icosahedral, as opposed to a Mercator, map projection to more realistically represent the size of continents, and to emphasize their near contact points where superconducting transmission lines could be run. Buckminster Fuller's ideal was a common electrical currency for the global economy of 1¢/kWe-hr. (Redrawn from Fuller, 1981. With permission.)

(04) Chapter 4 7/31/96 1:24 AM Page 245

op the delivery reach..."

Fuller's grid is an example of how important ideas can come from unanticipated directions when a creative mind is at work. Combined with plausible developments in superconductivity, his vision provides an electricity distribution approach that could enable terrestrial renewables to be effective global energy sources by the 21st century.

4.4.1.9 Energy Storage

In addition to new power transmission systems, more effective energy storage will be needed by a global nonfossil fuel civilization. Physics today defines matter in terms of elementary quarks and leptons and their antiparticles. Energy is stored in matter by the four fundamental forces through which these particles can interact — the strong and weak nuclear, electromagnetic and gravitational forces.

It follows from special relativity that the energy that can be stored per unit mass is limited to $\Delta E/m \leq c^2$, where m is the rest mass and c ~ 3.0×10^8 m s^{-1} the speed of light. The theoretical maximum (for matter-antimatter annihilation) is much larger than what has been obtained by any present-day storage system. Table 4.5 lists values of the specific energy, $\Delta E/m$, and the rest mass energy fraction, $\Delta E/(mc^2)$, for various materials and systems. The variation over so many orders of magnitude arises from the different strengths of the fundamental forces and the distances over which they act. In general, the higher specific energy systems require less material and land.

The most technologically ready large-scale electrical energy storage system is pumped-storage. In operation, a reversible hydroelectric generator-pump pumps water from a river or lake into a reservoir at higher elevation during periods of low power demand from utilities. During demand peaks, the system is reversed to recover the stored energy. The round-trip efficiency of pumped storage is ~ 65% (Culp, 1978, p. 428). However, the specific energy is a very low ~ 1.0×10^3 J kg^{-1} for a 100 meter head, leading to substantial land requirements. Pumped-storage also requires a specific type of topography which limits its applicability. In any event, the diversion of rivers and lakes for this purpose can lead to the same kinds of ecosystem disruption as hydroelectric reservoirs and dams discussed previously.

Batteries have more than 100 times the $\Delta E/m$ of pumped storage, ~ 1.2×10^5 J kg^{-1} for the conventional lead-acid automotive type, but are low in specific energy compared with fossil fuels. (This is the

(04) Chapter 4 7/31/96 1:24 AM Page 246

Table 4.5 Energy stored in matter (adapted from Culp, 1979, Table 9.1).

Material or System	$\Delta E/m$ [J kg^{-1}]	$\Delta E/(mc^2)$ [-]
Antimatter (M/A annihilation reaction)	9.0×10^{16}	1.0
Deuterium (D/D fusion reaction)	3.5×10^{14}	3.9×10^{-3}
Uranium 235 (fission reaction)	7.0×10^{13}	7.8×10^{-4}
Reactor fuel (2.5% enriched UO_2)	1.5×10^{12}	1.7×10^{-5}
80% plutonium-238 (radioactive decay)	1.8×10^{9}	2.0×10^{-8}
Hydrogen (lower heating value)	1.2×10^{8}	1.3×10^{-9}
Gasoline (lower heating value)	4.4×10^{7}	4.9×10^{-10}
Lead-acid battery (automotive)	1.2×10^{5}	1.3×10^{-12}
Flywheel (uniformly-stressed disk)	7.9×10^{4}	8.7×10^{-13}
Organic elastomer ("rubber band")	2.0×10^{4}	2.2×10^{-13}
Falling water (H = 100 m)	1.0×10^{3}	1.1×10^{-14}
Torsion spring (steel)	2.4×10^{2}	2.6×10^{-15}
Capacitor (communication electronics)	1.6×10^{1}	1.8×10^{-16}

main problem with battery-powered vehicles.) The U.S. Department of Energy (DoE) is supporting research on improved batteries for both automotive and stationary energy applications. Sodium-sulfur batteries are considered the most promising (Cohen, 1990, p. 265). A goal of the DoE is 80% efficiency over 2,500 charge-discharge cycles at a cost of \$90 per kW_e-hr capacity. This may be overly optimistic, but if it were attained the cost per charge-discharge cycle would be (\$90/2,500) x 0.8 ~ 2.8 ¢/kW_e-hr, not counting installation and maintenance. Materials and land use implications of a global electrical energy battery storage system have not been analyzed.

A global energy economy employing H_2 generated by electrolysis [$2H_2O \rightarrow 2H_2 + O_2$] as "currency" has been much discussed in recent years (Bockris, 1980; Winter and Nitsch, 1988; Ogden and Williams, 1989). Advantages of hydrogen as an energy carrier are storability at high specific energy and the absence of greenhouse gas emissions. Deriving H_2 from water-splitting is crucial to CO_2 emission reduction. Commercial H_2, as made today by steam reforming of natural gas [$2H_2O + CH_4 \rightarrow 4H_2 + CO_2$], is more likely to *increase* CO_2 emissions. Of course, the primary power source also needs to be non-fossil. Conventional water electrolysis with aqueous electrolytes employs an asbestos diaphragm and alkaline electrolytes to prevent corrosion of the metal components. H_2 production with alkaline electrolytes is well-established, but has been used thus far mainly for small to medium-sized plants (< 5 MW_e; Winter and Nitsch, 1988, p.

