



HAL
open science

Connections between simulations and observation in climate computer modeling. Scientist's practices and "bottom-up epistemology" lessons

Hélène Guillemot

► **To cite this version:**

Hélène Guillemot. Connections between simulations and observation in climate computer modeling. Scientist's practices and "bottom-up epistemology" lessons. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 2010, 41, pp.242 - 252. 10.1016/j.shpsb.2010.07.003 . hal-03825631

HAL Id: hal-03825631

<https://hal.science/hal-03825631>

Submitted on 5 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Manuscript Number: SHPMP-D-09-00139R1

Title: Connections between simulations and observations in climate computer modeling. Scientist's practices and "bottom-up epistemology" lessons.

Article Type: Special Issue: Modelling and Simulation

Keywords: Climate models; modeling; computer simulation; observation data; scientific practices; validation; evaluation

Corresponding Author: Mrs Helene Guillemot, Ph.D.

Corresponding Author's Institution: EHESS - CNRS

First Author: Helene Guillemot, Ph.D.

Order of Authors: Helene Guillemot, Ph.D.

Abstract: Climate modelling is closely tied, through its institutions and practices, to observations from satellites and to the field sciences. The validity, quality and scientific credibility of models are based on interaction between models and observation data. In the case of numerical modelling of climate and climate change, validation is not solely a scientific interest : the legitimacy of computer modelling, as a tool of knowledge, has been called into question in order to deny the reality of any anthropogenic climate change; model validation thereby bring political issues into play as well. There is no systematic protocol of validation : one never validates a model in general, but the capacity of a model to account for a defined climatic phenomenon or characteristic. From practices observed in the two research centres developing and using a climate model in France, this paper review different ways in which the researchers establish links between models and empirical data (which are not reduced to the the latter validating the former) and convince themselves that their models are valid. The analysis of validation practices - relating to parametrization, modes of variability, climatic phenomena etc .- allows us to highlight some elements of the epistemology of modeling.

Connections between simulations and observation in climate computer modeling. Scientist's practices and « bottom-up epistemology » lessons.

Hélène Guillemot

Introduction

Anthropogenic climate change has been studied in climate science laboratories over the past thirty years. Over the past twenty years, it has become the object of global expertise with the publication of IPCC reports to which thousands of scientists have contributed. More recently, the issue of climate change has climbed to the top of the international political and diplomatic agenda, leading to major evolutions in considerations of economy, geopolitical equilibrium, North-South relations, consumption and lifestyles. Even if the problem has become increasingly political, scientific expertise retains its central role. From the beginning, climate change was a largely “science-driven” issue, mostly resulting from climate modeling and simulations of future climate change.

Considering the importance of these political and economic stakes, climate modeling has not given rise to much research in the social sciences. In particular, there are few studies of the validation and evaluation of models through comparison with observational data, even though this is an essential domain that endows models with their scientific character as well as with the credence we lend to simulations¹. Indeed climate scientists spend much of their time testing their models, comparing them to measurements of the real climate, and these evaluations constitute, in their opinion, the principal guarantees of the skill of models, of their scientific validity and of the reliability of future climate projections.

This work is in large part based on a study carried out between 2003 and 2006 in the two most important institutions that develop and use climate models in France, the Laboratoire de Météorologie Dynamique (LMD) of the CNRS, in Paris, and the Centre de Recherche de Météo-France (the French organism for weather forecast) in Toulouse. This study primarily relied on two types of sources: internal laboratory records (activity reports, planning reports, internal journals, Web sites, colloquia and workshop reports etc.) and extensive interviews with scientists. In the first part of this article, I draw a brief review of studies on computer modeling and climate models, and try to frame the role of observation in validation of climate models. In the second part, after a quick survey of the observational data used in climate science, and an

¹ I should define the terms *model*, *simulation* and *numerical experiment* as they are used in this article, because these definitions vary depending on the discipline, language, uses and authors. Here, I adopt the terminology currently used by climate modelers. The word *model* (climate model, Global Circulation Model, meso-scale model, etc.) designates a program that is meant to run on a computer and carry out algorithms step by step. *Simulation* designates the results produced by a model's output, i.e. the product of the calculations performed by the computer by means of that model in simulating a particular climatic configuration that is provided as input (for example, the simulation of the climate in the 20th century with such increase of CO2 levels with the MétéoFrance's model, or simulation of the last glacial maximum...). A *numerical experiment* is a simulation with the objective of virtually exploring the climate's behavior by varying parameters or representation of phenomena.

insight of the fundamental difficulties that faces validation of climate models, I seek to describe and analyze what sorts of relationships modelers establish between data and model simulations, basing on observing actual scientific practices through several examples : top-down validation, local evaluation of parametrization against field-experiments, validation of climatic effects of a small scale processus, evaluation of a model's capacity to simulate particular phenomena... In part three, I attempt to derive a few epistemological lessons on climate modeling and what distinguishes it from more traditional theoretico-experimental sciences in regards to their relationship to data.

I- Studies on numerical models, climate modeling and observation, a brief review.

1- Models between theory application and the creation of knowledge

Computer modeling, which emerged after World War II, has since the 1970s and 1980s permeated nearly all sectors of science, technology, industry and economy. Yet, it was only recently that it started eliciting any notable research activity in philosophy, history of science or Science and Technology Studies (STS). For a long time, discourse on models privileged a “semantic” approach that focused on theories and according to which a model is first and foremost a representation that gives meaning to a mathematical formalism. Over the past few years, however, there has been a renewed reflection on models. The “cult of theory” has been criticized by authors who assert models’ superiority over theories as representations of the world (Cartwright 1983, Cartwright 1999, Sismondo 1999). According to Mary Morgan and Margaret Morrison (1999), models consist of elements other than just theories and data, and “it is precisely because the models are partially independent of both theories and the world that they have this autonomous component and so can be used as instruments of exploration in both domains.” As autonomous agents, models can thereby become active “mediators” between theories and the world, and may have more to teach us than theory does on its own. Owing to this mediating function, the modeling of complex physical systems - systems that are poorly understood even when theories of underlying processes exist - cannot be reduced to calculations that apply known laws: “The computer simulations (...) involves a complex chain of inferences that serve to transform theoretical structures into specific concrete knowledge of physical system. (...) This process of transformation is also a process of knowledge creation, and it has its own unique epistemology.” (Winsberg, 1999, p.275).

