A Reduced-Form Statistical Climate Model Suitable for Coupling with Economic Emissions Projections

by

Gregory S. Rabin

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Masters of Engineering in Electrical Engineering and Computer Science

at the

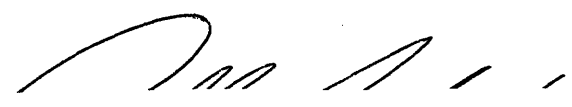MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2007

© Gregory S. Rabin, MMVII. All rights reserved.

The author hereby grants to MIT permission to reproduce and distribute publicly
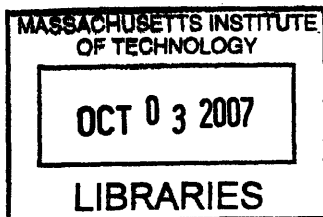paper and electronic copies of this thesis document in whole or in part.

Author . . . .
Department of Electrical Engineering and Computer Science
May 11, 2007

Certified by . . . .
Henry D. Jacoby
Co-Director, Joint Program on the Science and Policy of Global Change
Thesis Supervisor

Accepted by . . . .
Arthur C. Smith
Chairman, Department Committee on Graduate Theses

A Reduced-Form Statistical Climate Model Suitable for Coupling with Economic Emissions Projections

by

Gregory S. Rabin

Submitted to the Department of Electrical Engineering and Computer Science
on May 11, 2007, in partial fulfillment of the
requirements for the degree of
Masters of Engineering in Electrical Engineering and Computer Science

Abstract

In this work, we use models based on past data and scientific analysis to determine possible future states of the environment. We attempt to improve the equations for temperature and greenhouse gas concentration used in conjunction with the MIT Emissions Prediction and Policy Analysis (EPPA) model or for independent climate analysis based on results from the more complex MIT Integrated Global Systems Model (IGSM). The functions we generate should allow a software system to approximate the environmental variables from the policy inputs in a matter of seconds. At the same time, the estimates should be close enough to the exact values given by the IGSM to be considered meaningful.

Thesis Supervisor: Henry D. Jacoby
Title: Co-Director, Joint Program on the Science and Policy of Global Change

Acknowledgments

I would like to thank my supervisors Chris Forest, Henry Jacoby, and Ian Sue Wing, for all their help and encouragement, and also my parents and grandparents without whose help and support I would not be writing this thesis today.

Chapter 1

Introduction

During the last century, the amount of carbon dioxide and other greenhouse gases (CH_4, N_2O, etc.) industries, governments, and individuals emit into the atmosphere has been increasing greatly, threatening global warming and sea level rise.[1] If greenhouse gas emissions are not somehow controlled in the near future, the magnitude of these problems will continue to increase, leading to many undesirable climate changes.[2] Thus, it is imperative that we somehow limit emissions in a way that would stabilize atmospheric greenhouse gas concentrations within the next one or two centuries.[3] To do this, we must understand exactly how the variables that we (as humans) have control over, such as carbon emissions, influence the variables we wish to impact, such as temperature and atmospheric carbon concentrations.[4]

The MIT Joint Program on the Science and Policy of Global Change currently has a model of climate change and its impact on the economy: the Integrated Global System Model (IGSM).[5] This model has two submodels: the Emissions Prediction and Policy Analysis (EPPA) Model[6] simulating economic activity and a model of the geophysical climate system, including the physics and biogeochemistry of the atmosphere, ocean, and land systems.[5] The model has been developed to analyze interactions between human activity related to climate change and the resulting effects on the climate system. At the simplest scale, the IGSM economics model (EPPA) takes as input a scenario for future economic activity and produces an emissions scenario that will force the climate system. The climate system model then calculates the resulting concentrations, radiative forcings, and resulting response (changes in temperature, precipitation, winds, vegetation, etc.) to the given scenario. The IGSM is a complex model with a running time (up to the year 2100) of approximately 30 hours. This is sufficiently long to prevent running thousands of simulations. The model is calibrated to produce values close to reality for environmental variables, such as carbon concentration and temperature, given the input of policy values and random factors. (figure 1-1) The basic carbon cycle is determined by the transfers of carbon between atmosphere, ocean, land, and their

4

respective biospheres. In equilibrium, the fluxes between reservoirs balance, leading to relatively stable concentrations of CO_2 in the atmosphere.[7] Recently human industrial activity has greatly increased the emission of CO_2 and other greenhouse gases (GHGs), leading to an imbalance of the carbon cycle. The IGSM attempts to model the impact of this increase in GHGs by taking as input the expected industrial behavior, climate sensitivity (how sensitive atmospheric temperature is to changes in GHG concentrations), ocean sensitivity (how much excess heat and carbon the ocean is able to absorb), and aerosol forcing (how strong the radiative forcing is for a given aerosol concentration). The model includes submodels of human economic activity (EPPA), the atmosphere, urban area activity, ocean, and land, which combine into a large, complex model of the whole climate system. The IGSM produces outputs including GDP growth, energy use, global temperature, sea-level rise, greenhouse gas concentrations, air pollution levels, etc. In this work, we attempt to produce a simplified model of the "Earth System" part of the IGSM (brown box in figure 1-1). This simplified model will take as input the climate sensitivity, ocean diffusivity, aerosol forcing, and carbon emissions. They will yield the CO_2 concentration and average temperature.

The EPPA model, which has a running time of approximately 15 minutes, is the economics component of the IGSM. If we had simple mathematical functions, which could get approximately the same results for climate variables as the IGSM in less time, we could couple those functions with EPPA, thus getting a new less complex model, which is useful for fast calculations and estimation. Past researchers have been unsuccessful at determining a relation for climate variables that involves only simple mathematical operations that can be instantaneously computed. Such a function would be useful in policy and uncertainty studies when the full IGSM is not available due either to time constraints or to lack of access to software.

In the past, we attempted to couple the EPPA Model with algebraic fits for environmental variables in the MIT Toy Model.[8] These fits were made by estimating linear dependence of temperature increase as a function of the carbon emissions, temperature change values, and forcings in the previous period. However, the behavior of IGSM environmental variables is non-linear. Thus, while easily computable, the Toy Model was far from exact, as error terms which are fairly large in one period grow exponentially over the several time periods in the model.

We feel that this research is beneficial because the IGSM requires a long running time and a special software system. It is not practical to apply it every time we wish to analyze a new policy or uncertainty scenario. The EPPA Model requires fewer computational resources and so has a more reasonable running time; however it lacks the complex environmental calculations given in the IGSM. If we could couple EPPA with robust approximations of the complex non-linear behavior of the IGSM's climate component, we would have a better policy and uncertainty analysis tool.

In this study, we attempt to build a single climate model to satisfy the requirement for computational efficiency. To reduce the complexity of the system, we separate the world into three regions
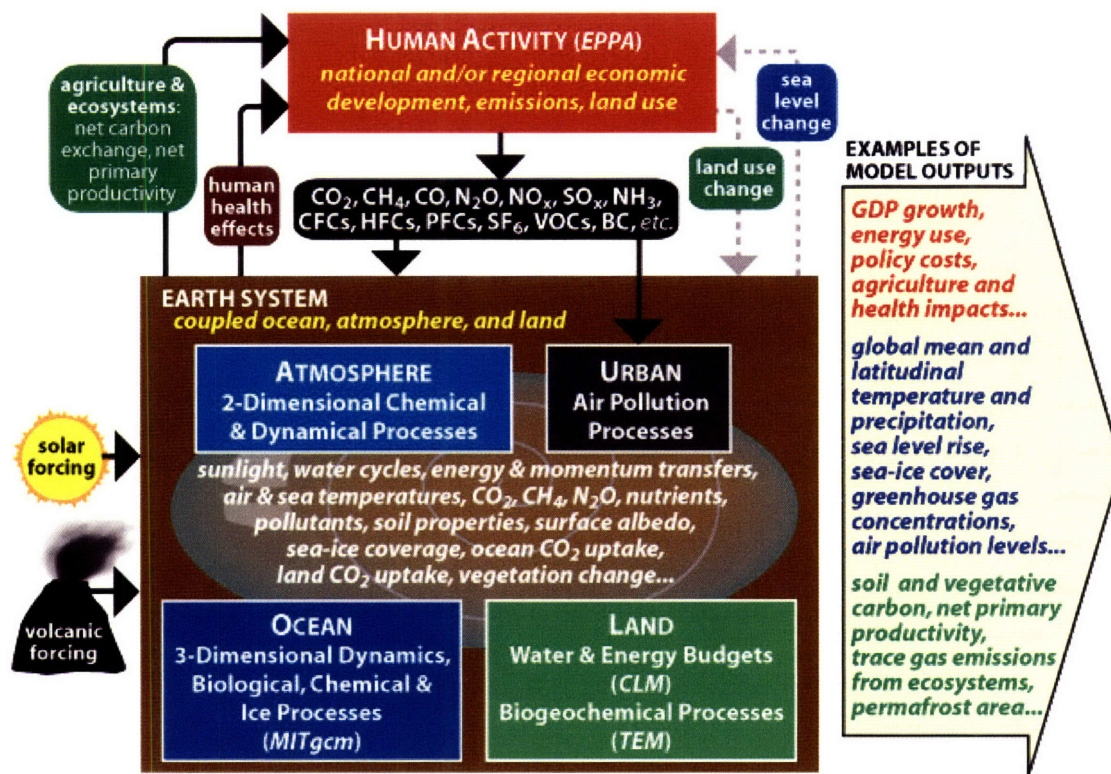
5

Figure 1-1: The schematic depicts the current framework and processes of the MIT Integrated Global System Model Version 2 (IGSM2). Feedbacks between the component models that are under development are shown as dashed lines.

based on altitude: the atmosphere and two ocean layers (the mixed layer and the deep ocean); and three regions based on latitude: southern high latitude, tropics, and northern high latitude. We then take 1000 reference runs of the IGSM (which were computed previously for other research and testing the model) and calculate the average temperature and carbon concentration in each region for each year between 1991 and 2100. These simulations have a wide range of values for input parameters and so provide a rich data set for calibrating such a simple climate model. We then formulate a model whose coefficients can be constrained by data and that will depend on input parameters from the IGSM estimating the temperature and carbon concentration. Our goal is to use results from the 1000 IGSM runs to calibrate a simple energy balance model (and carbon cycle) to use with the EPPA Model or in independent climate research.