(04) Chapter 4 7/31/96 1:24 AM Page 247

179).

The specific energy of H_2 fuel based on the mass of hydrogen is $\Delta E/m \sim 1.2 \times 10^8$ J kg^{-1} — three times higher than gasoline (Table 4.5). This doesn't count the mass of the storage tank, and the containment of gaseous hydrogen remains a major challenge for applications. Present-day technologies include cryogenic liquid hydrogen (LH_2) tanks and adsorption by metal hydrides such as iron-titanium ($FeTiH_x$). Hydrogen to total (tank plus hydrogen) mass fractions for H_2-powered vehicles were estimated by Carpetis (1988, Table 9.2) as ~ 0.10 kg(H_2)/kg for liquid hydrogen (LH_2) and ~ 0.018 kg(H_2)/kg for a metal hydride combination tank. Including these tankage factors decreases the $\Delta E/m$ of hydrogen to 1.2×10^7 J kg^{-1} (LH_2) and 2.2×10^6 J kg^{-1} (metal hydride)— less than gasoline but far better than storage batteries.

The main problems with hydrogen storage at this point are high cost and possible limitations from toxic (asbestos) and rare (titanium) materials. Relatively low power densities of electrolyzers and issues of water availability in remote locales may also pose problems for a global hydrogen economy.

The optimum energy storage system in a land-limited world may be that with the highest specific energy. There would, for example, be advantages to an economy in which energy was stored in antimatter and recovered through matter-antimatter reactions (Table 4.5). The technological challenge is major, but so is the payoff. When antiprotons annihilate with protons, the products are, on average, five elementary particles — three charged and two neutral pions. The charged pions live 26 ns and contain 60% of the rest mass energy (Forward, 1987). Thus ~ 5.4×10^{16} J kg^{-1} in kinetic energy of charged pions could be available for conversion to (say) electricity — eight orders of magnitude more than for hydrogen!

Antimatter is virtually nonexistent in the natural environment for reasons of interest mainly to cosmologists. But antiprotons are produced routinely today in particle accelerators, albeit in small quantities. When relativistic protons strike a dense metal target, their kinetic energy, which is many times their rest mass energy, is converted into a spray of particles, some of which are antiparticles. A magnetic field focuses and a selector separates antiprotons from the resulting debris and directs the antiprotons into a storage ring.

Storing antimatter is more difficult than storing hydrogen — the

(04) Chapter 4 7/31/96 1:24 AM Page 248

containment must avoid mutual annihilation on contact with normal matter. We do not know how to do it yet, but given its high ΔE/m, the energy stored as antimatter would make an excellent space propulsion energy source.

Again, science fiction may lead reality. The starship *Enterprise*, of the *Star Trek* series, is powered by matter/antimatter (M/A) reactions. According to the latest technical manual (Sternbach and Okuda, 1991, p. 60), "The key element in the efficient use of M/A reactions is the dilithium crystal... Dilithium permits the antihydrogen to pass directly though its crystalline structure without actually touching it, owing to the field dynamo effect created in the added iron atoms..." Although Arthur C. Clarke has reminded us that a sufficiently advanced technology is indistinguishable from magic, we cannot count on dilithium crystals or other "unobtainium" to provide antimatter containment in the near future.

Fortunately, there are options consistent with known physics (Forward, 1987). It is already possible to slow antiprotons in a particle accelerator storage ring to almost zero velocity and capture them in a small electromagnetic ion trap no larger than a thermos bottle, although the technique is limited to low ion densities. One possibility is to add positrons in the ion traps to slowly build up "cluster ions" of antihydrogen — large agglomerations of neutral hydrogen atoms clustered around a single antiproton. Forward (1987) proposes that the net electric charge of the cluster ion would permit it to be held in an ion trap until the mass is increased to the point where it could be electrostatically levitated without touching the walls of a cryogenically-cooled trap. This antimatter could then be transported to applications.

There is no guarantee that technologies speculated on here will become cost-effective in the next century. But it is likely that 21st century energy supplies will differ from today's more than what might expected by linear extrapolation of current trends. Given the explosive technology changes of the past century, it is at least possible that the low power fluxes of terrestrial renewables could be harvested with the help of superconducting global power grids and antimatter energy storage.

4.4.1.10 Land Use, Population Growth and Energy

Any effective global energy system should minimize the impact on land use. Humankind has already appropriated ~ 10% of the land sur-

(04) Chapter 4 7/31/96 1:24 AM Page 249

face for agriculture (Table 4.4) — a fraction that will almost certainly increase by the middle of the next century. At the same time, there will be increasing pressure on natural ecosystems for renewable energy conversion. We (and others) estimate this will require another 10% of the Earth's land surface. Humankind has already appropriated ~ 40% of the world's biological productivity, if one includes NPP used for food, the wood used for timber, paper and firewood, the diversion of organic matter harvested but not eaten, and modification of the land for human habitats intentionally (asphalt, pastures and cropland) and unintentionally (desertification). This proportion of terrestrial NPP diversion is the largest by any one species (and its servant species) since life colonized the land (Diamond, 1987).

