The development of general circulation models of climate

Spencer Weart

American Institute of Physics, College Park, MD, USA

ARTICLE INFO

Article history:
Received 30 October 2009
Received in revised form 19 May 2010
Accepted 25 June 2010

Keywords:
Simulation
Model
Computer
Climate
Atmosphere
Circulation

ABSTRACT

With the coming of digital computers in the 1950s, a small American team set out to model the weather, followed by attempts to represent the entire general circulation of the atmosphere. The work spread during the 1960s, and by the 1970s a few modelers had produced somewhat realistic looking models of the planet's regional climate pattern. The work took on wider interest when modelers tried increasing the level of greenhouse gases, and invariably found serious global warming. Skeptics pointed to dubious technical features, but by the late 1990s these problems were largely resolved—thanks to enormous increases in computer power, the number and size of the closely interacting teams that now comprised the international modeling community, and the crucial availability of field experiments and satellite data to set against the models' assumptions and outputs. By 2007 nearly all climate experts accepted that the climate simulations represented reality well enough to impel strong action to restrict gas emissions.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Climate is made by the general circulation of the atmosphere—the global pattern of air movements, with air masses rising in the tropics to descend farther north, semi-tropical trade winds, cyclonic storms that carry energy and moisture through middle latitudes, and so forth. Many meteorologists suspected that shifts in this pattern were a main cause of climate change. They could only guess about such shifts, for the general circulation was poorly mapped before the 1940s (even the jet streams remained to be discovered). The Second World War and its aftermath brought a phenomenal increase in observations from ground level up to the stratosphere, which finally revealed all the main features. Yet up to the 1960s, the general circulation was still only crudely known.1

This knowledge was strictly observational. From the 19th century forward, many scientists had attempted to explain the general pattern by applying the laws of the physics of gases to a heated, rotating planet. All their ingenious efforts failed to derive a realistic mathematical solution. The best mathematical physicists could only offer simple arguments for the character of the circulation, arguments which might seem plausible but in fact were mere hand-waving (see Lorenz, 1967, 59ff).

The solution would come by taking the problem from the other end. Instead of starting with grand equations for the planet as a whole, one might seek to find how the circulation pattern was built up from the local weather at thousands of points. But the physics of local weather was also a formidable problem.

2. Numerical weather prediction (1922–1965)

Early in the 20th century, Vilhelm Bjerknes developed a set of seven “primitive equations” describing the behavior of heat, air motion, and moisture. The solution of the set of equations would, in principle, describe and predict large-scale atmospheric motions (Nebeker, 1995; Friedman, 1989). Lewis Fry Richardson (1922) published a numerical system for weather prediction, using simplified versions of Bjerknes's equations. Richardson's idea was to divide up a territory into a grid of cells, each with its own set of numbers describing its air pressure, temperature, and the like, as measured at a given hour. He would then solve the equations that told how air behaved. He could calculate wind speed and direction, for example, from the difference in pressure between two adjacent cells. The number of computations was so great that Richardson scarcely hoped his idea could lead to practical weather forecasting. A practical test, which cost him
six weeks of pencil-work, ended in failure. Taking the warning to heart, meteorologists gave up any hope of numerical modeling (Nebeker, 1995, chap. 6).

The alternative to the failed numerical approach was to keep trying to find a solution in terms of mathematical functions—a few pages of equations that an expert might comprehend as easily as a musician reads music. Through the 1950s, some leading meteorologists worked with simplified forms of the physics equations, trying a variety of approaches to what we might call a mathematical simulation of the atmosphere. They were never able to convincingly show the features of the general circulation, not even something as simple and important as the trade winds.

What was hopeless with pencil and paper might possibly be made to work with the new digital computers, feverishly developed during the Second World War and leaping ahead in power as the Cold War demanded ever more calculations. In the lead, energetically devising ways to simulate nuclear weapons explosions, was the Princeton mathematician John von Neumann. Von Neumann saw parallels between his explosion simulations and weather prediction (both are problems of non-linear fluid dynamics). He assembled a small group of theoretical meteorologists at Princeton's Institute for Advanced Study to look into regional weather prediction; if that proved feasible, the group planned to move on to the extremely ambitious problem of modeling the entire global atmosphere. Von Neumann invited Jule Charney, an energetic and visionary meteorologist, to head the new group. Charney came from Carl-Gustaf Rossby's pioneering meteorology department at the University of Chicago, where the study of weather maps and fluids had developed a toolkit of sophisticated mathematical techniques and an intuitive grasp of basic weather processes.

Richardson’s equations were the necessary starting-point, but Charney had to simplify them if he hoped to run large-scale calculations in weeks rather than centuries. He began with a set of simplified equations that described the flow of air along a narrow band of latitude. By 1949, his group had results that looked fairly realistic—sets of numbers that you could almost mistake for real weather diagrams, if you did not look too closely. In one characteristic experiment, they modeled the effects of a large mountain range on the air flow across a continent. Modeling was taking the first steps toward the computer games that would come a generation later, in which the player acts as a god: raise up a mountain range and see what happens! Soon the group proceeded to fully three-dimensional models for a region (Charney, 1949; Charney & Eliassen, 1949).

To Charney (1949, pp. 371–372), all this was just an extension of normal theoretical analysis. “By reducing the mathematical difficulties involved in carrying a train of physical thought to its logical conclusion,” he wrote, “the machine will give a greater scope to the making and testing of physical hypotheses”. He did not imagine that computer models would be able to convey insights in a way that could not come from physics theory, nor a laboratory setup, nor the data on a weather map, but in an altogether new way. His challenge remained what it had been in the traditional style of physics theory: to combine and simplify equations until you got formulas that gave sensible results with a limited amount of computation. Developing usable combinations and approximations of meteorological variables took a rare combination of mathematical ingenuity and physical insight. And that was only the beginning.