Nevertheless, the cognitive approach to models is undoubtedly insufficient. The computer's massive influence, the emergence of new scientific objects, modeling practices that are more and more heterogeneous, as well as their increasing use in establishing expertise, makes “studying the activity of modeling in its institutional, technical and political environment, and without dissociating cognitive and social elements” (Armatte and Dahan, 2004) all the more necessary. According to Knuuttila *et al.* (2006), a convergence is already discernible between philosophers (up to this point focused on theories) and researchers in science studies (traditionally oriented towards the laboratory), who are sharing interest in the role of modeling and simulation in scientific practices.

2- What is at stake in studying climate models

Until very recently, climate modeling attracted few epistemologists, in part because it seems to fall under application, not under fundamental physics, bringing into play established physics theories and loosely-formalized heterogeneous elements. Moreover, it is embedded in expertise and embroiled in major political stakes. One would think that the latter characteristics would elicit the interest of the social sciences. This has been the case, but to a limited extent: the sociopolitical stakes inherent in climate models seem to have a contradictory influence in this regard.

In the context of climate change debates, models are often brought center-stage and placed into question -- how can they purport to predict the future of climate? Political controversies are frequently transposed to the scientific, and even epistemological, field: climate change skeptics shed doubt not only on projections but on the models themselves and seek to “stigmatize modeling as inferior science on philosophical grounds” (Norton and Suppe, 2001, p.67). Much of the time, such polemics focus on the relationship between models and observational data (Edwards, 1999). Computer models are opposed to “sounds science,” founded on data and solid theories; and the possibility of verifying models’ projections against data is questioned (Oreskes et al., 1994). In response to these criticisms, Paul Edwards has shown that if models and data are effectively in an relationship of interdependence – models being “data-laden” and data “model-filtered” – this relationship is not circular, but symbiotic, each gaining advantage from the other. In a detailed analysis, Stephen Norton and Frederic Suppe (2001) maintain that interdependence exists between theories and experiments as well, and that the absence of certainty, simplifying hypotheses, etc., are not the prerogative of models, nor do they constitute definitive obstacles to knowledge. These authors conclude that models may be trusted for the same reasons and to the same extent that traditional experiments are endowed with credibility or trusted.

If, despite these debates, climate modeling as a new mode of production of knowledge has attracted a small number of studies (comparatively to the economic or political side of the problem), it could be that the political role of these models makes their study more delicate, possibly risky. The research conducted in this domain has shown how the representation of uncertainty (Shackley and Wynne, 1996), the estimate of climate sensitivity (Van der Sluijs *et al.*, 1998), and recourse to flux adjustments (Shackley *et al.*, 1999) all result in negotiations or in scientists taking into account the expectations of policy makers. This has at times placed researchers in an uncomfortable position: emphasizing the co-construction between science and politics risks being co-opted by a critique of the objectivity of climate science and the validity of its results. Even when it comes to “defending the indeterminate character of the climate sciences,” (Shackley et al., 1999), any reference to uncertainty or ambiguity can be turned into a weapon in the hands of critics of the fight against global warming (Edwards, 1996). Tim Forsyth (2003) analyzed it well when he wrote: “these are controversial times for writing about environmental science and politics”. The political stakes can lead to closing black boxes more so than to analyzing their content; to prematurely transforming “matters of concern” into “matters of fact” (Latour, 2004).

3- Numerical experiments and the test of reality

In the historical and philosophical literature, the model's "credential" or "reliability" has been considered mostly in relation to theories and experiments – for example in the "identity" between simulations and experiments (Norton and Suppe 2001), or when it comes to the fact that for models, "the credentials develop over an extended period of time and become deeply tradition bound" (Winsberg 2003), just as various experimental techniques and instruments develop a tradition that give them their own internal stability, becoming « self vindicating » (Hacking, 1992). Less attention has been paid to the role of observation in evaluating the reliability of simulations and the validity of models (among these papers are Oreskes (1994), Edwards (1999) and Heymann (2006)). Before taking a look at scientists' practices, it might prove useful to clarify the role played by observations vis-à-vis numerical experiments in climate modeling.

The model's two-fold relationship to theories and experiments has puzzled philosophers who seek to understand "how simulation can have methodological and epistemological features in common with experimentation, while still playing the role of a form of scientific theorizing" (Winsberg 2003, p.106). The notion of a "numerical experiment" in this domain can be traced back to the first general circulation models (the climate model's ancestor). We find it under John von Neumann's pen, as well as in the works of many great names in meteorology (Nebecker, 1995). When Norman Phillips, who authored the first simulation of atmospheric circulation, was received at the Royal Society in 1956, he was presented as a "good experimenter"². Calculations based on hypotheses about atmospheric dynamics have been made in climate research since the beginning of the 20th century (Nebecker, 1995). However, the computer allowed for bringing into account multiple interacting physical phenomena under specified conditions. Now, a model that describes several processes interacting in a physical system, with conditions imposing specific limits, reproduces in numerical form certain traits and behaviors of that system. This is how modeling permits the exploration of virtual climates. These numerical experiments resemble laboratory experiments in their capacity to represent the system under study, by the possibility of varying parameters and testing theoretical hypotheses, as well as in the results they produce (data quantities to shape and interpret) (Galison, 1996). For some, there is no difference in principle between simulations and experiments (Norton et Suppe 2001). For others, simulations constitute a "second type of" experiment (Varenne, 2003): they coalesce and grow together heterogeneous formalisms, thus possessing the character of a "concrete experiment."