Tropical rainforests are being cut back for agriculture and habitats at rates that could destroy them by the middle of the 21st Century. Some 34 mostly poor countries with high fertility rates account for ~ 97.5% of tropical rainforests — Brazil, Zaire and Indonesia contain more than 50%. Meyers (1991, Table 1) estimates that only ~ 530 x 10^6 ha remain of ~ 1,360 x 10^6 ha of primary tropical rainforest. Thus [(1,360-530)/1,360] x 100% ~ 61% of the primary rainforests have already been destroyed. If the remaining 39% is deforested at 2% per year, then only $(.98)^{50}$ x 39% ~ 14% will survive fifty years hence. There is worldwide concern over the implications of this biomass and species destruction, but studies indicate that much tropical deforestation is the result of "slash and burn" shifting cultivation by the very poor in developing nations, who see no alternative to encroachment on the marginal environments of the tropical forests (Meyers, 1991).

As a society becomes more wealthy (as measured by its per capita income, GDP/N, or per capita energy consumption, E/N) the birth rate of its people tends to decline (Meadows et al., 1992, Fig. 2-7). All of the poorest nations experience birth rates between 20 and 50 per thousand people per year, while the rich nations generally have birth rates above 20 per thousand. There are some exceptions like China and Sri Lanka which have anomolously low birth rates for their level of income, and some Middle Eastern countries have anomolously high birth rates for theirs. But in general the inverse correlation between birth rate and GDP/N holds. The present-day population dynamics of tropical developing nations is characterized by a positive feedback loop in which poverty leads to overpopulation and overpopulation leads to further poverty, with concomitant impacts on natural ecosys-

(04) Chapter 4 7/31/96 1:24 AM Page 250

tems.

Protection of the global environment against adverse effects of deforestation, loss of biodiversity and climatic change is thus intimately linked to the economic level and energy use of developing countries. As Meadows et al. (1992, p. 40) observe, "Any positive feedback loop that grinds a system down, however, can be turned around to work the other way. More prosperity, widely distributed, can lead to slower population growth, which can lead to more prosperity." It was for precisely those reasons that Buckminster Fuller was motivated to find a more equitable global energy distribution system.

The development of sustainable nonfossil energy systems is a multifaceted problem that must be assessed in an integrated way. However desirable for other reasons, viable solutions must provide greenhouse-free energy supplies for developing nations which foster reasonable economic development along with population control. In the long run, that is the only way to prevent negative feedbacks from destroying the natural environment. In this section we analyzed advantages and limitations of renewable energy sources, and of transmission and storage systems, that could reduce greenhouse gas emissions. The low power fluxes of renewables can be most efficiently exploited by yet-to-be-invented energy distribution and storage systems. Such systems are going to be major challenges for 21st Century technology, and should be major research priorities now.

4.4.2 Nuclear Fission

Nuclear fission fuels are derived from heavy element radionuclides in the Earth's crust originally synthesized by exploding supernovae. The detritus of such an explosion, including nuclear fuel, was incorporated into the matter cloud from which our solar system formed. Of immediate relevance to global energy supply are the uranium and thorium resources residing in the Earth's crust which decay with half-lives ($t_{1/2}$) of order the solar system age (~ 5×10^9yr):

$^{238}U \rightarrow {}^{206}Pb + 8\,^4He, \quad t_{1/2} \sim 4.5 \times 10^9 yr$
$^{235}U \rightarrow {}^{207}Pb + 7\,^4He, \quad t_{1/2} \sim 0.7 \times 10^9 yr$
$^{232}Th \rightarrow {}^{208}Pb + 6\,^4He, \quad t_{1/2} \sim 14.0 \times 10^9 yr$

In the late 1930s, Hahn and Strassmann discovered that bombarding natural uranium with neutrons moving at energies of only a few electron volts (eV) would split the ^{235}U nucleus and release about 170

(04) Chapter 4 7/31/96 1:24 AM Page 251

million electron volts (MeV) — a ferociously exothermic reaction with output exceeding input by five orders of magnitude.

At the time of a supernova explosion, ^{235}U is produced more abundantly than ^{238}U, but the more rapid decay of the lighter (fissionable) isotope over the age of the solar system and the time prior to its formation has resulted in a present $^{235}U/^{238}U$ ratio in uranium ore of only ~ 0.72%. Economically recoverable fissionable uranium reserves available to present-day light water nuclear reactors (LWRs) is quite modest, less than oil or natural gas in energy equivalents at present energy prices (Table 4.1).