1.1 Objective

1.1.1 A Reduced-Form Model of Climate Response to Carbon Emissions

Our simple climate model divides the world into nine regions by latitude (southern high latitude, tropics, and northern high latitude) and altitude (atmosphere, mixed layer, and deep ocean). This allows heat and carbon to transfer from one region to adjacent regions.[9] So the temperature or carbon concentration of each region is a function of the temperature or carbon concentration in the previous period plus a constant multiplied by the difference in temperatures or carbon concentrations between all the adjacent regions and the region in question. The multiplicative constant is a function of global temperatures and carbon concentrations in the period. The generic equation for the carbon concentration in a box at time t is:

$$C_{b,t+1} = C_{b,t} + E_{b,t} - U_{b,t} + \sum_{b.adj} \alpha_{b.adj,b} \times (C_{b.adj,t} - C_{b,t})$$

Equation 1.1

C: carbon concentration in box b at time t

E: carbon emissions

U: uptake of carbon by land

The E and U terms are zero for the mixed layer and deep ocean boxes, as there are no emissions and no land uptake there. The temperature in these regions changes due to flux from the mixed layer only. The part of the equation consisting of C, E, and U is the same as the one box model. The additional term in the summation, $\alpha_{b.adj,b}$, represents the "transfer coefficient" for carbon between the two neigboring boxes b_{adj} and b. These α coefficients are the calibration coefficients and depend on climate sensitivity (CS), ocean diffusivity (KV), and aerosol forcing (FA). They serve to simplify the physics and chemistry of the nonlinear system. The dependency of the α terms on CS, KV, and FA provides a means to mimic the behavior of the full IGSM.

The generic equation for temperature has a similar form:

$$T_{b,t+1} = T_{b,t} + F_{b,t} + \sum_{b.adj} \alpha_{b.adj,b} \times (T_{b.adj,t} - T_{b,t})$$

Equation 1.2

T: temperature in box b at time t

F: temperature forcing at time t.

The problem here is one of uncertainty analysis. Given any carbon emissions, we cannot exactly determine the impact of these emissions on the climate. This depends on the value of certain environmental variables, such as climate sensitivity (how sensitive the climate is to changes in atmospheric CO_2 concentrations) and ocean diffusivity (how easily the ocean absorbs atmospheric CO_2 and heat). Each of the 1000 runs of the IGSM consists of an estimate for each such environmental variable. By taking all 1000 runs together, we are able to build a deterministic model for future values of temperature, atmospheric carbon content, etc. by calibrating the exchange rates for energy

and CO_2 between regions.[5]

As equations 1.1 and 1.2 show, the carbon concentration in box b at time t+1 is equal to the total of the carbon concentration in box b at time t and the sum of the flux between box b and each box neighboring b (either immediately to the left, to the right, above, or below b). The flux between b and $b_{adjacent}$ is the difference in carbon concentrations between b and $b_{adjacent}$ at time t multiplied by a transfer coefficient α. The form for the temperature equations is the same, as the "transfer" of temperature from a box to its neighbor could be thought of as a heat flux, thus allowing us to approach the temperature problem in the same manner as the carbon problem.

1.1.2 Goals of This Work

The goal of this work is to come up with improvements for estimates of the carbon concentration and temperature (as functions of inputs such as emissions, climate sensitivity, aerosol forcing, and ocean uptake) to apply in conjunction with the EPPA Model. This will allow for more precise values of those variables to be obtained more quickly, making EPPA an effective stand-alone teaching and policy analysis tool, improving the rate at which we make environmental and economic calculations, and increasing the efficiency of our research.

Our work will also be used outside EPPA. Environmental scientists can use this model to study climate uncertainty. Given probability distribution functions (PDFs) of unknown values such as climate sensitivity and ocean diffusivity, this work would allow us to quickly create PDFs of other variables, for example the global mean temperature in the year 2100.

This model will also provide a teaching tool for MIT environmental economics classes. The class is currently conducted by allowing students to solve problem sets by running simplified versions of climate and economics models. A fast-running, simple, yet accurate climate model, such as the one we hope to develop here, would be what these students need to study the concepts relevant to environmental science.

Chapter 2

Research Design

2.1 Stages of Climate Models

In this section, we give a brief overview of the climate models that are used by environmental researchers today and attempt to explain how our model is an improvement of the others.

2.1.1 Atmosphere Ocean General Circulation Model (AOGCM)

General Circulation Models (GCMs) are a class of computer-driven models for predicting climate change. The term "general circulation" refers to the large-scale circulations that exist in the atmosphere and ocean. GCM Versions designed for decade to century time scale climate applications were originally created by Syukuro Manabe and Kirk Bryan at the Geophysical Fluid Dynamics Laboratory in Princeton, NJ. These computationally intensive numerical models are based on the integration of a variety of fluid dynamical and chemical equations.[10]

There are both atmospheric GCMs (AGCMs) and ocean GCMs (OGCMs). An AGCM and an OGCM can be coupled together to form an atmosphere-ocean coupled general circulation model (AOGCM). With the addition of other components (such as a sea ice model or a model for evapotranspiration over land), the AOGCM becomes the basis for a full climate model. Within this structure, different variations can exist, and their varying response to climate change may be studied.[10]

AOGCMs operate by discretising the equations for fluid motion and integrating these forward in time. They also contain parametrisations for processes – such as convection – that occur on scales too small to be resolved directly. AOGCMs are used in lots of high-level policy analysis tasks and form one scientific basis for the Integovernmental Panal on Climate Change (IPCC). AOGCMs often do not include representations of the carbon cycle.[10]

AOGCMs represent the pinnacle of complexity in climate models and internalise as many processes as possible. They are the only tools that provide detailed regional predictions of future climate

change. However, given their attempt to capture all the nuances of the climate system, they are often difficult to analyze and explain. The simpler models, such as the simple energy balance equation described below, are generally susceptible to straightforward analysis and their results are generally easier to understand.[10]

2.1.2 Earth System Model of Intermediate Complexity (EMIC)

Earth System Models of Intermediate Complexity (EMICs) are similar to AOGCMs to the degree that they are very complex. However, the atmosphere and ocean components of EMICs are simpler than in AOGCMs and remove the least significant terms to improve computational efficiency and permit including better representation of biogeochemistry and the carbon cycle. EMICs are currently used for global (but not local) policy analysis, to determine the impact of policies such as the 1997 Kyoto Protocol on atmospheric carbon concentrations and worldwide temperature change.

EMICs have been developed by many research groups, including Global Analysis Integration and Modeling (GAIM)[11] and the International Geosphere-Biosphere Program (IGBP).[12] The EMIC used for most of the MIT's climate studies is the MIT Integrated Global System Model (IGSM).[5] The output of this model serves as the basis of the research presented here.

2.1.3 Simple Energy Balance Equation Model

In this work, we attempt to develop a simple energy balance equation model of the impact of climate change on the carbon cycle. This model would be simple enough to analyze by hand or with an easily written computer program, and could easily be explained to an intelligent individual familiar with climate science. However, unlike past simple energy balance equation models, we intend for our work to cover the movement of carbon dioxide and heat across latitude and altitude bands by introducing "box to box" carbon and heat fluxes. This would allow us to analyse the physics and chemistry unique to certain parts of the Earth, as in EMICs such as the IGSM, while leaving our works simple enough for others to easily understand.

2.1.4 Simple Statistical Models

Simple statistical models attempt to find a simple mathematical relationships between human-controlled variables (e.g. emissions) and environmental variables (e.g. temperature, CO_2 concentration). Some of these functions are used in the MIT Toy Model.[8] It is the goal of this work to replace those functions with the stronger, physics-based energy balance equations. While it is possible to write statistical models in a way that would produce little error for an individual data set (using regression methods), the generated equations are of little use when the input (emissions, climate paramters, etc.) is slightly larger or smaller than the data from which the model was built.