However, there is an obvious and essential difference. A computer experiment is based on symbols and digits; it does not confront the real world directly. A point in common with lab experiments is the capacity for manipulation; however, it does not possess the other essential function of experiments –the litmus test of the world (or at least a well-specified and prepared part of the real world). If we oversimplify: in theorico-experimental sciences, there is on one hand a theory that brings about calculations, and on the other, a lab experiment that produces measured results. *In fine*, experimental results are confronted with theoretical calculations. This principle is certainly highly schematic: in order to link theory and experiment (two terms designating very various realities), one needs "articulations" (Kuhn, 1977), "calculations"

² "By using a simple model and initial conditions which never occur in the real atmosphere he has been able to isolate, and study separately, certain fundamental properties of atmospheric motion – the kind of procedure adopted by all good experimenters" (Lewis, 1998, p.52)

(Hacking, 1989) or models -the term here having a much broader significance than numerical model (Cartwright, 1999). However, the important point is that the two processes result from distinct - even if interconnected - types of practices, and are under the charge of different communities with specific objectives and techniques (Galison, 1987). In modeling, on the contrary, the hypothesis to be tested (for example, a new climate mechanism represented by a new parametrization) and the numerical experiment are in a continuum : to verify a hypothesis, one first has to translate it into algorithmic form and insert it into the model; then one carries out a simulation by running this model. In this way, the model is at once theoretical and experimental... but it is a limited virtual experiment, cut off from the experiment's fundamental capacity for confronting the real world. The numerical experiment increases the range of explorable domains, but it doesn't produce "objective facts" and it postpones the moment of comparison with reality. The expansion of exploration that modeling allows does not exempt it from the necessity of confronting the world. Computer experimentation alone is insufficient; in addition to the "modeling" branch, an "observation" branch is needed.

II- The practices of climate modeling

1- Old ties between modeling and observation

The relation between climate modeling and observation can be traced back to the origins of this scientific discipline. Climate modeling is closely related to numerical modeling of the weather, which also emerged with the first electronic computers, just after World War II. While the invention of the computer was decisive, other ingredients were also necessary for modeling the weather and the climate, in particular progress in the observational instruments and networks that fed data to initial models and specified the characteristics of atmospheric circulation (Rochas and Javelle, 1993). Examples include the invention of radiosonde in 1927, as well as the data collected by metrological ships that starting in the 1920s circulated in order to aid the first regular transatlantic flights. However, above all it was World War II that gave rise to major developments in meteorology. Armies chartered metrological ships and airplanes, and installed observation stations that, for the first time, provided access to data on altitude winds, the stratosphere and cloud formation.

Without going into this vast domain, one must recall the central role of powerful and well-organized international scientific institutions in the establishment of powerful earth, ocean and satellite observation networks (Miller, 2001). The principal of these organizations was the International Meteorological Organization, created in 1873, becoming a United Nations agency in 1950 under the name World Meteorological Organization (WMO), which was behind many international programs (Edwards, 2006). It is important to emphasize that the computer did not only make modeling possible: it radically reconfigured the meteorology by unifying its three branches – observation, theory and forecasting – which were hitherto separate (Nebecker, 1995). All three branches had recourse to numerical computation and could therefore be compared and jointly used. In particular, the growing influx of observational data exploited the computer's capacity for collecting, classifying, statistically analysing and assimilating data in models.

The influence of observations, of experimental know-how and of measurement campaigns

on modeling also manifested itself in laboratories studied in France. The Laboratoire de Météorologie Dynamique (LMD) has, since its creation in 1968, built on the synergy between the various scientific tools used in the study of the atmosphere : theory, modeling, observation and instrumentation. In other laboratories that participate in the development of a GCM together with LMD (and that studied other components of the climate such as the ocean, the biosphere, etc.), the experimental or instrumental dimension is as important as modeling and precedes it. The same can be said for Météo-France. All these laboratories have strong and autonomous identities, focused on an object of research tackled using several tools. Traditionally, their scientists were often more interested in understanding climate (or ocean...) mechanism and developing parametrization close to the physics of the process than in building coupled global models. Modelers attach importance to links between observation and modelisation, and try to develop these relationships- even if they constituted a distinct community, separated from communities conducting observation, and if communication difficulties exist between the two. That culture of observation constitutes a shared value among French laboratories in this domain³.

2- Corrected, homogenized and reanalyzed observational data

What are the observational data with which simulations are confronted ? Three categories can be distinguished : data from vast networks of meteorological stations, on land and across the oceans; data provided by instruments on satellites (meteorological, oceanographic or scientific satellites); and finally, data collected during field experiments focused on a specific region, phenomenon and time period. Dozens of field campaigns have been organized since the 1970s, requiring massive human and technical investments (airplanes, ships, balloons), and bringing together numerous organizations (from meteorology, but also transportation, defense, forestry...)

It is important to insist on one essential characteristic of these data -- where they come from: they have been heavily reconstructed (Edwards, 1999). In the case of the climate sciences, “raw data” are not usable. What a satellite instrument is measuring is very much removed from the physical parameters which are needed by meteorologists or climate scientists: for example, in order to obtain the surface temperature of a point on the Earth’s surface from a signal captured by a meteorological satellite, no fewer than three intermediate models are necessary that combine initial data with numerous other factors (Kandel, 2002). Information coming from surface stations are not as indirect, but they are more heterogeneous and scattered, and therefore always have to be completed and harmonized with the aid of computer models. On the other hand, obtaining broad and homogenous data series over many decades (which is indispensable for detecting signatures of climate change), demand important reconstruction efforts: data must be corrected and re-calibrated to take into account changes in measuring technique or detectors. Météo-France developed a program that could locate the “ruptures in homogeneity” in measurements it had collected over a century (these ruptures were largely due to modifications in the instrument’s surroundings). This allowed Météo-France to reconstruct a homogenous evolution of climate in France throughout the 20th century (Moisselin *et al.*, 2002).