The fuel of LWRs is mostly ^{238}U enriched with ~ 3% readily fissionable ^{235}U — about 4 times the natural abundance. Bombs require much more enrichment (close to 100%). Stimulated by a perceived Nazi threat to develop the first atomic bomb, U.S. Manhattan Project scientists in Los Alamos during World War II made weapons from both highly-enriched ^{235}U and ^{239}Pu — a man-made plutonium isotope bred from uranium. The first atomic explosion on July 16, 1945 at the Trinity site in the New Mexico desert had an imploding plutonium core. The "Little Boy" bomb dropped on Hiroshima at the end of WW II incorporated ~ 50 kg of highly purified ^{235}U (Serber, 1992)[2]. The possible proliferation of nuclear weapons from nuclear power plants has been of concern ever since. Subsequent to the Gulf War of 1990, Iraq informed the International Atomic Energy Agency (IAEA) that it had on hand ~ 14 kg of 80% enriched and 11.3 kg of 93% enriched ^{235}U from its "peaceful" nuclear energy programs (Norman, 1991) — of the order to produce a nuclear weapon. Concern over nuclear proliferation has now shifted to North Korea. The so-called n-country problem in which progressively more nations acquire weapons-grade materials from their nuclear energy programs is clearly upon us, and getting worse. Some analysts argue that global nuclear energy should be developed only by politically "stable" societies, and that an electricity or electrolytic hydrogen be product marketed to others, but this is clearly a value-laden position.

[2]Three critical masses < 20 kg each were distributed between a ring fixed around the muzzle of a small cannon and a "bullet" inserted into the cannon breech ahead of the bags of cordite and fired at the appropriate time up the barrel to complete the supercritical assembly. Los Alamos considered Little Boy to be sufficiently conservative that it delivered the weapon without further testing. Smaller plutonium bombs are possible: A solid sphere of ~ 5 kg ^{239}Pu surrounded by a thick shield (tamper) explodes immediately on assembly (Serber, 1992).

(04) Chapter 4 7/31/96 1:24 AM Page 252

While environmentalists tend to favor renewable energy sources, there are good engineering arguments for the high power densities of nuclear reactors over the generally much lower energy fluxes available from solar, wind and biomass energy. On the anti-nuclear side is also a public concern about nuclear safety. A severe accident occurred in March 1979 at the Three Mile Island (TMI) power plant, an LWR of the type that now generates 85% of the world's nuclear-generated electricity. Meltdown of about half of the fuel occurred though the public was protected from hazardous radiation releases by the containment building. A much more serious accident occurred at the Chernobyl reactor complex in the Soviet Union in April 1986 resulting in some 30 prompt fatalities and a potentially much larger number of adverse health effects from the radiative contaminant released. There are also unresolved problems in disposing of radioactive wastes and spent fuel from power plants when they are decommissioned. Advocates of nuclear power argue that inherently safe nuclear power can be developed, and cite global warming by fossil fuel carbon dioxide as a motivation for doing so (Weinberg, 1992).

An approach that overcomes low ^{235}U abundance is *breeder reactors* — reactors that "breed" fissionable fuels from nonfissile uranium and/or thorium. A critical parameter for both nuclear fission and nuclear fuel breeding is η — the ratio of neutrons produced by a nuclear reaction to neutrons absorbed by the fuel (Table 4.6). In order to sustain a chain reaction, we need $\eta > 1$. In order to "breed" nuclear fuel we need $\eta > 2$; one neutron to continue the fission chain, and one neutron to be absorbed by the fertile nucleus to produce a new fuel atom. It follows from the values in Table 4.6 that for thermal neutrons ^{233}U is the best thermal breeder reactor fuel, and that it may be possible to build a thermal breeder reactor operating on the thorium-uranium-233 fuel cycle. The value of η for fast neutrons is significantly > 2 for all three fuel isotopes, which has motivated so-called fast breeder reactors.

France is singular in its commitment to breeder technology, having constructing the Rapsodie, Phenix and Super Phenix breeder reactors to explore its potential. These are important experiments. But a long time may be needed to breed enough plutonium and to build plutonium-burning reactors to replace fossil energy on a global scale — perhaps 50 years. It may already be too late to replace fossil energy with breeder reactors at the rates needed to mitigate global warm-

(04) Chapter 4 7/31/96 1:24 AM Page 253

Table 4.6 Neutrons produced per neutron absorbed by the fuel for common nuclear fuel isotopes, η (from Culp, 1979, Table 3.1. With permission.).

Fuel Isotope	Thermal Neutrons, (KE ~ 0.25 eV)	Fast Neutrons (KE > 1.5 MeV)
^{233}U	2.28	2.60
^{235}U	2.10	2.65
^{239}Pu	2.09	3.04

ing.

Alvin Weinberg, pioneer of "inherently safe" worldwide breeder reactors as a solution to the fossil fuel greenhouse, expresses his case profoundly to nuclear engineering colleagues — but sidesteps the thorny issues of nuclear proliferation and disposal of radioactive wastes (Weinberg, 1992):

"...There was no heavenly-ordained requirement that the age of fossil fuel be replaced by the age of fission, nor a cosmically-ordained anthropic cosmological principle that required η to exceed 1 so that a chain reaction be possible, or exceed 2 so that breeding is possible. Had fission not been developed we would, willy-nilly, be conserving energy at a much faster rate than now, we would be pushing fusion even harder — and as a last resort, we would turn to solar energy. A solar world in which primary energy is 3 times as expensive as it is today is hardly an impossible world, especially since in such a world, energy would be much more strongly conserved than now. We would not be relegated to Malthusian poverty, were the only reactor upon which man depended for energy located 150 x 10^6 kilometers away from the Earth.