To know when you were getting close to a realistic model, you had to compare your results with the actual atmosphere. For that you would need an unprecedented number of measurements of temperature, moisture, wind speed, and so forth. Largely because of military needs, during the war and afterward networks had been established to send up thousands of balloons that radioed back measurements of the upper air. For the first time the atmosphere was seen not as a single layer, as represented by a surface map, but in its full three dimensions. By the 1950s, the weather over continental areas, up to the lower stratosphere, was being mapped well enough for comparison with results from rudimentary models.3

The first serious weather simulation Charney's team completed was two-dimensional. Their model divided the atmosphere into a grid of cells, covering North America with 270 points about 700 km apart. Starting with real weather data for a particular day, the computer solved all the equations for how the air should respond to the differences in conditions between each pair of adjacent cells. Taking the outcome as a new set of weather data, it stepped forward in time (using a step of three hours) and computed all the cells again. Charney, Fjörtoft, and von Neumann (1950, p. 245) remarked that between each run it took them so long to print and sort punched cards that “the calculation time for a 24-hour forecast was about 24 hours, that is, we were just able to keep pace with the weather”. The resulting forecasts were far from perfect, but they turned up enough features of what the weather had actually done on the chosen day to justify pushing forward (Platzman, 1979). The U.S. Weather Bureau and units of the armed forces established a Joint Numerical Weather Prediction Unit, which in May 1955 began issuing real-time forecasts in advance of the weather. The results were encouraging, although it would be well over a decade before the accuracy of computer forecasts began to reliably outstrip the subjective guesswork of experienced human forecasters (Cressman, 1996; Nebeker, 1995).

These early forecasting models were regional, not global in scale. It was not for practical reasons of weather prediction that a few meteorologists wanted to push on to model the entire general circulation of the global atmosphere. To be sure, Von Neumann and the military agencies that funded him speculated that understanding the climate might make it possible someday to manipulate it, as rain-makers claimed to do, perhaps even as a form of warfare. But for the foreseeable future, the scientists' true ambition was strictly theoretical, the hope of understanding at last how the climate system worked—why there are deserts in some places, wet zones in others, and so forth. That computation became a holy grail for theoretical meteorologists.

Norman Phillips in Princeton took up the challenge. He was encouraged by “dishpan” experiments carried out in Chicago: a pan of water, rotating on a turntable and heated at the edge (as the Earth is heated in the tropics)—a physical simulation of the atmosphere. The simulation produced something roughly like the convection cells of the atmosphere's general circulation and even tiny storms (Fultz, 1949, 1952). For Phillips (1956) this proved that “at least the gross features of the general circulation of the atmosphere can be predicted without having to specify the heating and cooling in great detail.” If such an elementary laboratory system could model a hemisphere of the atmosphere, should not a computer be able to do as well? To be sure, the computer at Phillips’s disposal was as primitive as the dishpan (its RAM held all of five kilobytes and its magnetic-drum memory held ten). So his model had to be extremely simple. By mid-1955 Phillips had developed improved equations for a two-layer atmosphere. To avoid mathematical complexities, his grid covered not a hemisphere but a cylinder, 17 cells high and 16 in circumference. He drove circulation by putting heat into the lower half, somewhat like the dishpan experimenters only with

---

2 The first experiment in a GCM (raising up the Himalayas) was Mintz (1965).

numbers rather than an electrical coil. The calculations turned out a plausible jet stream and the evolution of a realistic-looking weather disturbance over as long as a month.

This settled an old controversy over what processes built the pattern of circulation. For the first time scientists could see, among other things, how giant eddies spinning through the atmosphere played a key role in moving energy and momentum from place to place. Phillips’s model was quickly hailed as a “classic experiment”—the first true General Circulation Model, or “GCM” (Smagorinsky, 1963, p. 100).

Three-dimensional models were not the only way to approach problems of climate. Early models of the global atmosphere had compressed every process into one or two dimensions. The three-dimensional models were indeed made up of one-dimensional vertical columns, each exchanging air and energy with its neighbors; such columns continued to attract much research attention over subsequent decades. Meanwhile scientists would develop an ever richer variety of approaches, finding many ways to slice up the total number of arithmetic operations that a computer could run through in whatever time you could afford to pay for. This essay does not cover these developments, but concentrates on the three-dimensional GCMs, the countless descendants of Phillips’s model.

Encouraged by the model’s success, Von Neumann drummed up government funding for a long-term project. The effort got underway that same year, 1955, under the direction of Joseph Smagorinsky at the Weather Bureau near Washington, DC. Smagorinsky’s goal was the one first envisaged by von Neumann and Charney, a general circulation model of the entire global atmosphere built directly from the primitive equations. In 1958, he invited the young Japanese meteorologist Syukuro Manabe to join the lab. Smagorinsky and Manabe put into their model the rainfall fell on the surface and evaporated, they put in how radiation passing through the atmosphere interacted not only with water vapor but also with ozone and carbon dioxide gas (CO2), they put in how the air exchanged water and heat with simplified ocean, land, and ice surfaces, and much more. Manabe spent many hours in the library studying such esoteric topics as how various types of soil absorbed water. The huge complexities of the modeling required contributions from several others. “This venture has demonstrated to me,” said Smagorinsky (1963, p. 151), “the value if not the necessity of a diverse, imaginative, and dedicated working group in large research undertakings”.

As decades passed this necessity would drive the community of researchers to grow by orders of magnitude without ceasing to collaborate closely.

By 1965 the group had a reasonably complete three-dimensional global model that solved the basic equations for an atmosphere divided into nine levels. This was still highly simplified, with no geography—land and ocean were blended into a single damp surface, which exchanged moisture with the air but could not take up heat. Nevertheless, the way the model moved water vapor around the planet looked gratifyingly realistic. The printouts showed a stratosphere, a zone of rising air near the equator (creating the doldrums, a windless zone that becalmed sailors), a subtropical band of deserts, and so forth. Manabe’s model (Arakawa, Venkateswaran, & Wurtele, 1994). By 1964 Mintz and Arakawa had produced a climate computed for an entire globe, with only a two-layer atmosphere but including realistic geography—the topography of mountain ranges was there, and a rudimentary treatment of oceans and ice cover. Although the results missed some features of the real world’s climate, the basic wind patterns and other features came out more or less right (Arakawa, 1970; Mintz, 1965). The model, packed with useful techniques, had a powerful influence on other groups.