³ We don’t find in our study the distinction observed by M. Sundberg (Sundberg, 2007) between experimenters’ and modelers’ aims. On the one hand, we have not studied the community of researchers working on observations – only the community of modelers. On the other hand, the French modeler’s objective is (at least) as much to improve the understanding of climatic processes as to creating representations of processes that can be implemented in models.

Finally, an increasingly common type of data is the “reanalysis data sets”, which takes the interdependence between models and data even further. These are series of meteorological data from the past entirely reprocessed and completed with a climate model through a procedure inspired by the “variational assimilation” technique, which is routinely used in weather forecast in order to introduce observational data in near real-time into projection models – in other words, a kind of dynamic interpolation of past observations in a model. Thus reanalyzed data fill in the gaps and correct the shortcomings of observations: the data become complete, homogenous and coherent. Veritable hybrids of observation and modeling, these reanalyses add new uncertainties that are linked to the shortcomings of the initial data, to the uncertainties of the model, and most of all, to the sparsity of observations that are artificially filled in by modeling (Bengtsson *et al.* 2004)⁴. When using them to validate models, modelers try to take into account these uncertainties by, for example, lending less credence to the reanalysis of past periods than they grant to those of more recent periods (interview with A, Météo-France, january 2004).

3- Fundamental difficulties in the validation of models

Before turning to validation practices, let us quickly describe climate models. A numerical climate model (like a weather forecast model) is a Global Circulation Model (GCM): it seeks to simulate atmospheric circulation, represented by a three-dimensional grid, over the course of time. Within each grid cell and for each time step, the computer calculates the parameters that are characteristic of the atmosphere’s state from their values in the preceding time step by running the algorithms that constitute the model. The model is composed of two large parts: a part that describes the movement of air masses, whose algorithms are derived from fluid dynamics equations; modelers call it *dynamic* part. The other part, so-called *physics part*, calculates the forcing of atmospheric circulation ; it deals with vertical exchanges between atmosphere and outer space or with the Earth’s surface⁵ (the ocean, continental land mass or ice). These exchanges – of radiation, energy, water, etc. – occur at a much smaller scale than that of the model (which is on the order of several hundred kilometers) and they are represented by “parameterizations” that statistically reproduce, at the scale of a grid cell, the climatic effects of the phenomena under consideration. The parameterizations, veritable smaller models nested within the main model, are extremely diverse; certain more or less directly arise from physical theories (when it comes to radiation, for instance, the parameterizations synthesize quantum physics calculations), others are more empirical or phenomenological representations (vegetation

⁴ The reanalysis can be classified depending on the relative influence of the observational data and the model : « An A indicates that the analysis variable is strongly influenced by observed data, and hence it is in the most reliable class (e.g., upper air temperature and wind). The designation B indicates that, although there are observational data that directly affect the value of the variable, the model also has a very strong influence on the analysis value (e.g., humidity, and surface temperature). The letter C indicates that there are no observations directly affecting the variable, so that it is derived solely from the model fields forced by the data assimilation to remain close to the atmosphere (e.g., clouds, precipitation, and surface fluxes).” Kalnay et al., The NCEP/NCAR 40-Year Reanalysis Project, *Bulletin of the American Meteorological Society* 1996; 77: 437-471

⁵ Calculations of dynamic part are carried out on the three-dimensional grid, whereas physics part may be seen as a juxtaposition of air columns that don’t interact. Dynamic and physics parts deal with different kinds of variables, and have different time steps (the time step of the physics part being larger). Cf. Instruction manual for Atmosphere Global Circulation Model of LMD (2005) : <http://web.lmd.jussieu.fr/~lmdz/manuelGCM/main/node3.html>

ecosystems, for example) (Guillemot 2007, Sundberg 2007).

Meteorological models and climate models are therefore rather similar. However, they are used differently. In everyday forecasts, the atmosphere's future state is calculated in a deterministic fashion, starting from its state as measured a few days or hours earlier⁶. For longer ranges of a few weeks, the atmosphere's dynamic is chaotic, and only statistical properties are predicted⁷. Weather is not predicted for November 30, 2077; instead, averages (or variabilities) are provided for temperature, precipitation, etc., for the period between 2070 and 2100. This brings us to the first difficulty of evaluating climate models. For meteorological models, validation seems rather simple, at least in principle: forecasts can be compared the very next day to the characteristics of actual weather. In climate modeling, however, the simulations only have a statistical value, and the comparison can only be based on data series for present or past climate. Validating projections of future climate is even more difficult and indirect (I will return to this later).

The evaluation of climate models faces another fundamental difficulty: the entanglement of climate processes in the model makes validating the representation of a particular phenomenon extremely delicate. When modeling first began, the manageability of models gave rise to certain illusions about their capacity to explain the climate's mechanisms. Joseph Smagorinsky, the "father" of the first GCM, wrote that "The main advantage in diagnosing model simulations is that we know a great deal about the mathematical distortions we have introduced, and right or wrong, we have all the variables defined everywhere and all of the time" (cited in Nebecker 1995 p. 178). Jule Charney, another important pioneer of climate modeling who founded and directed the American Numerical Meteorology Project in 1948, noted with optimism : "When a computer simulation successfully synthesizes a number of theoretically-predicted phenomena and is in accord with reality, it validates both itself and the theories" (cited in Nebecker 1995 p. 180).