"But Hahn and Strassmann's discovery of 50 years ago and God's providence in adjusting the nuclear constants so as to make a power breeder practical have given us another option. We nuclear engineers of the first nuclear era have had successes, yes, with our 500 commercial nuclear reactors, and our practical breeders. But the job is only half finished. The generation that follows us must resolve the profound technical and social questions that are convulsing nuclear energy. The challenge is clear, even the technical paths to meet the challenge are clear. All of us old-timers wish we will be here to see how these challenges are met; but even if we shall not be here, we wish the new generation well in fashioning an acceptable Second Nuclear Era!"

A reactor design developed by DoE that addresses nuclear safety

(04) Chapter 4 7/31/96 1:24 AM Page 254

and waste disposal but not breeding issues is the Integral Fast Reactor (IFR) being readied for test near Idaho Falls by Argonne National Laboratory *(Discover,* 1994). The IFR uses high-energy fast neutrons to trigger the chain reaction, in contrast to conventional light-water reactors that slow their neurons with graphite rods. Because fast neutrons can cause many more types of elements to undergo fission, the IFR is not limited to the uranium and plutonium that conventional reactors use as fuel. It can burn highly radioactive elements with half-lives of tens of thousands of years that are waste by-products of uranium and plutonium fission. In contrast to existing breeders, the IFR burns plutonium rather than producing it, thus precluding, in principle, the possibility that a cache of weapons-grade plutonium will fall into the hands of unstable regimes or terrorists. IFR project manager, Yoon Chang, claims that the reactor safely system is designed to be invulnerable to catastrophic loss-of-coolant accidents that crippled the Chernobyl and Three Mile Island reactors.

Even if the new generation of breeder reactors prove themselves inherently safe, a coordinated worldwide strategy for conversion of nuclear reactors would be needed to deal with the prospect of ^{235}U running out before plutonium was bred. The classical reactor strategy would have required that prior to the year 2000 breeders be installed to compliment the less efficient burner reactors. For this to have occurred, there would have to have been worldwide fast breeder programs already in place comparable to that of the French (Hafele et al., 1981, p. 55). An alternate tack is a converter-breeder strategy which can be fueled with ^{233}U. It remains to be seen whether the IFR or a comparable innovative reactive design can fit such a strategy, and provide safe nonfossil energy supplies on a global scale in the 21st Century and beyond.

REFERENCES

(04) Chapter 4 7/31/96 1:24 AM Page 255

Allen, C.W. (1976) *Astrophysical Quantities*. Athlone Press, London, 310 pp.

APS (1979) *Principle Conclusions of the American Physical Society Study Group on Photovoltaic Energy Conversion*, H. Ehrenreich, Chair, American Physical Society, New York, 190 pp.

Ausubel, J.H. (1994) Technical progress and climatic change. In Nakicenovic, N., W.D. Nordhaus, R. Richels and F.L. Toth (eds.), *Integrative Assessment of Mitigation, Impacts, and Adaptation to Climate Change*, PC-94-9, pp. 501-512, International Institute for Applied Systems Analysis, Laxenburg, Austria, 669 pp.

Bednorz, J.G., and K.A. Muller (1986) *Z. Phys.*, B64, 189.

Bockris, J.O'M. (1980) *Energy Options: Real Economics and the Solar Hydrogen System*. John Wiley & Sons, NY, pp. 212-233.

Boden, T.A., R.J. Sepanski and F.W. Stoss (1991) *Trends '91: A Compendium of Data on Global Change*. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 665 pp.

Bolin, B. (1994) Next step for climate-change analysis. *Nature*, 368, 94.

Boulton, J.R., and D.O. Hall (1991) The maximum efficiency of photosynthesis. *Photochem. Photobiol.* 53, pp. 545-548.

Boyle, S. (1994) Toward a fossil free future: The technical and economic feasibility of phasing out global fossil fuel use. In Naki´cenovi´c, N., W.D. Nordhaus, R. Richels and F.L. Toth (eds.), *Integrative Assessment of Mitigation, Impacts, and Adaptation to Climate Change*, CP-9S9, pp. 353-378, International Institute for Applied Systems Analysis, Laxenburg, Austria, 669 pp.

Bradley, R.A., E.C. Watts and E.R. Williams (1991) *Limiting Net Greenhouse Gas Emissions in the United States, Vol. 1: Energy Technologies*. Report DoE/PE-0101, U.S. Department of Energy, Office of Environmental Analysis, Washington, DC.

Brower, M.C. (1990) *Cool Energy: The Renewable Solution to Global Warming*. Union of Concerned Scientists, Cambridge, MA, 89 pp.

Budyko, M.I. (1974) *Climate and Life*. Academic Press, New York, 508 pp.

Cambel, A.B., and F.A. Koomanoff (1989) High-temperature superconductors and CO_2 emissions. *Energy*, 14, pp. 309-322.

Carpetis, C. (1988) Storage, transport and distribution of hydrogen. In Winter, C-J., and J. Nitsch (eds.) *Hydrogen as an Energy Carrier: Technologies, Systems, Economy*. pp. 249-289, Springer-Verlag, New York, 377 pp.

Cohen, B. (1990) *The Nuclear Energy Option:* An Alternative for the 90s. Plenum Press, New York, 338 pp.