This is because we attempt to produce a locally (for the data from which we are building the model) linearized simulation of a non-linear system, and outputs of non-linear sytems are highly sensitive to changes in initial conditions or parameters. Thus our locally correct simulation cannot be globally extrapolated.

2.2 Key Features to Capture in a Reduced-Form Model

Before the Industrial Revolution, the Earth had a roughly constant amount of atmospheric carbon dioxide for many millenia. However, the last few centuries brought industries that greatly increased the amount of carbon dioxide emissions per year. This resulted in an increase in the Earth's atmospheric and oceanic CO_2 concentrations. It also lead to climate change due to the greenhouse effect (as the concentration of greenhouse gases increases, the atmosphere is able to absorb more of the heat radiated from the sun). In the future, as industry continues to expand, we expect more and more carbon to be stored in the atmosphere and the ocean, and consequently more and more global warming. Understanding the magnitude of the CO_2 concentrations and global temperature changes is essential for policy makers to realize the scope of the problem and offer potential solutions.

2.2.1 More About the Process We Are Trying to Capture

The carbon cycle is the biogeochemical cycle by which carbon is exchanged between the biosphere, geosphere, hydrosphere, and atmosphere of the Earth. It is usually thought of as four major reservoirs of carbon interconnected by pathways of exchange. The reservoirs are the atmosphere, the terrestrial biosphere (which usually includes freshwater systems and non-living organic material, such as soil carbon), the oceans (which includes dissolved inorganic carbon and living and non-living marine biota), and the sediments (which includes fossil fuels). The annual movements of carbon, the carbon exchanges between reservoirs, occur because of various chemical, physical, geological, and biological processes. The ocean contains the largest active pool of carbon near the surface of the Earth, but the deep ocean part of this pool does not rapidly exchange with the atmosphere.[10]

Since the industrial revolution, humans have altered the carbon cycle by burning more fossil fuels, thus increasing the rate of transfer of CO_2 from sediments to the atmosphere. Some of this CO_2 remains in the atmosphere, some is uptaken by the terrestrial biosphere, and some dissolves in the ocean, where it is converted to dissolved inorganic carbon (DIC). As atmospheric CO_2 concentrations rise, global warming takes place as the Earth becomes able to absorb more and more heat.[10]

There is uncertainty in the system because, while we can determine through economic analysis estimates of how much CO_2 industries will emit absent government regulation, we do not know the values of key variables that are involved, such as climate sensitivity (how sensitive atmospheric heat capacity is to CO_2 emissions), ocean diffusivity (how easily the ocean absorbs atmospheric CO_2 and

heat), and aerosol forcing (how strong the radiative forcing is for a given aerosol concentration).[13] By studying past data, we can estimate these values using fingerprinting techniques that filter "noise" from large data sets and produce probability distribution functions (PDFs) for them.[14] Thus, predicting future global carbon concentrations and temperatures is similar to spinning a roulette wheel (to determine the values of the uncertain variables). Depending on the values of climate sensitivity, ocean diffusivity, and aerosol forcing, a given CO_2 emission regime can either lead to insignificant changes in the global average temperature or increases as large as 5 degrees Celcius per century.[14]

In order to accurately estimate future atmospheric carbon concentrations and global temperatures, we need to know the values of variables such as ocean diffusivity, climate sensitivity, and aerosol forcing. In 2002, Forest et al.[15] presented results of a procedure for estimating these values. Forest et al. applied an optimal fingerprint detection algorithm to functions of past data on atmospheric and ocean temperatures and carbon concentrations to provide a 5-95 percent confidence interval estimate for each variable above that is consistent with the observed climate record. The 1000 runs of the IGSM were generated using 1000 combinations of climate sensitivity, ocean diffusivity, and aerosol forcing values from within Forest et al.'s bounds and also with 1000 unique emissions scenarios.[16] From each run, we calculate future values of CO_2 content and temperature based on the given variable values. The combination of all of these runs can be used to calibrate a deterministic model of the future values of environmental variables.[10]

The system we are modeling operates as follows. To obtain energy for heat, production, or transportation, households burn fossil fuels into carbon dioxide and water vapor. This excess CO_2 is then absorbed in the local atmosphere, and eventually spreads around the world. As atmospheric carbon concentrations increase, a certain amount is absorbed in the ocean and land, partially offsetting the increase by human emissions. The CO_2 and greenhouse gases that remain in the atmosphere cause the Earth's temperature to rise due to the increased greenhouse effect. This can lead to impacts on human activity, such as the destruction of agricultural crops and the rising of the sea level.[10]

2.2.2 Movement of Carbon Dioxide and Heat

Here we describe the energy balance climate model and the simple climate cycle model. As in other energy balance climate models, the world is divided into several regions or "boxes". This makes modeling task significantly simpler, as one can focus on individual boxes, instead of the whole Earth. However, it also makes the work more difficult as it requires us to account for the movement of carbon dioxide and heat.[10]

The "single box" model used before this work[8] treats the Earth as one continuous unit. Carbon and heat are added to this unit through forcing and taken out through uptake. Thus, the form of the mathematical equations representing the Earth's CO_2 concentration in a given time period is:

$C_{t+1} = C_t + E_t - U_t$

Equation 2.1

or

$\frac{dC}{dt} = E - U$

Equation 2.2

C: carbon dioxide

E: emissions

U: uptake of carbon by land

t: current time period

t+1: next time period.

The "single box" above contains land, air, and water in the tropical and high latitude regions. Each of these has different physical properties that limit its ability to hold carbon. Thus our function, while statistically accurate (according to our data) is doomed to fail with other data sets, as it unable to capture the science behind the process it is modeling.

In order to resolve the problem of capturing the chemistry and physics behind the system we are simulating, we attempt to divide the world into nine "boxes" (figure 2-1). The boxes are separated by altitude: atmosphere (above sea level), mixed layer(0 through 475 meters below sea level), and deep ocean (more than 475 meters below sea level); and by latitude: southern high latitude (-90 to -30 degrees), tropics (-30 to 30 degrees), and northern high latitude(30 to 90 degrees). The form of the equation representing the carbon concentration of each box is available in Appendix I: Carbon and Temperature Equations of this work. In general terms, these equations are written as:

$C_{b,t+1} = C_{b,t} + E_{b,t} - U_{b,t} + \sum_{b.adj} \alpha_{b.adj,b} \times (C_{b.adj,t} - C_{b,t})$

Equation 2.3

C: carbon concentration

E: carbon emissions

U: uptake of carbon by land

$\alpha_{b.adj,b}$: "transfer coefficient" for carbon between the two neigboring boxes b_{adj} and b

The E and U terms are zero for the mixed layer and deep ocean boxes, as there are no emissions and no land uptake there. The temperature in these regions changes due to flux from the mixed layer only. The part of the equation consisting of C, E, and U is the same as the one box model. The additional term in the summation $\alpha_{b.adj,b}$ represents the "transfer coefficient" for carbon between the two neigboring boxes b_{adj} and b. We assume that the total amount transferred, or "transfer flux" is equal to the product of the transfer coefficient and the difference of concentrations in the two boxes. In this work, we attempt to estimate values for these α terms using the data from the 1000 runs of the IGSM.

The temperature equations operate in a manner similar to the carbon equations. Except the

13

emissions and uptake terms are replaced with a radiative forcing term (in the atmosphere and mixed layer only), which the sum of the individual temperature forcings due to all the greenhouse gases and aerosol forcings. The equation for the temperature forcing, F_{temp} is:

$F_{temp} = F_{T,CO2} + F_{T,CH4} + F_{T,N2O} + F_{T,othergases} + F_{T,aer}$

Equation 2.4

$F_{T,gasI}$: temperature forcing due to gasI alone.

The temperature equations are thus written as:

$T_{b,t+1} = T_{b,t} + F_{b,t} + \lambda \times T_{b,t} + \sum_{b.adj} \alpha_{b.adj,b} \times (T_{b.adj,t} - T_{b,t})$

Equation 2.5

$T_{b,t}$: temperature in box b at time t.

$F_{b,t}$: represents the temperature forcing at time t.

λ: net feedback factor that determines equilibrium temperature change. It is equal to 3.8 divided by the climate sensitivity (CS).

Equation 2.5 is derived from the differential equation applied to each box:

$\frac{dT}{dt} = F + \lambda \times T + \phi$

Equation 2.6

ϕ: temperature flux into the box.

Since temperature forcing takes place only in the atmosphere and mixed layer, we assign this term a value of zero for deep ocean boxes. Note that our functions assume a "temperature flux" from b_{adj} to b. This is really a heat flux. We assume that the heat capacity of each of our boxes is constant, and temperature is therefore linearly related to heat. Thus, equation 2.5 could more formally be written as:

$T_{b,t+1} = T_{b,t} + F_{b,t} + \lambda \times T_{b,t} + \frac{1}{Q} \sum_{b.adj} \alpha_{b.adj,b} \times (H_{b.adj,t} - H_{b,t})$

Equation 2.7

Q: heat capacity

H: total amount of heat.