Arakawa was becoming especially interested in a problem that was emerging as a main barrier to progress—accounting for the effects of cumulus clouds. The smallest single cell in a global model that a computer can handle, even today, is far larger than an individual cloud. Thus the computer calculates none of the cloud’s details. Models have had to get by with a “parameterization,” a set of numbers (parameters) that represent the net behavior of all the clouds in a cell under given conditions. Through the decades, Arakawa and others would spend countless hours developing and exchanging ways to attack this critical problem.

Modeling techniques and entire GCMs spread by a variety of means. In the early days, as Phillips recalled, modelers had been like “a secret code society.” There were so many subtleties that a real grasp required an apprenticeship on a working model. Commonly, a new modeling group began with some version of another group’s model. A post-doctoral student might take a job at another institution, bringing along his old team’s computer code. The new team he assembled would start off working with the old code and then set to modifying it. Others built new models from scratch. Americans dominated the field during the first postwar decades, but through the 1960s and 1970s, important GCM groups also emerged at institutions from England to Australia.

3. Credible climate prediction (1965–1979)

Although the modelers of the 1950s and early 1960s got results good enough to encourage them to persevere, they were still a long way from reproducing the details of the Earth’s actual climate zones. Even if the computers had been vastly faster, the simulations would still have been unreliable. For they were running up against that famous limitation of computers, “garbage in, garbage out.” To diagnose the failings that kept GCMs from being more realistic, scientists needed an intensified effort to collect and analyze data showing the actual profiles of wind, heat, moisture, and so forth, at every level of the atmosphere and all around the globe. The data in hand were still deeply insufficient. For example, for the atmosphere’s crucial water and energy balances, the expert Edward Lorenz (1967) estimated that the commonly used numbers might be off by as much as 50%. Smagorinsky (1970, p. 33) put the problem succinctly: “We are now getting to the point where the dispersion of simulation results is comparable to the uncertainty of establishing the actual atmospheric structure”.

4 Arakawa & Schubert (1974) was a major step, and briefly reviews the history.


6 For a “family tree” of models, see Edwards (2010, Fig. 7.4, p. 168).
Lorenz (1967, p. 10) cautioned that “even the trade winds and the prevailing westerlies at sea level are not completely explained”. He and a few others began to suspect that the problem was not merely difficult, but impossible in principle. They worried that climate was not a well-defined system, but only an average of the ever-changing jumble of daily storm fronts. Would computer modelers ever be able to say they had “explained” the general circulation? Many scientists looked askance at the new method of numerical simulation. People were attacking many kinds of scientific problems by taking a set of basic equations and running them through hundreds of thousands of computations. Their results were simply stacks of printout with rows of numbers. That was no “explanation” in the traditional sense of a model in words or diagrams or equations, something you could write down on a few pages, something your brain could grasp intuitively as a whole. The numerical approach “yields little insight,” Lorenz (1967, p. 8) complained. “The computed numbers are not only processed like data but they look like data, and a study of them may be no more enlightening than a study of real meteorological observations”.

Yet the computer scientist could “experiment” in a sense, by varying the parameters and features of a numerical model. You couldn’t put a planet on a laboratory bench and vary the sunlight or the way clouds were formed, but wasn’t playing with computer models functionally equivalent? Through many such trials you might eventually come to understand how the real world operated. Indeed you might be able to observe the planet more clearly in graphs printed out from a model than in the clutter of real-world observations, so woefully inaccurate and incomplete (Kellogg & Schneider, 1974, p. 1166).

A new viewpoint was spreading along with digital computing. Climate was not changed by any single cause, the modelers said. It was the outcome of a staggeringly intricate complex of interactions—something that could only be comprehended in the working-through of the numbers. The printouts were all the explanation there was.

The growing community of climate modelers was strengthened by the progress of other teams that carried out detailed calculations on short time-scales for weather prediction. This progress required much work on parameterization—schemes for representing cloud formation, interactions between waves and winds, and so forth. Such studies accelerated as the 1970s began. The weather forecasting models also required data on conditions at every level of the atmosphere at thousands of points around the world. Such observations were now being provided by the balloons and sounding rockets of an international World Weather Watch, founded in the mid-1960s. The weather predictions became accurate enough—looking as far as three days ahead—to be economically important. That built support for the meteorological measurement networks and computer studies necessary for climate work.

An example of the crossover could be found at NASA’s Goddard Institute for Space Studies in New York City. A group there under James Hansen had been developing a weather model as a practical application of its mission to study the atmospheres of planets. For one basic component of this model, Hansen developed a set of equations for the transfer of radiation through the atmosphere, based on work he had originally done for studies of the planet Venus. The same equations could be used for a climate model, by combining them with an elegant method for computing fluid dynamics that Arakawa had developed.

In the 1970s, Hansen assembled a team to work up a model that would be both fast-running and realistic. An example of the kind of detail they pursued was a simple equation they devised to represent the reflection of sunlight from snow. They included the age of the snow layer (as it gradually melted away) and the “masking” by vegetation (snowy forests are darker than snowy tundra). They managed to get a quite realistic-looking climate that ran an order of magnitude faster than some rival GCMs, permitting the group to experiment with multiple runs, varying one factor or another to see what changed. In such studies, the global climate was beginning to feel to researchers like a comprehensible physical system, akin to the systems of glassware and chemicals that experimental scientists manipulated on their laboratory benches (Hansen et al., 1983).

Groups continued to proliferate, borrowing ideas from earlier models and devising new techniques of their own. In their first decade or so of work the GCM modelers had treated climate as a given, a static condition. But in the 1960s a few modelers began to take an interest in global climate change as a problem over the long term. The observations by Keeling (1960, 1970) that the level of CO₂ in the atmosphere was rising prompted hard thinking about greenhouse warming.

Manabe had a long-standing interest in the effects of CO₂, not because he was worried about the future climate, but simply because the gas at its current level was a significant factor in the planet’s heat balance. First, however, he had to deal with the most powerful greenhouse gas: H₂O. Manabe and his colleagues were building the first model that took full account of water in all its forms, realistically including the feedback between the air’s temperature and the amount of water vapor the air would hold. In particular, Manabe’s group calculated the way rising columns of moisture-laden air conveyed heat from the surface into the upper atmosphere. The required computations were so extensive that Manabe stripped down the model to a single one-dimensional column, which represented the atmosphere averaged over the globe. His aim was to get a system that could be used as a basic building-block for a full three-dimensional GCM.