Cleveland, C.J., R. Costanza, C.A.S. Hall and R. Kaufmann (1984) Energy and the U.S. economy: A biophysical perspective. *Science*, 225, pp. 890-897.

Culp, A.W., Jr. (1979) *Principles of Energy Conversion*. McCraw-Hill, New York, 199 pp.

Dale, S.J., S.M. Wolf and T.R. Schneider (1990) *Energy Applications of High Temperature Superconductors, Vol. 1*, Electric Power Research Institute, Palo Alto, CA, 89 pp.

Diamond, J.M. (1987) Human use of world resources. *Nature*, 328, pp. 479-480.

Dickinson, R.E. (ed.) (1987) *The Geophysiology of Amazonia: Vegetation and Climate Interactions*. John Wiley & Sons, New York, 526 pp.

Discover (1994) Breakthroughs: fast neutrons in Idaho. *Discover*, April 1994, p. 23.

(04) Chapter 4 7/31/96 1:24 AM Page 256

Dugger, G.L., D. Richards, E.J. Francis and W.H. Avery (1983) Ocean thermal energy conversion: Historical highlights, status and forecast. *J. Energy,* 7 (4), pp. 293-303.

Early, J.T. (1989) Space-based solar shield to offset greenhouse effect. *J. Brit. Interplan. Soc..,* 42, pp. 567-569.

Edmonds, J. and J.M. Reilly (1985) *Global Energy: Assessing the Future.* Oxford University Press, New York.

Edmonds, J. and J.M. Reilly (1986) *The IEA/ORAU Long-Term Global Energy CO_2 Model: Personal Computer Version A84PC,* Institute for Energy Analysis, Oak Ridge Associated Universities, Oak Ridge, Tennessee (Available from National Technical Information Service, Springfield, Virginia).

Engström, P.C., et al. (1978) Direct-current power transmission. In Fink, D.G., and H.W. Beaty (eds.), *Standard Handbook for Electrical Engineers,* Chapter 15, McCraw-Hill, New York, 2,448 pp.

Faires, V., and C.M. Simmang (1978) *Thermodynamics,* Sixth edition. Macmillan Publishing Company, New York, 646 pp.

Forward, R.L. (1987) *Advanced Space Propulsion Study: Antiproton and Beamed Power Propulsion.* AFAL TR-87-070, Air Force Astronautics Laboratory, Air Force Systems Command, Edwards Air Force Base, CA, 29 pp. plus seven appendices.

Fuller, R.B. (1981) *Critical Path.* St. Martins Press, New York, 471 pp.

Giese, R.F., R.A. Thomas and E.B. Forsyth (1988) AC transmission. In Wolsky, A.M. et al. (eds.), *Advances in Applied Superconductivity: A Preliminary Evaluation of Goals and Impacts,* ANL/CNSV-64, pp. 69-100, Argonne National Laboratory, Argonne, IL; available from National Technical Information Service, U.S. Department of Commerce, Springfield, VA, 192 pp.

Gilder, G. (1993) When bandwidth is free. Interview in *Wired,* September/October 1993, pp. 38-41.

Glaser, P.E. (1968) Power from the sun: Its future. *Science,* Vol. 162, pp. 857-861.

Glaser, P.E., F.P. Davidson and K.I. Csigi (1993) *Solar Power Satellites: The Emerging Energy Option.* Ellis Horwood, New York, 300 pp.

Goldemberg, J., T. B. Johansson, A.K.A. Reddy and R.A.H. Williams (1988) *Energy for a Sustainable World,* Wiley Eastern, New Delhi, India, 517 pp.

Gore, A. (1992) *Earth in the Balance: Ecology and the Human Spirit.* Houghton Mifflin Co., New York, 408 pp.

Grubb, M.J., and N.I. Meyer (1993) Wind energy: resources, systems, and regional strategies. In Johansson, T.B., H. Kelly, A.K.N. Reddy and R.H. Williams (eds.), *Renewable Energy: Sources for Fuels and Electricity,* pp. 157-212, Island Press, Washington, DC, 1160 pp.

Hackmann, W. (1993) Tesla's sparks of imagination. *Nature,* 363, page 592.

Häfele, W., et al. (1981) *Energy in a Finite World: Paths to a Sustainable Future.* Vol. 1, Ballinger Publishing Company, Cambridge, Massachusetts, 225 pp.

Hall, D.O., H.E. Mynick and R.H. Williams (1991) Cooling the greenhouse with bioenergy. *Nature,* 353, pp. 11-12.

Hammond, A.L. (ed.) (1990) *World Resources 1990-91.* Oxford University Press, New York, 383 pp.

Hammond, A.L. (ed.) (1994) *World Resources 1993-94.* Oxford University Press,

(04) Chapter 4 7/31/96 1:24 AM Page 257

New York, 403 pp.

Harvey, L.D. (1990) Managing atmospheric CO_2: Policy implications. *Energy*, 15, pp. 91-104.

Hoffert, M.I. (1990) Climate sensitivity, climate feedbacks and policy implications. In Mintzer, I.M. (ed.), *Confronting Climate Change: Risks, Implications and Responses*, pp. 33-54, Cambridge University Press, New York, 382 pp.