In our model, we assume that all heat is represented as temperature. Thus, $T = \frac{H}{Q}$, and the heat flux can be represented as a temperature flux.

The three box model, as shown in figure 2-2 on page 15 is based on the same idea as the nine box model. However, the former attempts to simplify the latter by removing some of the boxes, thus decreasing the total number of (b, b_{adj}) pairs. Thus, we removed the latitude separation among boxes and kept only three boxes by altitude: atmosphere, mixed layer, and deep ocean. The form of the equations remains the same, but there are fewer variables and functions per timestep. The rationale for this simplification is that, while the temperatures differ by latitude, the basic structure of the atmosphere, mixed layer, and deep ocean does not change much along the north-south axis. Thus, it is reasonable to treat these "boxes" as a single unit.

Figure 2-1: Regions in the Nine Box Model. In each time step, we assume that heat and carbon are transferred between adjacent boxes only. I.e. Heat can be transferred from Region 8 to Region 9, but not from Region 5 to Region 9.

Figure 2-2: Regions in the Three Box Model. In each time step, we assume that heat and carbon are transferred between adjacent boxes only. I.e. Heat can be transferred from Region 2 to Region 3, but not from Region 1 to Region 3.

2.2.3 Rationale for Moving from Simple Equation to "Box" Model

In this work, we model the world as a set of nine or three boxes, rather than using a single equation for the whole system. The reasoning for this is that each of the three or nine boxes has its own unique set of physical characteristics, which we want to include. Also, we are interested in a model that would capture the gradual diffusion of carbon emissions and heat throughout the world.

The three altitude boxes: atmosphere, mixed layer, and deep ocean, all have different physical characteristics. The atmosphere consists of air. That is where most of the CO_2 emissions from households and industry are directly placed. Also, this is where the heat or temperature forcing due to greenhouse gases occurs. The mixed layer, or the upper layer of the ocean, consists of water. Since this water is close to the surface, it absorbs heat and CO_2 from the atmosphere directly. Because the deep ocean is separated from the surface by 475 meters, it takes longer for changes in atmospheric CO_2 concentrations and temperatures to propagate to this level, as the carbon or heat must first be absorbed by the mixed layer. The reasoning for the separation of the mixed layer and deep ocean is that the mixed layer responds quickly to changes in the atmospheric temperature or CO_2 concentration and is "mixed" (dissolved molecules spread out through the whole mixed layer) fairly quickly (within one month for most GHGs). The deep ocean mixes and responds to changes in atmospheric and mixed layer CO_2 and temperature conditions much more slowly. As such, the mixed layer must be treated as separate from the deep ocean for biogeochemical constituents.

Heat and carbon are also forced or emitted at different rates in the northern, tropical, and southern regions of the atmosphere and mixed layer. In our three latitude boxes, we attempt to capture the movement of heat and carbon along the north-south axis.

Chapter 3

Steps in the Development of the Reduced-Form Model

3.1 Input Data

After we determined the basic form of the reduced-form model, there were still many steps to take before we were able to run and test our design.

3.1.1 Source of Data

To run our model, we needed future (up to the year 2100) values of carbon concentration and global average temperature given certain climate parameters (climate sensitivity, ocean diffusivity, and aerosol forcing). We obtained these numbers from the MIT Integrated Global System Model (IGSM).[5] As mentioned earlier, the IGSM is an Earth System Model of Intermediate Complexity (EMIC). In this work, we attempt to estimate the IGSM's output. If that model is later modified (for example, following the discovery of a new variable relevant to climate change), our project will have to be repeated to produce a model for the new data.

3.1.2 Data Preparation

The first major challenge we faced with this work was preparing the data. The IGSM data set is approximately 188 Gb in size, and required extraction of the relevant data for our project.

In our simulation, we originally accounted for carbon emissions (from industry, residential energy consumers, etc.) and ocean uptake. However, we also required the uptake of atmospheric CO_2 by land, as obtained from the IGSM.

3.1.3 Required Conversions

As mentioned earlier, all the data for this project was obtained from the IGSM. The latter model does not divide the earth into the nine boxes that we used. Instead, the IGSM uses 24 latitude bands and 20 vertical layers (nine in the atmosphere and 11 in the ocean). To determine the values we needed for each of our "boxes", we used weighted averages, with weights assigned according to either the surface area (for atmospheric data) or ocean volume (for ocean data). A weighted average using weights W for a data set X is given below:

$$X_{avg} = \sum_i (W_i \times X_i) / \sum_i W_i$$

Equation 3.1

Also, while most of the IGSM data is displayed in the one year time steps that we used, many of the GHG concentration and forcing data sets used five year time steps. For those data, we assumed linear growth during each five year period (e.g. between 2000 and 2005) and used linear interpolation to obtain estimates for the desired values during the years that were not covered. An example of the interpolation estimate for the value of variable X in 2001, given the values in 2000 and 2005 appears below:

$$X_{2001} = 0.8 \times X_{2000} + 0.2 \times X_{2005}$$

Equation 3.2

3.1.4 Dependence of Data on Climate Sensitivity, Ocean Diffusivity, and Aerosol Forcing

The increase in CO_2 concentration and temperature depends not only on emissions but also on the values of the uncertain climate parameters climate sensitivity (CS), ocean diffusivity (KV) and aerosol forcing (FA). We do not know these values. However, as described earlier, Forest et al.[15] provide a 5-95 precent confidence interval estimating them through the use of fingerprinting techniques that filter "noise" from large data sets of past values of atmospheric and ocean temperatures and carbon concentrations.[14] Each of the 1000 runs of the IGSM uses one such estimate. In this work, we test the hypothesis that global temperature depends on CS, KV, and FA. We find that a correlation does exist. (figures 3-1, 3-2, and 3-3) As demonstrated in figure 3-1 on page 19, given constant values of KV, FA, and CO_2 emissions, as CS increases, the rate of temperature change rises. Figure 3-2 on page 19 shows that, when CS, FA, and CO_2 emissions are held at fixed values, the rate of temperature change falls as KV increases. Figure 3-3 on page 20 illustrates the trend for FA. As aerosol forcing rises, the rate of temperature change falls, assuming constant values of the other parameters. Based on these observations, we include these values in our fits for mixed layer temperature.
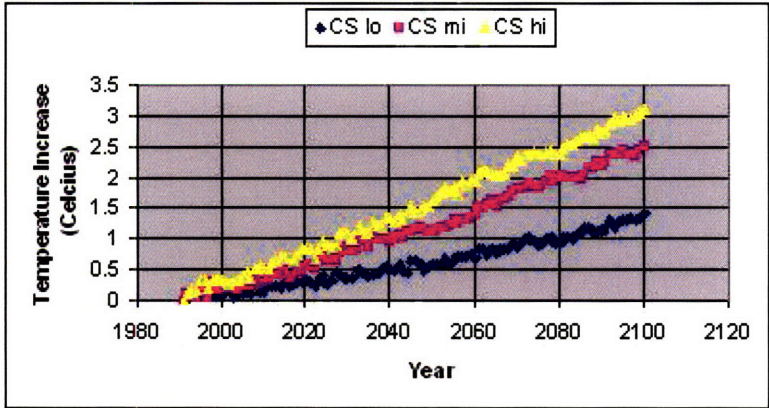
Figure 3-1: Dependence of mixed layer temperature on CS (CSlo=1.7, CSmi=2.5, CShi=3.8, KV=12, FA=0.58, atmospheric CO_2 concentration in 2100 = 720 ppm). In these three IGSM runs, KV and FA are set to their average values, while CS is varied between high, medium, and low values. Note that higher CS leads to greater temperature increases.
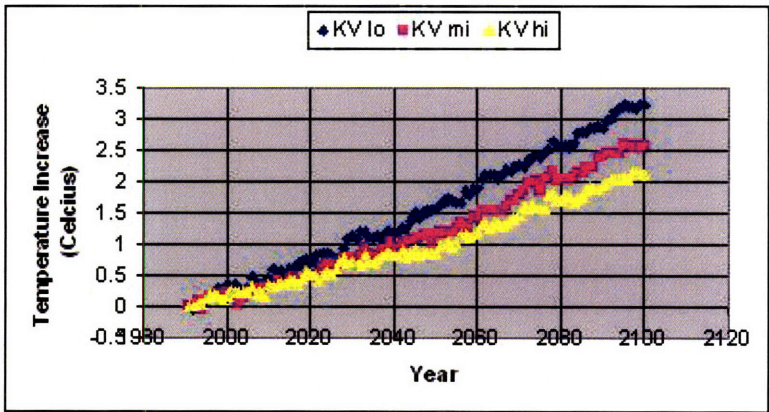

Figure 3-2: Dependence of mixed layer temperature on KV(CS=2.5, KVlo=9.3, KVmi=12, KVhi=16, FA=0.58, atmospheric CO_2 concentration in 2100 = 720 ppm). In these three IGSM runs, CS and FA are set to their average values, while KV is varied between high, medium, and low values. Note that higher KV leads to greater temperature increases.
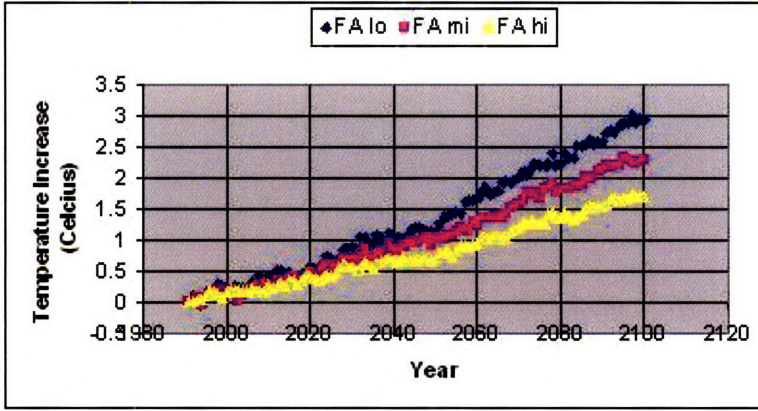
19

Figure 3-3: Dependence of mixed layer temperature on FA (CS=2.5, KV=12, FAlo=0.35, FAmi=0.58, FAhi=0.81, atmospheric CO_2 concentration in 2100 = 720 ppm). In these three IGSM runs, CS and KV are set to their average values, while FA is varied between high, medium, and low values. Note that higher FA leads to lesser temperature increases.