However, as models became more complex, it was necessary to observe that Charney's statement could backfire: the apparent validation of a model through observation could result from the combination of a false hypothesis and a faulty representation in the model, which together give a good climate simulation due to what modelers call "error compensation." A simulation could thereby appear "correct but for the wrong reason." Model validation proved to be more delicate than theory validation: theories are also under-determined by experiments, in accordance with Duhem's thesis, but the difficulty is pushed to a higher degree in models, which bring into interaction numerous intricate hypothesis and representations of processes.

As models incorporated a growing number of elements and processes, scientists were confronted with the difficulty of interpreting their simulations, of attributing a cause to a phenomenon and understanding multiple feedbacks. The large number of interacting mechanisms, which constitutes the specificity of models and makes numerical simulations and experiments possible, is also what makes their validation particularly arduous (Lenhard and

⁶ Nowadays, meteorological agencies also predict the weather 4 to 10 days in advance based on probabilistic "ensemble" forecasts.

⁷ These differences in scale can introduce certain differences between weather forecast models and climate models. For example, ocean currents or polar ice dynamics are represented in climate models, but not necessarily in meteorological models. Reciprocally, local details are important for weather forecasting, but not for climate modeling.

Winsberg, 2008)⁸. More and more, the researchers' work came to consist of understanding what happens in their model, which parameter influences such and such result and how. The model itself, "almost as difficult to understand as the real climate", became an object of study. "Climate modeling means working on what comes into the model and what comes out of it, it means understanding what there is in the model, how it reacts, what are its characteristics" explains a researcher at the LMD (interview with B, LMD, July 2003).

4- Evaluation on all scales

Modelers tend to distrust the term "validation" and prefer to use expressions like "evaluation" -- the norm in IPCC reports (Randall et al., 2007). The concept of validation was criticized in a *Science* article (Oreskes *et al.* 1994) that maintained that numerical climate models could be neither verified nor validated in a rigorous fashion, but could only be confirmed under certain conditions. The article gave rise to many debates, and modelers, prompted to a modesty of sorts, declared more humbly that it was a question of "evaluating the model's performance" or of showing that a model is "sufficiently good to be useful".

Beyond these nuances, the first essential point is that a model is not evaluated in general. What is evaluated is its capacity to account for a particular climatic characteristic or a defined phenomenon. Evaluations are performed at all temporal and spatial scales and at all levels of the model: they might concern characteristics of average global climate, a geographically limited phenomenon, the model's capacity to represent a specific feedback, etc. The type of validation depends on how the model is used: a model might be appropriate for medium-term simulations (for the study of interannual variability, for instance), but not for simulations spanning one hundred years. Consequently, there is no systematic protocol for evaluating models. An evaluation supposes a preliminary question, the definition of a problem for which an appropriate procedure for confronting models and data must be imagined. "It requires astuteness and creativity more so than respect for a method," summarizes a LMD researcher (interview with C, LMD, July 2003). Modelers must know how to make use of everything at their disposal in order to "constrain the system" (interview with D, Météo-France, January 2004). For example, the eruption of Mount Pinatubo in 1991, presented researchers with a unique opportunity to validate the representation of aerosols in models.

Despite the plurality of evaluations, two greater types of approaches are distinguishable, one "top-down" and one "bottom up." Top-down validations consist of comparing simulations to global data series. This is an older type of validation: models were initially validated in relation to average climate, by comparing a map of the simulated climate with a map of climatic averages based on observational data, in order to see if the model correctly reproduced large climatic characteristics (temperature, wind, main phenomena). Subsequently, it was also necessary to validate the climate's variability: seasonal cycles, interannual and decennial variability, monsoons, extreme event frequency, El Niño... Model evaluation concerned itself with larger and larger variability domains. In this way, Météo-France studies the capacity of its model

⁸ Lenhard and Winsberg (2008) develop a similar analysis of what they call « confirmation hollism » caused by the « entrenchment » of climate models, which makes them « analitically impenetrable ».

Arpège to reproduce the variability spectrum of precipitation, including extreme precipitation, by studying the distribution of rain intensity down to the daily scale.

Since the 1990s, another kind of validation (bottom-up validation) has been used for testing parameters against field experiments. The procedure consists of studying a climate phenomenon in those geographic locations where it is predominant (for example, the monsoon in India, or winter storms in the North Atlantic) and using this case study to test the manner in which this phenomenon is represented (or “parametrized”) in the model (Chaboureaud and Bechtold, 2002). This methodology has been used to test many parametrizations through numerous field experiments – for example parametrizations of clouds in several GCM in Europe have been tested within the framework of program EUCREM and EUROCS⁹. The procedure can be decomposed into three steps, within which intervene no less than three different models. The first step consists of carrying out a measurement campaign within a zone that is monitored by observation stations. Then –this is the second step - a meso-scale model, also called « cloud resolving model » (with grid cells measuring a few kilometers) is used to simulate weather evolution in this zone during the time period under consideration: meso-scale specialists enter into their models parameters measured at the beginning of this period, and define limit conditions so that the simulated climate resembles the climate observed during the campaign. In this way, the meso-scale model is validated by observations in a detailed fashion. Finally, climate modelers test the studied parametrization in a simplified, one-dimensional version of the model. This so-called “column” model consists of a single horizontal grid cell with all of the vertical layers superimposed ; it contains all of the « physics part » of a GCM but no dynamic, so, it is far less difficult to use. This one-dimensional model is provided in input with the external climate data from the meso-scale model, then it runs and its simulation is compared to that produced by the meso-model, which permits modelers to validate its parametrization, and eventually, to ameliorate it (interview with B and E, LMD, july 2003 and july 2005).