Houghton, J.T., G. Jenkins and J.J. Ephraums (eds.) (1990) *Climate Change: The IPCC Scientific Assessment.* Cambridge University Press, New York, 365 pp.

Jain, A.K., H.S. Kheshgi, M.I. Hoffert and D.J. Wuebbles (1994) Distribution of radiocarbon as a test of global carbon cycle models. *Global Biogeochem. Cycles*, 9(1), 153-166.

Kaku, M. (1994) *Hyperspace.* Oxford University Press, New York, 359 pp.

Kessler, G.W. (1978) Steam boilers. In Baumeister, T., E.A. Avalone and T. Baumeister III (eds.), *Mark's Standard Handbook for Mechanical Engineers, Eight Edition*, pp. 9.7-9.35, McGraw-Hill, New York, 1,864 pp.

Krause, F., W. Bach and J. Koomey (1989) *Energy Policy in the Greenhouse*, Vol. 1, International Project for Sustainable Energy Paths (IPSEP), El Cerrito, California.

Leggett, J., W.T. Pepper and R.J. Stewart (1992) Emission scenarios for the IPCC. In Houghton, J.T., B.A. Callander and S.K. Varney (eds.) *Climatic Change 1992: The Supplementary Report to the IPCC Scientific Assessment*, pp. 69-95, Cambridge University Press, New York, 200 pp.

Lovins, A.B. (1977) *Soft Energy Paths: Toward a Durable Peace.* Harper Colophon Books, New York, 239 pp.

MacCracken, M.C.(ed.) (1990) *Energy and Climate Change: Report of the DoE Multi-Laboratory Climate Change Committee.* Lewis Publishers, Chelsea, Michigan, 161 pp.

Maier-Reimer, E., and K. Hasselmann (1987) Transport and storage of CO_2 in the ocean — An inorganic ocean-circulation carbon cycle model. *Clim. Dyn.*, 2, pp. 63-90.

Marland, G. (1991) A commentary on: "Tropical forests and the greenhouse effect: A management response." *Climatic Change* 19, pp. 227-232.

Matthews, R.B., E.P. Coomes and E.U. Khan (1994) Hierarchical analysis of options for lunar-surface power. *Journal of Propulsion and Power*, 10, pp. 425-440.

Manne, A.S., and R.G. Richels (1992) *Buying Greenhouse Insurance: The Economic Costs of Carbon Dioxide Emission Limits.* MIT Press, Cambridge, Massachusetts, 182 pp.

Meadows, D.H., D.L. Meadows and J. Randers (1992) *Beyond the Limits*, Chelsea Green Publishing Co., Post Mills, VT, 300 pp.

Meyers, N. (1991) Tropical forests: Present status and future outlook. *Climatic Change* 19, pp. 3-32.

Meyers, N. and T.J. Goreau (1991) Tropical forests and the greenhouse effect: A management response. *Climatic Change* 19, pp. 215-225.

Naki´cenovi´c, N., W.D. Nordhaus, R. Richels and F.L. Toth (eds.) (1994), *Integrative Assessment of Mitigation, Impacts, and Adaptation to Climate Change*, CP-94-9, pp. 501-512, International Institute for Applied Systems Analysis, Laxenburg, Austria, 669 pp.

(04) Chapter 4 7/31/96 1:24 AM Page 258

NAS (1992) *Policy Implications of Greenhouse Warming: Mitigation, Adaptation, and the Science Base.* Panel on Policy Implications of Greenhouse Warming, National Academy of Sciences, National Academy Press, Washington, DC, 918 pp.

Nordhaus, W.D. (1991) The cost of slowing climate change: a survey. *Energy J.*, 12 (1), pp. 37-65.

Nordhaus, W.D. (1992) An optimal transition path for controlling greenhouse gases. *Science*, 258, pp. 1315-1319.

Norman, C. (1991) Iraq's bomb program: A smoking gun emerges. *Science*, 254, pp. 644-645.

Ogden, J.M., and R.H. Williams (1989) *Solar Hydrogen: Moving Beyond Fossil Fuels.* World Resources Institute, Washington, DC, 123 pp.

Oort, A.H. (1983) *Global Atmospheric Circulation Statistics.* NOAA Professional Paper 14, p. 157, F40-F43, National Oceanic and Atmospheric Administration, Geophysical Fluid Dynamics Laboratory, Princeton New Jersey, 180 pp. plus microfiche data tables (F01-F47).

Ovshinsky, S.R., and A. Madan (1983) Amorphous semiconductors equivalent to crystalline semiconductors. United States Patent 4,409,605, Oct.11,1983.

Peixoto, J.P., and A.H. Oort (1992) *Physics of Climate.* American Institute of Physics, New York, 520 pp.

Reddy, A.K., and J. Goldemberg (1990) Energy for the developing world. *Scientific American*, 263 (Sept. 1991), pp. 111-118.

Revelle, R. (1985) Soil dynamics and sustainable carrying capacity of the Earth. In Malone, T.F., and J.G. Roederer (eds.), *Global Change*, Cambridge University Press, New York, pp. 465-473.