3.2 Function of the Forward and Estimation Models

We decided to approach the linear climate estimation problem in two steps: an estimation step, in which we determine the coefficients of the model equations; and a forward step, in which we calculate future data from the inputs.

The "forward model" tests the transfer flux coefficients produced in the estimation model (discussed in detail later in this document). The forward model's input includes the estimated parameters, and the CO_2 and DIC concentrations, temperatures, and forcings from only the first time step in the dataset. For the remaining time steps, the forcings are predicted as functions of concentrations in the previous period, and the concentrations and temperatures are predicted from the transfer coefficients and previous concentrations and forcings. Our model has a built-in test. If the estimation model is successful at producing reliable coefficients, the output of the forward model would be approximately equal to the original values (used as input to the estimation problem). Our forecasting equations calculate the CO_2 and temperature fluxes and add their total to the emissions or forcing terms to determine the next period's temperature and CO_2 concentration for each box. The formulas are given below.

$C_{b,t+1} = C_{b,t} + E_{b,t} - U_{b,t} + \sum_{b.adj} \alpha_{b.adj,b} \times (C_{b.adj,t} - C_{b,t})$

Equation 3.3

$T_{b,t+1} = T_{b,t} + F_{b,t} + \lambda \times T_{b,t} + \sum_{b.adj} \alpha_{b.adj,b} \times (T_{b.adj,t} - T_{b,t})$

Equation 3.4

The "estimation model" must be run before the forward model in order to generate the tranfer (α) coefficents that are used as input for the latter. The former program takes as input the carbon and temperature forcings, and CO_2 concentration, dissolved inorganic carbon (DIC) concentration,

20

and temperature for each "box". This data is then used to estimate the box-to-box carbon and heat transfer fluxes as either first or second order Taylor functions of climate sensitivity (CS), ocean diffusivity (KV), and aerosol forcing (FA). The estimation model is written in the GAMS programming language because the mathematical problem solving features in that langage allowed us to focus on the structure of the problem, rather than the mathematical and computational features of the solution. The GAMS software attempts to solve the carbon and temperature equations described in the previous section for the unknown α transfer coefficients. The equations used in the software are given below. Note that these are the same equations as in the forward model, except there is an additional ERROR term. GAMS attempts to select values for the α coefficients which minimize this ERROR quantity.

$C_{b,t+1} = C_{b,t} + E_{b,t} - U_{b,t} + \sum_{b.adj} \alpha_{b.adj,b} \times (C_{b.adj,t} - C_{b,t}) + ERROR_{C,b,t}$

Equation 3.5

$T_{b,t+1} = T_{b,t} + F_{b,t} + \lambda \times T_{b,t} + \sum_{b.adj} \alpha_{b.adj,b} \times (T_{b.adj,t} - T_{b,t}) + ERROR_{T,b,t}$

Equation 3.6

The CO_2 concentrations, temperatures, emissions, uptake, and forcing terms are loaded from data files from the IGSM and used as input for the solver. The ERROR term is added to the equations above because, since there are more equations than α terms, we will not be able to obtain an exact solution. A linear optimization program is used to find optimal values of the α terms such that the sum of squared errors (for all values of b and t) is minimized.

Given difficulties in assessing errors in the nine-box model, the three-box model was developed. The latter was modified further to take as input, rather than attempt to estimate, the carbon and temperature forcing terms, as we were unable to determine adequate fits for those values (an "adequate" fit would have given us forcing terms within 25 percent of those given in the IGSM), so we later modified the forward model to simply read in the forcing terms from the IGSM (like the estimation model). Our hope was that the differences between the original and model-produced carbon and temperature values would lead to negligible differences in forcing. After we did this, our three-box forward model accurately predicted the mixed layer temperatures (average error of 7 percent). However, we were unable to come up with adequate fits for deep ocean temperatures, or any CO_2 or DIC concentrations. Fixing this remains a task for future research.

The equation for estimating the value of the temperature forcing, which we later replaced with simply reading in the exact forcing coefficients from the IGSM, is given below.

$F_{temp} = F_{T,CO2} + F_{T,CH4} + F_{T,N2O} + F_{T,othergases} + F_{T,aer}$

Equation 3.7

$F_{T,gas}$: temperature forcing due to that gas alone.

The formulas for the temperature forcing due to CO_2, CH_4 and N_2O are given below:

CO_2: $\Delta F = 5.35 \ln(C/C_0)$

21

Equation 3.8

CH_4: $\Delta F = 0.036 \times (\sqrt{M} - \sqrt{M_0}) - (f(M,N_0) - f(M_0,N_0))$

Equation 3.9

N_2O: $\Delta F = 0.12 \times (\sqrt{N} - \sqrt{N_0}) - (f(M_0,N) - f(M_0,N_0))$

Equation 3.10

Where:

$f(M,N) = 0.47 \ln [1 + 2.01 \times 10^{-5} \times (MN)^{0.75} + 5.31 \times 10^{-15} \times M(MN)^{1.52}]$

Equation 3.11

C: CO_2 in ppm

M: CH_4 in ppm

N: N_2O in ppm

3.2.1 Why This Approach?

After obtaining the coefficients, we needed to quickly test whether they were accurate. Since we had the data from the 1000 IGSM runs, we thought that a program that took as input the data from the first time step and attempted to use the coefficients to reproduce the remaining time steps would be most effective, as it could easily be compared to the actual data and analyzed for error.

Chapter 4

Results

4.1 Comparison of Our Three and Nine Box Models with Simple Statistical Model

Our previous attempt to "simplify" the IGSMs climate model was a set of equations based on statistical analysis used in the MIT Toy Model.[8] While these formulas were reasonably accurate (usually within 30 percent of the IGSM data), they were based on a mathematical regression of the data, not on the science of the climate system. (To obtain the Toy Model equations, we allowed a mathematical program, i.e. MATLAB or STATA, to read the input values – CS, KV, FA, carbon emissions – and output values – CO_2 concentration and temperature – and asked it to produce "best fit" functions for the output in terms of the input.) Thus, the Toy Model formulas would fail if tested on a policy scenario which involved significant reductions of future emissions (e.g. the Kyoto Protocol).

While the model developed here is scientifically sound, it is only accurate for mixed layer temperatures, as the results it gives for other values are much greater than those predicted by the IGSM. The carbon cycle portion has not been implemented. Fortunately, the deep ocean temperatures are not used for policy research and can be ignored so long as the other parts of the model are correct. Future research could improve on this work to come up with a more accurate model.

4.2 Error Analysis

In this section we attempt to determine why our three and nine box models are coming up with such large errors and offer ideas future researchers could use to try to fix it. We use the following function to determine the error for a single term, $X_{r,b,t}$:

ERRinX$_{r,b,t}$ = (X$_{FWD,r,b,t}$ - X$_{IGSM,r,b,t}$) / X$_{IGSM,r,b,t}$

Equation 4.1

ERRinX$_{r,b,t}$: Error in variable X for run r, box b, time t.

X$_{FWD,r,b,t}$: Value of variable X for run r, box b, time t according to our forward model.

X$_{IGSM,r,b,t}$: Value of X for run r, box b, time t according to IGSM.

To determine the **total error** for a single variable (CO_2 concentration or temperature), we took the sum of squared errors (as defined above) for the variable for each run and box in the last year of the model (2100) only. We defined the **average error** as the total error divided by the number of samples it comprises.

Unfortunately, most of our results were not accurate. The average error was greater than 25 percent for all values except the mixed layer temperature for the three box model, which had an average error of 7 percent. In the sections that follow, we discuss several attempts to fix the errors, the reasoning behind the attempts, and the results.