In 1967, Manabe and a collaborator, Richard Wetherald, used the one-dimensional model to test what would happen if the level of CO₂ changed. Their target was the climate’s “sensitivity,” often defined as the change in average global temperature if the CO₂ level doubled. Their answer was that global temperature would rise roughly 2 °C (Manabe & Wetherald, 1967).

This was no more than a first baby step toward a realistic three-dimensional model of the changing climate. The next important step was also taken by Manabe’s group, which in 1968 moved to Princeton. They now used a GCM with nine atmospheric levels, but it was still highly simplified. In place of actual land and ocean geography they pictured a geometrically neat planet, half damp surface (land) and half wet (a “swamp” ocean). They could not predict cloudiness but just held it unchanged at the present level when they calculated the warmer ocean. However, they did incorporate the movements of water, predicting changes in soil moisture and snow cover on land, and they calculated sea surface temperatures well enough to show the extent of sea ice. The results, published in 1975, looked quite realistic overall. As one might expect on basic physical grounds, the warmer model with increased CO₂ held more moisture in the air, with an intensified hydrological cycle of evaporation and precipitation, as hot soil dried out in one region and more rain came down elsewhere. The model also showed greater warming in the Arctic than in the tropics; this too could be predicted from simple reasoning, for less snow and ice meant more absorption of sunlight by ground and sea. But it took a calculation to show that what sounded reasonable on elementary principles was indeed likely to happen in the real world—or at least this simulation of it.
Averaged over the entire planet, for doubled CO2 the computer predicted a warming of around 3.5 °C. It all looked plausible. The results made a considerable impact on scientists, and through them on policy-makers and the public.

**Manabe and Wetherald (1975, p. 13)** warned that “it is not advisable to take too seriously” the specific numbers they published. They singled out the way the model treated the oceans as a simple wet surface. On our actual planet, the oceans absorb large quantities of heat from the atmosphere, move it around, and release it elsewhere. And there remained the old vexing problem of clouds. As the planet got warmer the amounts of cloudiness would probably change at each level of the atmosphere in each zone of latitude, but change how? There was no reliable way to figure that out. Worse, the net effect depended on the types of cloud and how high they floated in the atmosphere. A better prediction of climate change would have to wait on general improvements.

Progress was steady, thanks to the headlong advance of electronic computers. From the mid-1950s to the mid-1970s, the power available to modelers increased by a factor of thousands. That meant modelers could put in more factors in more complex ways. The models no longer had gaping holes that required major innovations, and the work settled into a steady improvement of existing techniques. At the foundations, modelers devised increasingly more sophisticated and efficient schemes of computation. As input for the computations they worked endlessly to improve parameterizations. From around 1970 on, many journal articles appeared with ideas for dealing with convection, evaporation of moisture, reflection from ice, and so forth (Nebeker, 1989).

The most essential element for progress, however, was better data on the real world. Strong efforts were rapidly extending the observing systems. In 1969 NASA's Nimbus 3 satellite began to broadcast measurements which, although designed primarily for meteorology, provided a fundamental check on model results. A main reason Manabe's 1975 model planet was seen as successful was the reasonably good agreement between its reflection of sunlight at each latitude and the actual numbers measured by Nimbus 3.

Also encouraging was a 1972 model by Mintz and Arakawa (unpublished, like much of their work), which managed to simulate in a rough way the huge changes as the sunlight shifted from season to season. During the next few years, Manabe and collaborators published a model that produced entirely plausible seasonal variations. Seasons provided a convincing test of the model's validity. It was almost as if a single model worked for two seasons. During the next few years, Manabe and the other by Hansen—elaborate three-dimensional models that used different physical approaches and different computational methods for many features. The panel found differences in detail but solid agreement for the main point: the world would get warmer as CO2 levels rose. The Charney panel announced that they had rather high confidence that as CO2 doubled in the atmosphere, the planet would warm up by about three degrees, plus or minus fifty percent: in other words, 1.5–4.5 °C (2.7–8 °F). They concluded dryly, “We have tried but have been unable to find any overlooked or underestimated physical effects” that could reduce the warming (National Academy of Sciences, 1979, pp. 2, 3).


Several groups pressed ahead toward more realistic models. By the early 1980s they were using a reasonable facsimile of the Earth’s actual geography, and replaced the wet “swamp” surface with an ocean that could exchange heat with the atmosphere. They were coming to believe that their work might be of more than academic interest. When they introduced a doubled CO2 level into their improved models, they consistently found the same few degrees of warming (Manabe & Stouffer, 1980).

Many scientists, especially ones outside the modeling fraternity, were skeptical. The treatment of clouds remained a central uncertainty. Who could say that as the world warmed, clouds would not spread, reflecting sunlight to prevent further warming? Another great unknown was the response of the oceans. Oceanographers were coming to realize that large amounts of energy were carried through the seas by a myriad of whirls of various types, from tiny convection swirls up to sluggish eddies a thousand kilometers wide. Calculating these whirls, like calculating all the world’s individual clouds, was beyond the reach of the fastest computer. Again parameters had to be devised to summarize the main effects, only this time for entities that were far worse observed and understood than clouds.

Manabe was keenly aware that if the Earth’s future climate were ever to be predicted, it was “essential to construct a realistic model of the joint ocean–atmosphere system” (Manabe, Bryan, & Spelman, 1979, p. 394). He shouldered the task in collaboration with Kirk Bryan, an oceanographer with meteorological training, who had been brought into the group back in 1961 to build a stand-alone numerical model of ocean circulation. Manabe’s winds and rain would help drive Bryan’s ocean currents, while in return Bryan’s sea-surface temperatures and evaporation would help drive the circulation of Manabe’s atmosphere. Bryan and Manabe not only incorporated both oceans and atmosphere, but added into the bargain feedbacks from changes in sea ice. Moreover, they included a detailed scheme that represented, region by region, how moisture built up in the soil, or evaporated, 

---

8 Manabe had a rough seasonal simulation by 1970 and published a full seasonal variation in Manabe, Hahn, & Holloway (1974).
or ran off in rivers to the sea. They had enough confidence in their model to undertake a heroic computer run, some 1100 hours long (more than 12 full days of computer time devoted to the atmosphere and 33 to the ocean). They published the results in an unusually short paper (Manabe & Bryan, 1969).