To summarize: modelers compare the climate simulated by a one-dimensional model equipped with the parametrization to be tested, to a climate simulated by the meso-model that is validated by observation. In this sophisticated methodology, the meso-scale model plays the intermediary between the data and the large-scale model, by providing this large-scale model with a complete set of «predigested » data reconstituted from observations. Such studies of parametrization through field experiments have multiplied over the past fifteen years or so due to the availability of operational meso-scale models (such as the Méso-NH model developed by Météo-France).

In following these steps, what remains to be done is the implantation of the new (or ameliorated) parameterization in the GCM, and it’s here that the real difficulties begin (joke the modelers), since one invariably obtains a simulated climate that is catastrophic! This has to do with the error compensation mentioned above: even though the preceding model was equipped with a parameterization that is farther from the physical mechanism, the model’s errors and approximations compensated each other in order to engender an accurate climate. The parameterizations in a model constitute « a family » (interview with A, Météo France, january 2004), and the implantation of a new parameterizations requires much research, testing and control.

⁹ See EUROCS website : <http://www.cnrm.meteo.fr/gcss/EUROCS/EUROCS.html>

5- Evaluating long-term feedback between clouds and radiation

Nevertheless, validation cannot be reduced to global evaluation of a model by comparing it to global databases on the one hand, and validation of parameterizations by local case studies on the other. In particular, one must evaluate the so-called « climatic effects » of some parameterizations. Small-scale processes represented by parameterizations have consequences that cannot be perceived during field study, but that become apparent in the long-term and over longer distances. These climatic effects also must be reproduced through parameterization. One of the most important is the feedback between clouds and radiation. Clouds are the most influential element of climate, and their effects are particularly complex and difficult to model, specially their interaction with radiations - which they can reflect or absorb depending of the wavelength, type of clouds and other parameters. While cloud formation, convections or precipitations can be studied through local campaigns, the interaction between clouds and radiation are not proximately observable (interview with A, Météo-France, january 2004).

It was the global warming issue that made necessary the study of the long-term impacts of these parameterizations. Indeed, climatic changes are nothing other than distant and large-scale consequences of a small scale warming, provoked by an increase in green-house gas concentration. This warming provokes a disequilibrium in the atmosphere's lower layers that interacts with numerous small-scale mechanisms, that are represented by parameterizations : turbulence, cloud microphysics, etc. (Le Treut, 1999). A number of studies have established that different cloud representations are primarily responsible for the uncertainty of models as well as for differences between different model's projections of climate change. It is therefore crucial to understand how cloud-radiation feedback evolves with climate change, and to evaluate if models correctly reproduce this evolution. Because they cannot observe future climate, climate scientists have to settle with exploring the correlations between climatic variables with increases in temperature. However, they face an additional difficulty: temperature increase principally affects atmospheric dynamics, which favors certain types of clouds. This first-order effect masks the impact of temperature increases on cloud microphysics and on convection heights – in short, on the physical evolutions of clouds which have important, long-term climatic impacts because they affect cloud-radiation interaction (interview with F, LMD, july 2003).

A key point in modeling is to wisely choose the variables that characterize the phenomenon under study, called “diagnostics” by the modelers. During a stay at the Godard Institute for Space Studies (GISS) in New York, a young LMD researcher (F) developed an original method for exploring cloud-radiation feedback and its sensitivity to warming (Bony *et al.*, 2004). Drawing on all sorts of observational data on clouds (from satellite databases to reanalyzed data, etc.) and using statistical analysis, F sought to establish the diagnostics that showed the effect of temperature on cloud physics, in given dynamic conditions. Her new diagnostic has been used later to test the capacity of cloud parametrization of the LMD climate model to represent this feedback. Let us expose briefly the principle of this diagnostic.

Instead of representing clouds according to longitude and latitude, F decided to represent them along a single axis, as a function of the vertical speed of the air. “In this way, we can represent on a single axis, on one side where the air rises, on the other, where it goes down,” she

explains, “Over there are all of the convective clouds, here the low clouds [...] In classifying all of the zones as a function of vertical speed, we obtain continuous variation in properties of clouds and of the water vapor ; everything is organized in a simple manner... It’s much easier to analyze what is happening [...] Instead of looking at things regionally, we try to classify them in a synthetic manner.”

By coming up with this diagnostic (widely used ever since), which provides the specialist with a new synoptic panorama of cloud organization, the scientist was able to extract new information from a flood of available data. The search for new ways of seeing is described as creative and ludic: “One can choose to look at things in more detail, by looking at different variables, by carrying out statistical analyses (...) What interests me is looking at things otherwise; in a somewhat more twisted way...”

6- Making simulations and data “speak the same language”

A diagnostic is a means of analyzing observational data, but also the output of the model’s simulations. Diagnostics therefore constitute an evaluation method : applying the same diagnostic to simulations and observations renders them commensurable and allows for the comparison of two schemes constructed in accordance with the same criteria. Some climatologists enforce this rule when evaluating models: “use observations and simulations in the exact same way” (interview with F, LMD, July 2003). For example, they may choose to study the radiation sent back into outer space by the outer layer of the atmosphere because satellite instruments measure this flux. However, they would avoid working on the radiation’s vertical distribution, which the model can calculate, but which cannot be measured. In this case, researchers voluntarily limit the field of numerical experiments so that it only includes those that can be directly compared to observation.