4.2.1 Increase Time Step from One to Five Years

When we first encountered the unexpectedly large errors in the predicted data, we thought that they were due to "noise" in our system – i.e. random variation in CO_2 concentrations and temperatures from year to year. In order to minimize these fluctuations, we replaced the 110 annual samples we had (for the years 1991-2100) with one average sample for each five year period (i.e. one sample for 1991-1995, one for 1996-2000, etc.). We hoped that this would "even out" the noise, and result in a smoother data set that was more susceptible to our linear fits. Unfortunately, implementing this idea did not decrease the errors as we hoped. Thus, we returned to our earlier annual time step model and continued to look for other potential error sources.

4.2.2 Add Carbon Monoxide and Methane to the System

Carbon dioxide is not the only gas that contributes to global temperature increase. Other greenhouse gases, such as carbon monoxide (CO) and methane (CH_4) are also active contributers. Thus, we thought that the error might be decreased if we included these gases in our nine box and three box models. We attempted to produce functions for the atmospheric CO and CH_4 concentrations based on emissions, land and ocean uptake, and an initial condition. However, we were unsuccessful at doing this, as our functions had high error, as defined above.

We did, however, take advantage of the idea that gases other than CO_2 could be involved in global warming. As mentioned earlier, our equation for temperature forcing accounts for forcing due to CO2, CO, N_2O, other greenhouse gases, and aerosols. By coupling this temperature forcing equation with IGSM values for gas concentrations, we were able to fairly accurately (average error of 7 percent) predict the mixed layer temperature in the three box model.
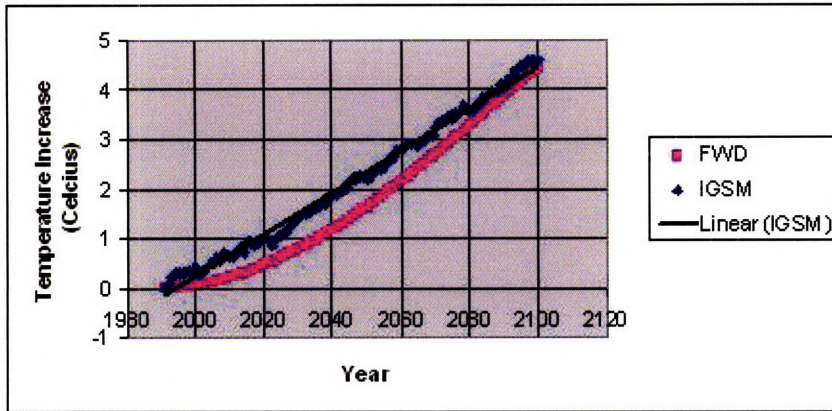
Figure 4-1: Graph of temperature increase vs time from IGSM and Forward Model (Run 1, CS=3.7, KV=2.6, FA=0.38, atmospheric CO2 concentration in 2100 = 763 ppm). The blue line represents the temperature increase predicted by the IGSM. The pink line represents our forward model. The black line is a basic linear regression fit to the IGSM data.

4.2.3 Separate Carbon Concentration into Carbon Dioxide and Dissolved Inorganic Carbon

Carbon is not stored in the same form throughout the regions in our model. In the atmosphere, most atmospheric carbon is gaseous CO_2. Deep ocean carbon is dissolved in water. It is mostly in the form of dissolved inorganic carbon (DIC). In the mixed layer, CO_2 and DIC coexist. We thus separated our original "carbon concentration" data set into two sets "CO_2" and "DIC", by obtaining and analyzing appropriate files from the IGSM.

4.2.4 Account for "Spin Up" in the System

Our simplified nine and three box models use "spin up" to the degree that they adjust based on carbon and temperature forcing and assume that there was no forcing before the starting year (1991), at which point an instantaneous forcing takes place. When we examined some of the mixed layer temperature vs time graphs, (figure 4-2) we noticed that, while the mathematical function our forward model (pink line) was coming up with for the IGSM mixed layer temperature (blue line) was reasonable, in some cases a closer linear fit (black line) was imaginable. Notice how the forward model (pink) curve becomes approximately parallel to the IGSM (blue) curve around the year 2040. This happens because by 2040, the forcing adjustment due to spin up is completed and the forcing (slope of the curve) is finally correct.

In theory, if we started the model in an earlier year, for example in 1860, rather than in 1991, the spin up problem would resolve by 1991 and allow us to come up with a better prediction of the temperature. To attempt this modification, we obtained the 1860-1990 data on CO_2 emissions and concentration and temperature and attempted to calculate our model coefficients from the full

25

1860-2100 dataset. Unfortunately, the 1860-2100 forward model produced temperatures many times higher than those suggested by the IGSM. Thus, we were compelled to abandon this idea.

4.3   Analysis of Nine Box Model Performance

The nine box model attempts to use simple mathematical functions to predict carbon concentrations and temperatures from emissions for nine world regions separated by latitude and altitude. The model was written from data in the Integrated Global System Model (IGSM), and we attempt to minimize the error, the difference between the actual IGSM data and that predicted by the simple nine box model.

However, the carbon concentration and temperature from the nine box model do not match what we expected. The nine box forward model output values are several times larger than those in the IGSM, and the multiplicative error increases with time. This error rate is not acceptable, and a better fit will have to be found if we are to use this simplified model.

Our task was to create a simple mathematical model that would predict the output of the complex IGSM within reasonable bounds. This is essentially a statistical estimation problem, with a structure determined by the physics and chemistry of the carbon cycle. When we determined our model, it made sense based on our understanding of the carbon cycle and our rough overview of the data. It seemed that there were relatively few variables that we were leaving out, and those variables could only have had a small impact on the atmospheric temperatures and carbon concentrations.

However, after we ran the estimation model and built the forward model, the results we were getting were several orders of magnitude greater than those produced by the IGSM. The several modifications that we attempted, discussed above, were unsuccessful at alleviating this problem. Thus, we conclude that the carbon cycle is full of complexities that cannot be efficiently represented in a simple mathematical system such as the one we have here.

4.4   Justification for Simpler Three Box Model

In the three box model, we attempt to take some of the complexity out of the nine box model while still retaining its basic structure. Thus, while keeping the box-to-box heat and carbon transfer equations the same, we modify the box division structure. We keep the three altitude regions (atmosphere, mixed layer, and deep ocean), but no longer subdivide them into latitude bands. (See figure 2-1 on page 15 and figure 2-2 on page 15 for an illustration of the different box structures.) This allows us to ignore north-south fluxes of carbon and heat and only focus on up-down transfers. Also, the three box model allows us to ignore the fact that heat always flows from the tropics to the poles, since the tropics have a much higher average temperature. It was our hope that the simplicity

of the three box model would make error analysis and correction simpler than it was in the nine box case.

4.5   Analysis of Three Box Model Performance

4.5.1   Three Box Model with Endogenous Carbon Cycle

In our first attempt at the three box model, we tried to make the simulation "capture" the carbon cycle; that is, we tried to input a set of CO_2 concentrations and temperature forcings values for the starting year (1991) and then calculate future values from those variables from previous timesteps.

For example, the 2010 CO_2 concentration would be a function of the 2009 CO_2 concentration and the 2009 CO_2 emissions. The 2010 temperature would be calculated from the 2009 temperature and temperature forcing (which is a function of the 2009 CO_2 concentration, aerosol forcing, and other variables, as mentioned earlier).

Unfortunately, this model produced output several orders of magnitude greater than the IGSM values. We noticed that this was due, in part, to our forcing estimates, which were also several order of magnitudes greater than desired. Thus, we decided to write a new version of the three box model, which reads as input (from IGSM values) rather than attempting to calculate the forcing terms.

4.5.2   Three Box Model with Exogenous Carbon Cycle

The new version of the three box model did not attempt to endogenously "capture" the carbon cycle. Instead, it inputted all the carbon cycle values (i.e. CO_2 concentrations, temperature forcings, etc.) from the IGSM data, giving us an exact replica of the IGSM's carbon cycle. This change led to significant improvements in our model and allowed us to accurately estimate (with 7 percent average error, as defined above) the mixed layer temperature.

Some graphs of the differences between the IGSM temperature and forward model temperature are available in figure 4-2 on page 28, figure 4-3 on page 29, and figure 4-4 on page 29. Notice that, in these graphs, the output of our forward model (pink dots) is close enough to the original IGSM values (blue dots) to serve as a reliable estimate. However, as discussed earlier, the forward model does not necessarily produce the "best fit curve" for the IGSM values. As mentioned in the "spin up" section of this work, we attempted to resolve this issue but encountered too many new errors in the process. We are still unable to figure out a suitable equation for deep ocean temperature and carbon concentration. For those values, there is a substantial multiplicative error that increases with time. The rate of increase (and thus the amount of error) for these values in the three box model is much smaller than that for the nine box model.