Bryan (1969, p. 822) wrote modestly that “in one sense the experiment is a failure”. For even after a simulated century, the deep ocean circulation had not nearly reached equilibrium. It was not clear what the final climate solution would look like. Yet it was a great success just to carry through a linked ocean–atmosphere computation that was at least starting to settle into equilibrium. The result looked like a real planet—not our Earth, for in place of geography there was only a radically simplified geometrical sketch, but it had ocean currents, trade winds, deserts, rain belts, and snow cover, all in roughly the right places. Unlike our actual Earth, so poorly observed, in the simulation one could see every detail of how air, water, and energy moved about.

Following up, in 1975 Manabe and Bryan published results from the first coupled ocean–atmosphere GCM that had a roughly Earth-like geography. For example, it showed the Sahara and the American Southwest as deserts, but plenty of rain in the Pacific Northwest and Brazil. Manabe and Bryan had not shaped their equations deliberately to bring forth such features. These were “emergent features,” emerging spontaneously out of the computations. The computer’s output looked roughly like the actual climate only because the modelers had succeeded in roughly representing the actual operations of the atmosphere upon the Earth’s geography. By 1979, Manabe and Bryan had mobilized enough computer power to run their model through more than a millennium while incorporating seasons. (Bryan, Manabe, & Pacanowski, 1975; Manabe, Bryan, & Spelman, 1975; Manabe et al., 1979).

Meanwhile a team headed by Warren Washington at the National Center for Atmospheric Research (NCAR) in Boulder, Colorado developed another ocean model, based on Bryan’s, and coupled it to their own quite different GCM. It was gratifying that again the patterns of air temperature, ocean salinity and so forth came out roughly correct, albeit with noticeable deviations from the real planet, such as tropics that were too cold (Washington, Semtner, Meehl, Knight, & Mayer, 1980). Through the 1980s, these and other teams continued to refine coupled models, occasionally checking how they reacted to increased levels of CO₂. These were not so much attempts to predict the real climate as experiments to work out methods for doing so.

The results, for all their limitations, said something about the predictions of the atmosphere-only GCMs. Including a somewhat realistic ocean did not turn up anything that would alter the basic prediction of future warming. Evidently simple models were good enough to point in the right direction. On the other hand, even very simple reasoning showed that the models were still inadequate for good predictions. In particular, nobody had actually simulated a climate change. Manabe and others had run their model twice to compute two equilibrium states, one with current conditions and one with doubled CO₂. In the real world, the atmosphere would pass through a series of changes as the level of the gas rose. There were hints that a model could end up in different states depending on just what route it took.

Hansen’s group and a few others therefore undertook protracted computer runs to compute the entire “transient response” while the CO₂ level rose. They plodded through a century or more simulating from one day to the next. Hansen’s coupled ocean–atmosphere model, which incorporated not only the observed rise of CO₂ but also the historical record of aerosols from volcanic explosions, turned out a fair approximation to the observed global temperature trend of the previous half century. Pushed into the future, the model showed sustained global warming (Hansen et al., 1988). By 1988 Hansen had enough confidence to issue a strong public pronouncement, warning of an imminent threat. Meanwhile a major international scientific conference in Toronto issued a call for the world’s governments to restrict greenhouse gas emissions. The predictions of computer models had become a matter of concern to policymakers—and politicians.

The climate changes computed for doubled CO₂ by different GCMs, noted reviewers Schlesinger and Mitchell (1987, p. 795), “show many quantitative and even qualitative differences; thus we know that not all of these simulations can be correct, and perhaps all may be wrong”. Skeptics pointed out that GCMs did not really represent even the present climate successfully from first principles. Anything slightly unrealistic in the initial data or equations could be amplified a little at each step, and after thousands of steps the entire result usually veered off into something impossible. To get around this, the modelers had kept one eye over their shoulder at the real world. They adjusted various parameters, for example, numbers describing cloud physics, running the models again and again, “tuning” them until the output resembled the real climate. Some of these adjustments were not calculated from physical principles, nor were they pinned down precisely by field studies. They were fiddled, within the limits set by the known physics and data, until the model became realistic and stable. But if models were tuned to match the current climate, the critics asked, how reliably could they calculate a future, different state?

 Debates over climate models also helped stimulate philosophers of science, who explained that a computer model, like any other embodiment of a set of scientific hypotheses, could never be “proved in the absolute sense one could prove a mathematical theorem. What models could do was help people sort through countless ideas and possibilities, offering evidence on which were most plausible. Eventually the models, along with other evidence and other lines of reasoning, might converge on a representation of climate that—if necessarily imperfect, like all human knowledge—could be highly reliable (Norton and Suppe, 2001; Oreskes, Shrader-Frechette, & Belitz, 1994).

Undeniably there remained points where the models stood on shaky foundations. For some features, no calculation could be trusted until more observations were made. Even the actual cloudiness of various regions of the world had been measured in only a sketchy fashion. (Until satellite measurements became available later in the decade, most models used data from the 1950s that only gave averages by zones of latitude, and only for the Northern Hemisphere). It did not solve a problem if you happened to understand all the physics. As Manabe regretfully explained, so much physics was involved in every raindrop that it would never be possible to compute absolutely everything. “And even if you have a perfect model which mimics the climate system, you don’t know it, and you have no way of proving it.”

Reliable progress would require more work on fundamental elements, to improve the sub-models that represented clouds, snow, vegetation, and so forth. Modelers settled into a long grind of piecemeal improvements.

5. Results for policymakers (1988–2007)

“There has been little change over the last 20 years or so in the approaches of the various modeling groups,” an observer remarked in 1989. Modelers felt they had a basic grasp of the

---

9 Manabe, interview by Weart, December 1989 (AIP).
main forces and variations in the atmosphere. Their interest was shifting from representing the current climate ever more precisely to studies of long-term climate change. The research front accordingly moved from calculating stable systems to representing the transient response to changes in conditions. Running models under different conditions, sometimes through simulated centuries, with rising confidence the teams drew rough sketches of how climate could be altered by changes in greenhouse gases. Many were now reasonably sure that they knew enough to issue stern warnings of future global warming.