Richards, J.F. (1990) Land transformation. In Turner III, B.L. et al. (eds.), *The Earth As Transformed by Human Action: Global and Regional Changes in the Biosphere over the Past 300 Years*, pp. 163-201, Cambridge University Press, New York, 713 pp.

Riding, A. (1994) Back to the present in a long-lost novel by Verne. *The New York Times*, September 27, 1994, p. C15.

Savage, M. (1994) *The Millennial Project.* Little, Brown and Co., New York, 508 pp.

Scheraga, J.D., and N.A. Leary (1992) Improving the efficiency of policies to reduce CO_2 emissions. *Energy Policy*, May 1992, pp. 394-403.

Schlesinger, W.H. (1991) *Biogeochemistry: An Analysis of Global Change.* Academic Press, New York, 443 pp.

Schockley, W., and H.J. Queisser (1961) Detailed balance limit of efficiency of p-n junction solar cells. *J. Appl. Phys.*, 32, pp. 510-519.

Serber, R. (1992) *The Los Alamos Primer: The First Lectures on How to Build An Atomic Bomb.* University of California Press, Berkeley, CA, 98 pp.

SERI (1990) *The Potential of Renewable Energy: An Interlaboratory White Paper.* SERI/TP-260-3674, Solar Energy Research Institute, Golden, CO.

Shepard, M., J.B. Chaddock, F.H. Cocks and C.M. Harmon (1977) *Introduction to Energy Technology.* Ann Arbor Science, Ann Arbor, MI, 300 pp.

Shine, K.P., R.G. Derwent, D.J. Weubbles and J-J. Morcrette (1990), Radiative forcing of climate. In Houghton, J.T., G.J. Jenkins and J.J. Ephraums (eds.) *Climate*

(04) Chapter 4 7/31/96 1:24 AM Page 259

change: The IPCC Scientific Assessment, pp. 41-68, Cambridge University Press, New York, 364 pp.

Smil, V. (1991) *General Energetics: Energy in the Biosphere and Civilization.* John Wiley & Sons, New York, 370 pp.

Sternbach, R., and M. Okuda (1991) *Star Trek: The Next Generation Technical Manual.* Pocket Books, New York, 184 pp.

Sundquist, E.T. (1993) The global carbon dioxide budget. *Science,* 259, pp. 934-940.

Sverdrup, H.U., M.W. Johnson and R.H. Fleming (1970) *The Oceans, Their Physics, Chemistry and General Biology.* Prentice-Hall, Englewood Cliffs, New Jersey, 1087 pp.

Terborgh, J. (1992) *Diversity and the Tropical Rain Forest.* W.H. Freeman & Co., New York, 242 pp.

Thorndike, E.H. (1976) *Energy and Environment: A Primer for Scientists and Engineers,* Addison-Wesley, Reading, MA, 286 pp.

Tissot, B.P., and D.H. Welte (1978) *Petroleum Formation and Occurrence.* Springer-Verlag, New York, 538 pp.

Tolmazin, D. (1985) *Elements of Dynamic Oceanography.* Allen & Unwin, Boston, MA, 181 pp.

Turcotte, D.L. and G. Schubert (1982) *Geodynamics: Application of Continuum Physics to Geological Problems.* John Wiley & Sons, New York, 450 pp.

Vos, A.D. (1980) Detailed balance limit of the efficiency of tandem solar cells. *J. Phys. D: Appl. Phys.,* 13, pp. 839-846.

Waldrop, M.M. (1992) *Complexity: The Emerging Science of Order and Chaos,* Touchtone, New York, 380 pp.

Walker, J.C.G., and J.F. Kasting (1992) Effects of fuel and forest conservation on future levels of atmospheric carbon dioxide. *Paleogeogr., Paleoclimatol., Paleoecol.* (Global and Planetary Change Section), 97, pp. 151-189.

Wallace, J.M., and P.V. Hobbs (1977) *Atmospheric Science: An Introductory Survey.* Academic Press, New York, 467 pp.

Weinberg, A.M. (1992) The first and second fifty years of nuclear fission. In Kuliasha, M.A., A. Zucker and K.J. Ballew (eds.), *Technologies for a Greenhouse-Constrained Society,* pp. 227-237, Lewis Publishers, Boca Raton, FL, 835 pp.

Welford, W.T., and R. Winston (1981) Principles of optics applied to solar energy concentrators. In Kreider, J.F., and F. Kreith (eds.), *Solar Energy Handbook,* Chapter 3, McGraw-Hill, New York, 1,120 pp.

Wigley, T.M.L. (1995) Global-mean temperature and sea level consequences of greenhouse gas stabilization. *Geophys. Res. Lett,* 22, pp. 45-48.

Williams, R.H. (ed.) (1980) *Toward a Solar Civilization,* MIT Press, Cambridge, MA, 250 pp.

Wilson, E.O. (1992) *The Diversity of Life.* Harvard University Press, Cambridge, Massachusetts, 424 pp.

Winter, C.J., and J. Nitsch (eds.) (1988) *Hydrogen as an Energy Carrier: Technologies, Systems, Economy.* Springer-Verlag, New York, 377 pp.

Yergin, D. (1991) *The Prize: The Epic Quest for Oil, Money and Power,* Simon &

Schuster, New York, 877 pp.