Even the three box model was more complex than we had envisioned and our forcing, uptake,
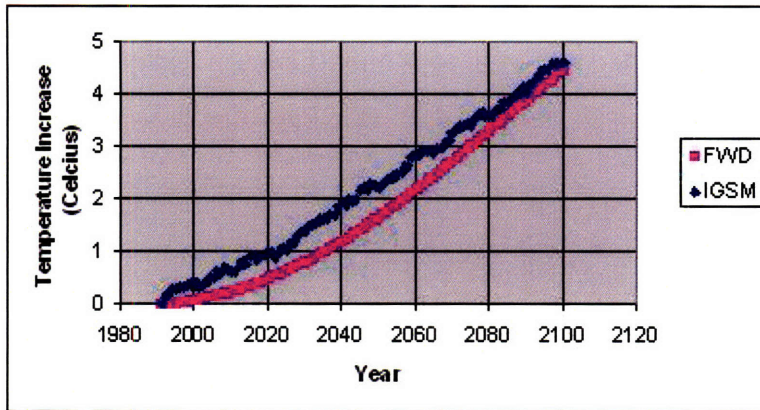
Figure 4-2: Graph of temperature increase vs time for Run 1 (CS=3.7, KV=2.6, FA=0.38, atmospheric CO2 concentration in 2100 = 763 ppm). This run has an average error (sum of squares of absolute differences between IGSM and forward temperatures). The blue line represents the temperature increase predicted by the IGSM. The pink line represents our forward model.

and transfer coefficients could not cover it effectively, as they were producing results several orders of magnitde greater than the IGSM data for all values except the mixed layer temperature. Some additional terms are necessary to explain the impact that carbon dioxide emissions are having on global temperatures and carbon dioxide concentrations. These additional terms might be the emissions of other gases and the rates of their conversion to CO_2, the impact these additional gases have on climate change, or something else.

Future work in this modelling task should focus on the three box model as it is simpler than the nine box model and the system is easier to understand (as north-south fluxes need not be accounted for). Unfortunately, our experience shows that simplification is not enough to make this problem easily solvable.
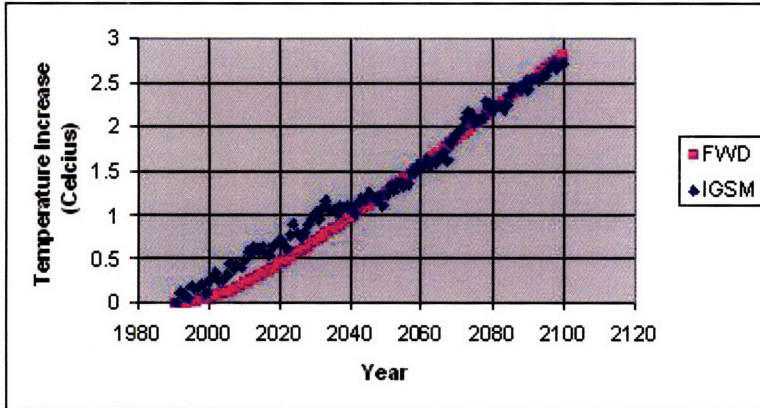
Figure 4-3: Graph of temperature increase vs time for Run 10 (CS=2.3, KV=3.8, FA=0.67, atmospheric CO2 concentration in 2100 = 715 ppm). This run has a lower error (sum of squares of absolute differences between IGSM and forward temperatures) than most other runs. The blue line represents the temperature increase predicted by the IGSM. The pink line represents our forward model.
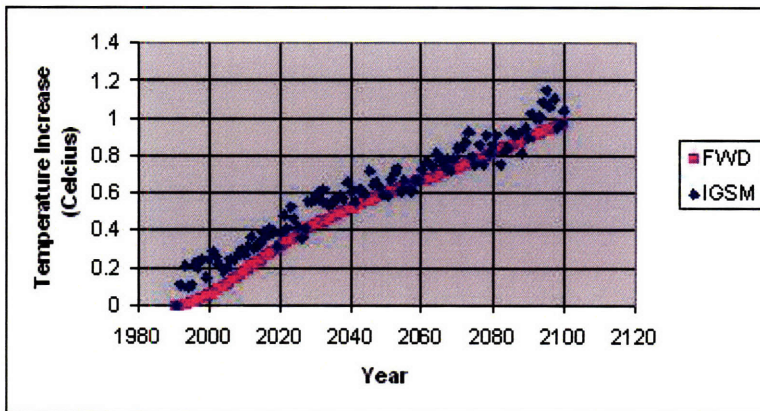

Figure 4-4: Graph of temperature increase vs time for Run 100 (CS=2.2, KV=32, FA=0.71, atmospheric CO2 concentration in 2100 = 500 ppm). This run has a higher error (sum of squares of absolute differences between IGSM and forward temperatures) than most other runs. The blue line represents the temperature increase predicted by the IGSM. The pink line represents our forward model.

Chapter 5

Future Research

In this work, we attempt to develop a parametrized reduced-form model for estimating global CO_2 concentrations and temperatures given data regarding CO_2 emissions. We use 3x3 and 3x1 box models of the Earth, with three altitude levels (atmosphere, mixed layer, and deep ocean) and either three or one latitude levels (southern high latitude, tropics, and northern high latitude). While we believe that our work accounts for most of the variables relevant to climate change, we may have inadvertently left out some important values. In the future, additional quantities, such as emissions, uptake, and concentrations of other gases, rates of conversion of other gases to CO_2, etc. should be added to the model in an attempt to make it successfully estimate important climate coefficients. Also, this work could be repeated using different sets of boxes (for example, 2x1 or 2x3, with the two altitude boxes representing "atmosphere" and "ocean"). This could assist in simplifying the model and make it easier to accurately implement.

The basic approach used in our work was to first come up with equations based on climate science, and then use statistical techniques to fit the IGSM data to these equations. Future researchers might wish to try a different technique. For example, the data could be loaded into a statistical analysis program, such as MATLAB or STATA. The software in these programs could be used to find a "best fit" equation for outputs (i.e. current CO_2 concentration and temperature) in terms of inputs (i.e. past CO_2 concentration and temperature, emissions, and forcing). A forward model, similar to ours could then be written to test the effectiveness of the equations produced by the statistical software. After an efficient model is created (from statistical analysis), it would be the task of climate scientists to interpret the coefficients in the model and their relationship to climate, geographic, chemical, and physical constants.

Chapter 6

Conclusion

In this work, we attempt to develop an energy balanced climate model in which the control parameters have been calibrated to depend on global climate system properties. These control parameters (coefficients) will then depend on climate sensitivity, ocean diffusivity, and aerosol forcing, and provide a simple climate model that reproduces the behavior of the more complex IGSM. This project is an improvement over previous works of this nature because it treats the Earth as a set of several "boxes" separated by latitude and altitude, rather than as a single unit. As such, we hoped to be able to successfully represent industrial emissions (carbon forcing) and account for the movement of heat and carbon across latitude and altitude bands using simple mathematical equations which could easily be implemented in a fast-running computer program.

After it is successfully completed, this simulation, along with a similar one for economic variables, could be used to analyze potential policies aimed at reducing carbon dioxide emissions in the hope of preventing or slowing global warming and climate change. When successfully implemented, our model promises to have a much faster running time and smaller software requirement than the previously available MIT IGSM[5]. As such, it could prove to be a valuable tool for policy analysts.

Appendix I: Carbon and Temperature Equations

Nine Box Model Carbon Equations

$$
\begin{aligned}
C_{r,1,t+1} =& C_{r,1,t} + F_{r,t}^C \\
& - \alpha_{r,1,4}^C \times (C_{r,1,t} - C_{r,4,t}) \\
& - \alpha_{r,2,5}^C \times (C_{r,2,t} - C_{r,5,t}) \\
& - \alpha_{r,3,6}^C \times (C_{r,3,t} - C_{r,6,t})
\end{aligned}
\tag{6.1}
$$

$$
C_{r,2,t} = C_{r,1,t}
\tag{6.2}
$$

$$
C_{r,3,t} = C_{r,1,t}
\tag{6.3}
$$

$$
\begin{aligned}
C_{r,4,t+1} =& C_{r,4,t} \\
& + \alpha_{r,1,4}^C \times (C_{r,1,t} - C_{r,4,t}) \\
& - \alpha_{r,4,5}^C \times (C_{r,4,t} - C_{r,5,t}) \\
& - \alpha_{r,4,7}^C \times (C_{r,4,t} - C_{r,7,t})
\end{aligned}
\tag{6.4}
$$

$$
\begin{aligned}
C_{r,5,t+1} =& C_{r,5,t} \\
& + \alpha_{r,2,5}^C \times (C_{r,2,t} - C_{r,5,t}) \\
& + \alpha_{r,4,5}^C \times (C_{r,4,t} - C_{r,5,t}) \\
& - \alpha_{r,5,6}^C \times (C_{r,5,t} - C_{r,6,t}) \\
& - \alpha_{r,5,8}^C \times (C_{r,5,t} - C_{r,8,t})
\end{aligned}
\tag{6.5}
$$

	Southern High Lat	Tropics	Northern High Lat
Atmosphere	Region 1	Region 2	Region 3
Mixed Layer	Region 4	Region 5	Region 6
Deep Ocean	Region 7	Region 8	Region 9

Figure 6-1: Regions in the Nine Box Model.