As GCMs incorporated ever more complexities, modelers needed to work ever more closely with one another and with people in outside specialties. The clearest case centered around NCAR. It lived up to its name of a “National Center” (in fact an international center) by developing what was explicitly a “Community Climate Model.” In 1983 NCAR published all its computer source codes along with a “Users’ Guide” so that outside groups could run the model on their own machines. The various outside experiments and modifications in return informed the NCAR group. Subsequent versions incorporated many basic changes and additional features. NCAR had an exceptionally strong institutional commitment to building a model that could be run on a variety of computer platforms, but in other ways their work was not unusual. By now most models used contributions from so many different sources that they were all in a sense “community” models (Kiehl, Hack, & Bonan, 1996, pp. 1–2).

An important example of massive collaboration was a 1989 study involving groups in the United States, Canada, England, France, Germany, China, and Japan. Taking 14 models of varying complexity, the groups fed each model the same external forces and compared the results. The simulated climates agreed well for clear skies. But “when cloud feedback was included, compatibility vanished.” The models varied by as much as a factor of three in their sensitivity to the external forces, disagreeing in particular on how far a given increase of CO₂ would raise the temperature (Cess et al., 1989, 1990). A few respected meteorologists concluded that the modelers’ representation of clouds was altogether useless. Most experts nevertheless felt the GCMs were on the right track. In the multi-model comparisons, all the results were at least in rough overall agreement with reality.

These studies were helped greatly by a new capability to set their results against a uniform body of world-wide data. Specially designed satellite instruments were at last monitoring incoming and outgoing radiation, cloud cover, and other essential parameters. No less important, the sketchy parameterizations in the models were increasingly refined by field studies. Decade by decade the science community mounted ever larger fleets of ships, aircraft and balloons to observe actual processes in clouds and other key features of the climate system. The instrumental systems were increasingly oriented toward producing numbers meaningful to the models, and vice-versa; global data and global models were no longer distinct entities, but parts of a single system for representing the world (Edwards, 1999, 2010).

There was also progress in building aerosols into climate models. When Mount Pinatubo erupted in the Philippines in June 1991, sharply increasing the amount of sulfuric acid haze in the stratosphere world-wide, Hansen’s group declared that “this volcano will provide an acid test for global climate models.” Running their model with the new data, they predicted a “striking” (Carson, 1999, p. 10). The ability of modelers to not only reproduce but predict Pinatubo’s effects gave scientists a particularly strong reason for believing that adjusting the parameters within the GCMs had created some kind of reliable connection with reality, the actual planet.

Incorporating aerosols into GCMs helped to answer a major criticism. Skeptics had pointed out that the actual rise in temperature over the past century had not been as large as would be expected for the known rise of CO₂ in the atmosphere, given the 3 °C sensitivity that typical models calculated. Try as they might, the modelers had not been able to tune their GCMs to get the modest temperature rise observed. An answer came when advanced models put in the history of a global rise of aerosol pollution. The aerosols had increasingly exerted a cooling effect, which had tended to offset part of the greenhouse warming. In 1995, models at centers in the US, England and Germany all reproduced fairly well the overall trend of 20th-century temperature changes and even the observed geographical patterns (Mitchell, Johns, Gregory, & Tett, 1995). This reversed the significance of the earlier inability of models to reproduce the temperature trend. Apparently the models that had been tuned without aerosols had correctly represented a planet without aerosols; they had been grounded solidly enough in reality to resist attempts to force them to give a false answer.

The success in matching the historic temperature trend powerfully influenced the Intergovernmental Panel on Climate Change (IPCC), a collaboration of experts and officials appointed by the world’s governments. In its 1996 reports the IPCC noted that the pattern of geographical and vertical distribution of atmospheric heating that the models computed for greenhouse warming was different from the pattern that would result from other influences, for example changes in solar activity. The rough similarity between the computed greenhouse effect’s “signature” and the actual record of recent decades backed up the panel’s official conclusion: “The balance of evidence suggests that there is a discernible human influence on global climate” (IPCC, 1996, p. 22, see chap. 8).

The weaselly wording reflected the fact that many people found reasons for skepticism. In particular, the models that coupled atmospheric circulation to a full-scale ocean model, a type of model that now dominated attention, all tended to drift into unrealistic patterns. The only solution was to tune the models to match real-world conditions by adjusting various parameters. The simplest method was to fiddle with the flux of heat at the interface between ocean and atmosphere in regions where the models gave bad results. Modelers would likewise force transfers of water and so forth, formally violating basic laws of physics to compensate for their models’ deficiencies. The little community of modelers was divided, with some roundly criticizing flux adjustments as “fudge factors” that could bring whatever results a modeler sought (Dahan-Dalmedico, 2007, p. 142; Shackley, Risbey, Stone, & Wynne, 1999). If the models were arbitrarily forced to match the present climate, why believe they could tell us anything at all about a different situation? But by 1999 two computer groups managed to do away with flux adjustments while running their models through centuries. Their results were not severely different from the results of the earlier flux-adjusted models. Evidently the tuning had not been a fatal cheat. Models without flux adjustments soon became standard (Carson, 1999, pp. 13–17; Kerr, 1997).

Further raising confidence, in 2001 two groups using coupled models matched the rise of temperature that had been detected in the upper layers of the world’s oceans. They got a good match only by putting in the rise of greenhouse gases. By 2005, computer

---

11 The “signature” or “fingerprint” method was pioneered by Klaus Hasselmann’s group at the Max Planck Institute, e.g., Hasselmann (1993). See also Santer et al. (1996).
modelers had advanced far enough to declare that temperature measurements in every ocean basin over the previous four decades gave a detailed, unequivocal “signature” of the greenhouse effect. Nothing but the rise of greenhouse gases could produce such a warming pattern, not the observed changes in the Sun’s radiation, emissions from volcanoes, or any other proposed “natural” mechanism (Barnett et al., 2005).