On this theme, one can give two additional examples of the model-data relationship. LMD researchers who wished to evaluate their model’s reproduction of the monsoon in India, used a program for analyzing meteorological data that was developed by a researcher at the European Centre for Medium-Range Weather Forecasts in Reading (UK). This program selects from data the parameters that characterize monsoon depressions, thereby constructing trajectories; subsequently it produces a statistical representation of the properties of these perturbations. With the help of this program, LMD researchers have analyzed monsoon depressions simulated by their models in a region in India over nine (simulated) years. Among those virtual depressions, they’ve identified those that have the characteristics of a monsoon, and have compared them to real observations of Indian monsoons that are analyzed by the same program. Their conclusion is two-fold: “(1) LMD’s model is capable of simulating monsoon depressions that exhibit realistic circulation characteristics, (2) the “tracking” method is powerful and offers a panoply of robust quantitative statistical studies that make it possible to effectuate a precise analysis of the simulated depression systems”¹⁰. With this example, we can see that research does not limit itself to evaluating a model; it is also concerned with the method of analysis used, and with monsoon displacement. This work actually deals with three levels of analysis: the monsoon (whose

¹⁰ Extracts from LMD’s internal journal *LMDZ Info* n°0 (July 2000), p. 12 : « caractérisation des dépressions atmosphériques simulées par le modèles de circulation générale du LMD (MCG LMD6) ».

statistical characteristics are under study); the model (which is to be evaluated in relation to its capacity to reproduce the monsoon); and finally, the common method of analysis (whose efficacy is under evaluation).

Another example, which will also keep us in India: this time it's a matter of validating the model's capacity to simulate a tropical hurricane that hit the northeast of India on October 29, 1999. First, researchers introduce into the model the atmospheric conditions of October 21, 1999, and then the model simulates the climate of the following ten days with the zoom function pointed at India. A program made it possible to transform calculated climate characteristics by simulating the signal that would be detected by the Meteosat satellite if it observed the phenomenon. Thus researchers could easily compare "real" images of cyclones taken by Meteosat in October 1999 with simulated satellite images (this comparative method is named "from model to satellite."). In all of these examples, it's a matter of "making models and measurements speak the same language"¹¹.

III- Bottom-up epistemology of climate modeling

1- Some epistemological remarks on the relationships between simulations and observations

In order to be more thorough in this outline of the relationship between data and simulations, I would have to give more examples, notably examples of how simulations of past climate are compared to paleoclimatic data. Besides, intercomparison projects organised by the Program for Model Diagnosis and Intercomparison (Atmosphere Model Intercomparison Project, Paleoclimate Model Intercomparison Project, Coupled Model Intercomparison Project etc.) play an essential role in these evaluations, and in structuring the modeling community as well ; they deserve to be discussed in details (see Lenhard and Winsberg 2008). However, the cases described above, even if succinct and limited, do already point to a few reflections on a bottom-up epistemology of modeling that emerges from an analysis of practices.

The first remark concerns the manipulation of data. The computer does not merely engender virtual climates, it also allows researchers to manipulate observations or rather « to look at things otherwise, in a somewhat twisted way » (as F pointed out) in order to uncover hidden correlations. Due to open access to databases and reanalyses, playing with observations is an integral part of the modeler's work. We are quite far here from the opposition between fixed and untouchable data and readily malleable simulations. It's almost the inverse : having already undergone all sorts of transformations, the data can be sorted further and « twisted » in all directions since they come from observation and remain unswervingly linked to it. Simulations, on the other hand, demand a more cautions handling : because they have no link to the real world, they are bound to always being measured against data.

Certainly, the fact that data are not given, that they are calibrated, corrected, reduced, that they need theory and modeling – all of this is not new and has been studied by numerous authors

¹¹ Extract of « Rencontre Modèles – Données », LMD's internal journal *LMDZ Info* n°0 (July 2000)

(Galison 1987, Hacking 1989, among others). The general character of data manipulation was one of the arguments Norton and Suppe (2001) used to counter critics of model validation. Knorr Cetina (1992) remarks that “the theoretical relevance of laboratories” rests “upon the malleability of natural objects.” Within the domain of the field sciences, Latour (2001) impressively analyzed the “circulating reference” (the successive transformations of data) in pedology (science of soil). The work on climate observation that I have described recalls Latour’s description in this regard: Between science and its object, there is neither correspondence, nor gaps, Latour writes, but a succession of small displacements, « cascade of transformations ». Similarly, we have seen that cascades of transformations allow back and forth between data and climate simulations. « This chain has to remain reversible. The succession of stages must be traceable, allowing for travel in both directions » (Ibid p. 72). This claim applies of course to climate data¹². In both cases, we find the same capacity of combination and comparison of data, allowing new transformations : « Hardly surprising, then, that in the calm and cool office the botanist who patiently arranges the leaves is able to discern emerging patterns that no predecessor could see (...) Scattered through time and space, these leaves would never have met without her redistributing their traits into new combinations » (Ibid, p.38). If we stand in the leaves for clouds, and the botanist for the climatologist, we would think we are reading a description of F exploring cloud’s properties. Regarding the manipulation of data, there is no difference in principle between climate modeling and other sciences. Nevertheless, this data malleability is of a much higher degree in the climate sciences, due to the extensive use of computers on all levels.

The second remark is that exchanges between models and observations are more symmetrical than we would believe. This circulation continually moves in both directions, in an iterative back and forth, whereby both data and simulations are repeatedly compared and discussed. Therefore, even if customarily models are evaluated based on observations, the inverse also occurs and models serve to validate methods of interpreting the data. This is especially the case in paleoclimatology. The interpretation of data from the distant past is particularly delicate because it relies on relationships that are themselves depending on the climate. For instance, reconstructing climate from pollen makes use of relationships (between properties of pollen and parameters such as temperature and precipitation), that may have been very different in the past from what they are today. If divergences between the model and the data are found, they can be explained in many ways : model errors, a process poorly accounted for in modeling, but also errors in data, or moreover in data interpretation - or any combination of these different factors.