$$
\begin{aligned}
C_{r,6,t+1} =& C_{r,6,t} \\
&+ \alpha^{C}_{r,3,6} \times (C_{r,3,t} - C_{r,6,t}) \\
&+ \alpha^{C}_{r,5,6} \times (C_{r,5,t} - C_{r,6,t}) \\
&- \alpha^{C}_{r,6,9} \times (C_{r,6,t} - C_{r,9,t})
\end{aligned}
\tag{6.6}
$$

$$
\begin{aligned}
C_{r,7,t+1} =& C_{r,7,t} \\
&+ \alpha^{C}_{r,4,7} \times (C_{r,4,t} - C_{r,7,t}) \\
&- \alpha^{C}_{r,7,8} \times (C_{r,7,t} - C_{r,8,t})
\end{aligned}
\tag{6.7}
$$

$$
\begin{aligned}
C_{r,8,t+1} =& C_{r,8,t} \\
&+ \alpha^{C}_{r,5,8} \times (C_{r,5,t} - C_{r,8,t}) \\
&+ \alpha^{C}_{r,7,8} \times (C_{r,7,t} - C_{r,8,t}) \\
&- \alpha^{C}_{r,8,9} \times (C_{r,8,t} - C_{r,9,t})
\end{aligned}
\tag{6.8}
$$

$$
\begin{aligned}
C_{r,9,t+1} =& C_{r,9,t} \\
&+ \alpha^{C}_{r,6,9} \times (C_{r,6,t} - C_{r,9,t}) \\
&+ \alpha^{C}_{r,8,9} \times (C_{r,8,t} - C_{r,9,t})
\end{aligned}
\tag{6.9}
$$

Nine Box Model Temperature Equations

$$T_{r,1,t} = T_{r,4,t} \tag{6.10}$$

$$T_{r,2,t} = T_{r,5,t} \tag{6.11}$$

$$T_{r,3,t} = T_{r,6,t} \tag{6.12}$$

$$
\begin{aligned}
T_{r,4,t+1} = {} & T_{r,4,t} - \lambda_r \times T_{r,t}^{AVG} + F_{r,t}^T \\
& - \alpha_{r,4,5}^T \times (T_{r,4,t} - T_{r,5,t}) \\
& - \alpha_{r,4,7}^T \times (T_{r,4,t} - T_{r,7,t})
\end{aligned} \tag{6.13}
$$

$$
\begin{aligned}
T_{r,5,t+1} = {} & T_{r,5,t} - \lambda_r \times T_{r,t}^{AVG} + F_{T,r,t} \\
& + \alpha_{r,4,5}^T \times (T_{r,4,t} - T_{r,5,t}) \\
& - \alpha_{r,5,6}^T \times (T_{r,5,t} - T_{r,6,t}) \\
& - \alpha_{r,5,8}^T \times (T_{r,5,t} - T_{r,8,t})
\end{aligned} \tag{6.14}
$$

$$
\begin{aligned}
T_{r,6,t+1} = {} & T_{r,6,t} - \lambda_r \times T_{r,t}^{AVG} + F_{r,t}^T \\
& + \alpha_{r,5,6}^T \times (T_{r,5,t} - T_{r,6,t}) \\
& - \alpha_{r,6,9}^T \times (T_{r,6,t} - T_{r,9,t})
\end{aligned} \tag{6.15}
$$

$$
\begin{aligned}
T_{r,7,t+1} = {} & T_{r,7,t} \\
& + \alpha_{r,4,7}^T \times (T_{r,4,t} - T_{r,7,t})
\end{aligned} \tag{6.16}
$$

$$
\begin{aligned}
T_{r,8,t+1} = {} & T_{r,8,t} \\
& + \alpha_{r,5,8}^T \times (T_{r,5,t} - T_{r,8,t})
\end{aligned} \tag{6.17}
$$

$$
\begin{aligned}
T_{r,9,t+1} = {} & T_{r,9,t} \\
& + \alpha_{r,6,9}^T \times (T_{r,6,t} - T_{r,9,t})
\end{aligned} \tag{6.18}
$$

$$T_{r,t}^{AVG} = \frac{1}{9} \sum_{b=1}^{9} T_{r,b,t} \tag{6.19}$$

Nine Box Model Forcing Equations

$X_{r,t}$: CH4 concentration

$Y_{r,t}$: N2O concentration

$$F_{r,t}^{CO2} = 5.35 \times \log X_{r,t} \tag{6.20}$$

$$G_{r,t} = 0.47 \times \log(1 + 2.01 \times 10^{-5} \times (X_{r,t} \times Y_{r,t})^{0.75}$$
$$+ 5.31 \times 10^{-15} \times X_{r,t} \times (X_{r,t} \times Y_{r,t})^{1.52}) \tag{6.21}$$

$$F_{r,t}^{CH4} = 0.036 \times \sqrt{X_{r,t}} - G_{X,Y} \tag{6.22}$$

$$F_{r,t}^{N2O} = 0.12 \times \sqrt{Y_{r,t}} - G_{X,Y} \tag{6.23}$$

$$F_{r,t}^{F11} = 0.25 \times X_{r,t} \tag{6.24}$$

$$F_{r,t}^{F12} = 0.32 \times X_{r,t} \tag{6.25}$$

$$F_{r,t}^{C} = F_{r,t}^{CO2} + F_{r,t}^{CH4} \tag{6.26}$$

$$F_{r,t}^{T} = F_{r,t}^{CO2} + F_{r,t}^{CH4} + F_{r,t}^{N2O} + F_{r,t}^{F11} + F_{r,t}^{F12} \tag{6.27}$$

Bibliography

[1] W. D. Nordhaus, *Managing the Global Commons: The Economics of Climate Change.* Cambridge, MA: MIT Press, 1994.

[2] R. Alley, "Contribution of working group i to the fourth assessment report of the intergovernmental panel on climate change: Summary for policymakers," tech. rep., Intergovernmental Panel on Climate Change, Geneva, Switzerland, April 2007.

[3] F. Joos, G. Muller-Furstenberger, and G. Stephan, "Correcting the carbon cycle representation: How important is it for the economics of climate change?," *Environmental Modeling and Assessment*, vol. 4, pp. 133–140, 1999.

[4] H. Jacoby, R. Schmalensee, and I. S. Wing, "Toward a useful architecture for climate change negotiations," Tech. Rep. 49, MIT Joint Program on the Science and Policy of Global Change, Cambridge, MA, May 1999.

[5] A. Sokolov, C. A. Schlosser, S. Dutkiewicz, S. Paltsev, D. W. Kicklighter, H. D. Jacoby, R. G. Prinn, C. Forest, J. Reilly, C. Wang, B. Felzer, M. C. Sarofim, J. Scott, P. H. Stone, J. M. Melillo, and J. Cohen, "The mit integrated global system model (igsm) version 2: Model description and baseline evaluation," Tech. Rep. 124, MIT Joint Program on the Science and Policy of Global Change, Cambridge, MA, July 2005.

[6] S. Paltsev, J. Reilly, R. E. H.D. Jacoby, J. McFarland, M. Sarofim, M. Asadoorian, and M. Babiker, "The mit emissions prediction and policy analysis (eppa) model: Version 4," Tech. Rep. 125, MIT Joint Program on the Science and Policy of Global Change, Cambridge, MA, August 2005.

[7] W. D. Nordhaus and J. Boyer, *Warming the World: Economic Models of Global Warming.* Cambridge, MA: MIT Press, 2003.

[8] G. Rabin, *MIT Toy Model.* MIT Joint Program on the Science and Policy of Global Change, Cambridge, MA, March 2005.

[9] G. Holian, A. P. Sokolov, and R. G. Prinn, "Uncertainty in atmospheric co2 predictions from a parametric uncertainty analysis of a global ocean carbon cycle model," Tech. Rep. 80, MIT Joint Program on the Science and Policy of Global Change, Cambridge, MA, September 2001.

[10] J. T. Kiehl and V. Ramanathan, *Frontiers of Climate Modeling*. New York, NY: Cambridge University Press, 2006.

[11] D. Sahagian and B. Moore, "Toward the integration of global biogeochemical systems," *Global Biogeochemical Cycles*, vol. 10, no. 4, p. 162, 1996.

[12] J. Canadell, D. Pataki, and L. Pitelka, *Terrestrial Ecosystems in a Changing World*. Heidelberg, Germany: Springer-Verlag, 2007.

[13] J. Houghton, *Global Warming: The Complete Briefing*. New York, NY: Cambridge University Press, 3 ed., 2004.

[14] C. Forest, P. Stone, and A. Sokolov, "Estimated pdfs of climate system properties including natural and anthropogenic forcings," Tech. Rep. 126, MIT Joint Program on the Science and Policy of Global Change, Cambridge, MA, September 2005.

[15] C. Forest, P. Stone, A. Sokolov, M. Allen, and M. Webster, "Quantifying uncertainties in climate system properties using recent climate observations," Tech. Rep. 78, MIT Joint Program on the Science and Policy of Global Change, Cambridge, MA, July 2001.

[16] M. Webster, C. Forest, J. Reilly, D. Kicklighter, M. Mayer, R. Prinn, M. Sarofim, A. Sokolov, P. Stone, and C. Wang, "Uncertainty analysis of climate change and policy response," Tech. Rep. 95, MIT Joint Program on the Science and Policy of Global Change, Cambridge, MA, December 2002.