Yet if modelers now understood how the climate system could change and even how it had changed, they were far from saying precisely how it would change in future. Never mind how the mean global temperature would rise; citizens and policy-makers wanted to know what heat waves, droughts or floods were likely in their particular region. The attention of the community turned to making predictions in ever more geographical detail.

Such detail was unattainable without further refinements. For example, ocean–atmosphere GCMs had to be linked interactively with models for changes in vegetation. Dark forests and bright deserts not only responded to climate, but influenced it. Since the early 1990s the more advanced numerical models, for weather prediction as well as climate, had incorporated descriptions of such things as the way plants take up water through their roots and evaporate it into the atmosphere. Changes that pollution caused in the chemistry of the atmosphere also had to be incorporated, for these influenced cloud formation and more. Over longer time scales, modelers would also need to consider changes in ocean chemistry, ice sheets, ecosystems, and so forth. When people talked now of a “GCM” they no longer meant a “General Circulation Model,” built from the traditional equations for weather. “GCM” now stood for “Global Climate Model” or even “Global Coupled Model,” incorporating many things besides the circulation of the atmosphere. These simulations strained the resources of the newest and biggest supercomputers, some of which were built with climate modeling primarily in mind.

The range of modelers’ predictions of global warming for a doubling of CO₂ was still around 3 °C give or take 50%. If the real sensitivity was at the lower limit of what seemed reasonably possible, the world would have ample time to adjust its economy to the gathering threat; if at the upper limit, only prompt and vigorous efforts could ward off catastrophe. The ineradicable challenge to them was to produce a simulation that did not reasonably dismiss computer modeling in general. That would be saying something as the truth when it emerges ‘as the intersection of independent lies’. Although ‘lies’ may sound a bit too harsh for the models involved, both our approach and the large simulation models clearly have their shortcomings. Interpreting our results in this spirit, they enhance the credibility of global warming projections (Scheffer, Brovkin, & Cox, 2006, p. L10706, slightly misquoting Levins, 1966, p. 423).

For a climate pretty much like the present, however, all the significant mechanisms must have gotten incorporated somehow into the parameters. At worst, the models were all getting right results for wrong reasons—flaws that would only show up after greenhouse gases pushed the climate beyond any conditions that the models were designed to reproduce. If there were such deep-set flaws, that did not mean, as some skeptics implied, that there was no need to worry about global warming. If the models were faulty, the future climate changes could be worse than they predicted, not better. For all the millions of hours the modelers had devoted to their computations, in the end they could not say exactly how serious future global warming would be. They could only say that it was very likely to be bad, and it just might be an appalling catastrophe.

Those who still denied any serious risk of climate change could not reasonably dismiss computer modeling in general. That would throw away much of the past few decades’ work in many fields of science and engineering, and even key business practices. The challenge to them was to produce a simulation that did not show global warming. Now that personal computers were far more powerful than the most expensive computers of earlier decades, it was possible to explore thousands of combinations of parameters. But no matter how people fiddled with climate models, the answer was the same. If your model could reproduce something resembling the present climate, and then you added some greenhouse gases, the model showed serious global warming (Stainforth et al., 2005).

By the dawn of the 21st century, climate models had become a crucial source of information for policy-makers and the public. Struggling to provide better predictions, the community of modelers grew still larger and better organized. Projects to compare the models devised by different groups became a major ongoing activity. By 2007, groups were exchanging so much data that it would have taken years to transfer it on the internet, and they took to shipping it on terabyte hard drives. There were about a dozen major teams now and a dozen more that could make significant contributions. The pictures of overall global warming that their GCMs computed were converging. It was largely thanks to their work that, as the editor of Science magazine announced, a “consensus as strong as the one that has developed around this topic is rare in the history of science” (Kennedy, 2001).

Each computer modeling group normally worked in a cycle. When their model began to look outdated, and still more if they managed to acquire a new supercomputer, they would go back to basics and spend a few years developing a new model. It was no simple process, for introducing a new wrinkle (for example, a new way to calculate convection in the tropics) might introduce unexpected feedbacks that caused the entire model to crash. Once a team had persuaded their model to produce stable results that looked like the real world, they would spend the next year or two using it to analyze climate processes, gathering ideas for the next cycle.

After finishing their part of the IPCC’s 2001 report, the modeling community worked to synchronize the teams’ separate cycles. By early 2004, nearly all the major models simultaneously reached the analysis stage. That made it possible for 17 teams to share and compare data in time to produce results for the 2007 IPCC report (Schmidt, 2008).

The work was grueling. After a group had invested so much of their time, energy, and careers in their model, they could become reluctant to admit its shortcomings to outsiders and perhaps even to themselves. A frequent result was “prolonged and acrimonious conflicts” over whose approach was best (Lahsen, 2005, p. 916). Yet in the end they found common ground, working out a range of numbers that all agreed were plausible.

The numbers got an independent check. Other scientists were studying various kinds of evidence of past climates back to the ice ages, indeed back to the Cretaceous era and earlier still, lining up ancient temperatures with greenhouse gas levels and other influences such as volcanic aerosols. By 2006 they had arrived at fairly consistent numbers for how the mean global temperature connected with the level of CO₂. The numbers agreed comfortably with the IPCC computer modelers’ sensitivity consensus of 3 °C, give or take a couple of degrees (Annan & Hargreaves, 2006; Hegerl, Crowley, Hyde, & Frame, 2006). As one group that worked with ancient climates explained, “one is more likely to accept something as the truth when it emerges ‘as the intersection of independent lies’. Although ‘lies’ may sound a bit too harsh for the models involved, both our approach and the large simulation models clearly have their shortcomings. Interpreting our results in this spirit, they enhance the credibility” of global warming projections (Scheffer, Brovkin, & Cox, 2006, p. L10706, slightly misquoting Levins, 1966, p. 423).
As the charts of model outputs increasingly came to look like actual observed data, it became easy to forget which was which. As one modeler put it, “You start referring to your simulated ocean as ‘the ocean’—you know, ‘the ocean gets warm’... there is a tendency to forget that just because your model says x, y, or z doesn’t mean that that’s going to happen in the real world.” (Lahsen, 2005, p. 909) But it was small wonder if modelers were giving the products of their labor a new kind of ontological status. The modelers had long since reproduced not only simple geographical and seasonal averages from December to July and back, but also the spectrum of random regional and annual fluctuations in the averages—indeed it was now a test of a good model that a series of runs showed a variability similar to the real weather. Modelers had matched the record of climate changes for the past century. Exploring unusual conditions, they had reproduced the effects of volcanic eruptions and the ice ages. Above all, since the 1980s modelers had been predicting that a rise of mean global temperature would emerge from the “noise” of random climate shifts around the start of the 21st century; the prediction was fulfilled by a rise ominously faster and higher than anything in the historical record.

The models unanimously predicted much greater warming to come. In IPCC meetings virtually all the world’s climate experts and representatives of nearly all the world’s governments hammered out a consensus. Our civilization would probably face serious trouble, they announced unanimously in 2001—and again, more emphatically, in 2007—if we did not promptly begin to reduce our emissions of greenhouse gases. It was a triumph of rationality utterly without historical precedent: a drive to severely revise the entire world economy, relying chiefly on simulations whose workings only a few hundred experts actually understood.

References


Arakawa, A., Denk, S. L., & Wurtele, M. G. (1994). Yale Mintz, Arakawa, S., & Schubert, W. H. (1974). Interaction of a cumulus cloud ensemble with back, but also the spectrum of random regional and annual fluctuations in the averages—indeed it was now a test of a good model that a series of runs showed a variability similar to the real weather. Modelers had matched the record of climate changes for the past century. Exploring unusual conditions, they had reproduced the effects of volcanic eruptions and the ice ages. Above all, since the 1980s modelers had been predicting that a rise of mean global temperature would emerge from the “noise” of random climate shifts around the start of the 21st century; the prediction was fulfilled by a rise ominously faster and higher than anything in the historical record.

The models unanimously predicted much greater warming to come. In IPCC meetings virtually all the world’s climate experts and representatives of nearly all the world’s governments hammered out a consensus. Our civilization would probably face serious trouble, they announced unanimously in 2001—and again, more emphatically, in 2007—if we did not promptly begin to reduce our emissions of greenhouse gases. It was a triumph of rationality utterly without historical precedent; a drive to severely revise the entire world economy, relying chiefly on simulations whose workings only a few hundred experts actually understood.

References


As the charts of model outputs increasingly came to look like actual observed data, it became easy to forget which was which. As one modeler put it, “You start referring to your simulated ocean as ‘the ocean’—you know, ‘the ocean gets warm’... there is a tendency to forget that just because your model says x, y, or z doesn’t mean that that’s going to happen in the real world.” (Lahsen, 2005, p. 909) But it was small wonder if modelers were giving the products of their labor a new kind of ontological status. The modelers had long since reproduced not only simple geographical and seasonal averages from December to July and back, but also the spectrum of random regional and annual fluctuations in the averages—indeed it was now a test of a good model that a series of runs showed a variability similar to the real weather. Modelers had matched the record of climate changes for the past century. Exploring unusual conditions, they had reproduced the effects of volcanic eruptions and the ice ages. Above all, since the 1980s modelers had been predicting that a rise of mean global temperature would emerge from the “noise” of random climate shifts around the start of the 21st century; the prediction was fulfilled by a rise ominously faster and higher than anything in the historical record.

The models unanimously predicted much greater warming to come. In IPCC meetings virtually all the world’s climate experts and representatives of nearly all the world’s governments hammered out a consensus. Our civilization would probably face serious trouble, they announced unanimously in 2001—and again, more emphatically, in 2007—if we did not promptly begin to reduce our emissions of greenhouse gases. It was a triumph of rationality utterly without historical precedent; a drive to severely revise the entire world economy, relying chiefly on simulations whose workings only a few hundred experts actually understood.

References


Arakawa, A., Venkateswaran, S. V., & Wurtele, M. G. (1994). Yale Mintz, Arakawa, S., & Schubert, W. H. (1974). Interaction of a cumulus cloud ensemble with back, but also the spectrum of random regional and annual fluctuations in the averages—indeed it was now a test of a good model that a series of runs showed a variability similar to the real weather. Modelers had matched the record of climate changes for the past century. Exploring unusual conditions, they had reproduced the effects of volcanic eruptions and the ice ages. Above all, since the 1980s modelers had been predicting that a rise of mean global temperature would emerge from the “noise” of random climate shifts around the start of the 21st century; the prediction was fulfilled by a rise ominously faster and higher than anything in the historical record.

The models unanimously predicted much greater warming to come. In IPCC meetings virtually all the world’s climate experts and representatives of nearly all the world’s governments hammered out a consensus. Our civilization would probably face serious trouble, they announced unanimously in 2001—and again, more emphatically, in 2007—if we did not promptly begin to reduce our emissions of greenhouse gases. It was a triumph of rationality utterly without historical precedent; a drive to severely revise the entire world economy, relying chiefly on simulations whose workings only a few hundred experts actually understood.

References


Arakawa, A., Venkateswaran, S. V., & Wurtele, M. G. (1994). Yale Mintz, Arakawa, S., & Schubert, W. H. (1974). Interaction of a cumulus cloud ensemble with back, but also the spectrum of random regional and annual fluctuations in the averages—indeed it was now a test of a good model that a series of runs showed a variability similar to the real weather. Modelers had matched the record of climate changes for the past century. Exploring unusual conditions, they had reproduced the effects of volcanic eruptions and the ice ages. Above all, since the 1980s modelers had been predicting that a rise of mean global temperature would emerge from the “noise” of random climate shifts around the start of the 21st century; the prediction was fulfilled by a rise ominously faster and higher than anything in the historical record.

The models unanimously predicted much greater warming to come. In IPCC meetings virtually all the world’s climate experts and representatives of nearly all the world’s governments hammered out a consensus. Our civilization would probably face serious trouble, they announced unanimously in 2001—and again, more emphatically, in 2007—if we did not promptly begin to reduce our emissions of greenhouse gases. It was a triumph of rationality utterly without historical precedent; a drive to severely revise the entire world economy, relying chiefly on simulations whose workings only a few hundred experts actually understood.

References