Thus, contradictions between paleoclimatic simulations and data has sometimes raised doubts about the latter : for example, an estimate of ocean surface temperatures during the last glacial maximum was shown to be false after it happened to be incompatible with many models. The “reconciliation” between models and data is the result of dialogues and « iterative work » between data specialists and modelers (interview with G, Laboratoire des Sciences du Climat et de l’Environnement, june 2006) – this cooperation playing a notable part in paleoclimatology. The iterative dialogue between simulations and observations is facilitated by the availability of

¹² Without going into a fresh controversy that has opened up as this text is submitted for publication, and which does not encroach on the problematic of this article, the « ClimateGate » affair (the hacking of more than 1000 emails from the Climate Research Unit (UK) and the polemics that followed few days before the Climate Change summit in Copenhagen in December 2009), highlights the crucial importance of accessibility and traceability of scientific data in a highly political context.

powerful computers, which made it possible to make the two “speak the same language” to « use them in the same manner » (according to the rule followed by some modelers) ; in the examples given in Part two, the common language is that of the diagnostic, of a method of analysis or instrumental measurement. As in all sciences, here one obtains charts, curves, and histograms – produced by the computer, and arising from simulations as from data – that constitute what is compared, analyzed and discussed, and what constitutes proof (Latour 1989).

The third characteristic noticed here proceeds from this back and forth : models are at once (or alternatively) validated by observations and used to complete the data. They are the object to explore and the tool of exploration; they inseparably serve to understand the climate and make predictions about it. Because of this, it is impossible to strictly distinguish a models construction from its use. Scientists go back and forth between the development phase and the simulation phase, relying on the model to study and project the climate and on the climate to explore and ameliorate the model, following a multitude of different modalities. (Heymann (2006) made a similar point for atmospheric chemistry modeling). If model development involves its utilization (even if just for the sake of testing), using the model generally requires participating in its construction, or at least sufficiently knowing its content to be able to carry out sensitivity experiments or studies of variability. Stated differently: the model is not (yet) a black box, a tool easily accessible to all users, a machine for producing color images, in the words of a LMD scientist (who dreads such use of today’s models). The model remains an object of research.

There is a fourth point to raise. What is evaluated is the capacity to reproduce or account for a phenomenon (the monsoon, for instance), or a climatic characteristic (for example, variability of precipitation). And this evaluation is never an isolated task : every combination of phenomena, available data and urgent questions suggests certain methods of analysis to the scientists’ imagination. A general protocol for validation does not exist – no more than a universal norm for experimental testing (Atten and Pestre, 2002). Norms of validation are defined at the same time that the facts to be validated, as well as the analysis methods to which data and models will be subjected and that allows for their comparison. We could say that the researchers must construct the phenomenon to be validated. Evidently, the researchers do not construct the phenomenon called Indian monsoon; however, it is not the simulation of the monsoon “itself” that is under evaluation, rather the set of parameters characterizing the monsoon - that is effectively constructed.

Fifth remark : these modes of analysis, which bring out structures, correlations and retroactions in the models and the data, are also what allows modelers to explain and understand phenomena by simplifying them. Evaluations also have a heuristic function. The computer’s combinatorial and sorting capacity, in allowing modelers to bring order to observations and to classify clouds, helps triumph over tangle and multiplicity of data and overcomes (in part) the famous complexity of clouds and their climatic interactions. Here we recall Latour’s remark about the botanist sorting leaves in her lab : « ... a pattern emerges (...) and here again, it would be astonishing if it were not the case. Invention almost always follows the new handle offered by a new translation or transportation » (Latour 1999, p. 53).

2- The question of realism in climate modeling

The last remark deals with a particularly delicate point, the realism of models, which deserve a longer exposition. Models do not aim to “imitate” the real climate (which would hardly make sense). Every model is incomplete, simplifying, and it is the result of choices, definitions and conceptions made by its authors. Modelers, as we have seen, seek to look at things according to new frames of references. Is there anything less realistic than clouds that are lined up as if ready for battle, or monsoons reduced to their statistical characteristics? It is a matter of understanding climatic mechanisms, increasing confidence in models and projections of climate change, which does not entail producing realistic representations. Simulations and observations must “speak the same language,” but this is not the language of the real climate¹³. Even data do not constitute a faithful representation of the climate. The referent to which simulations are compared is not “the climate,” it is a set of remodeled data that are carefully selected from the hundreds of thousands of data sets furnished by instrument networks on land, sea and in outer space. Only an instrumented world is capable of providing the data that are used to test models.

Yet, the question of realism is at the heart of the image of models that scientists present to policy makers and the public. Models are in part constituted from algorithms that reproduce laws of dynamics and thermodynamics and in part from parametrizations that scientists continuously seek to bring “closer to the physics of the process.” An increasing number of phenomena and cycles that influence the climate are taken into account. Spatial and temporal resolutions are improving due to more powerful computers and means of observation, giving a better geographical representation of the globe. It is because computers allow « pragmatic interaction without theoretical background » of model components (Küppers and Lenhard 2006), because they allow all of these “formal heterogeneous systems” (Varenne 2007) to interact and work together, simulating their temporal development, that simulations have come to better reproduce some climatic phenomena (for instance, coupling with vegetation helps explain strong Asian monsoon in the Holocene epoch). Progress in this domain seems to move towards finer resolution, detail, and proliferation of elements, for an increasingly faithful representation of the climate. When climate scientists address the non-specialist public, they normally begin by presenting a sketch of this progress from primitive atmospheric models - bare globes with crude grids - from the 1970s to today’s “earth systems” populated by all sorts of chemical and vegetal species, mountains, seas, cities and rivers. This representation seems to be the exact opposite to that given by high-energy physicists – the famous “quest for the ultimate particle”: instead of a reductionist race towards the elementary and towards unification, climate modeling would advance towards an horizon of comprehensive representation of everything.

The image is nevertheless too simple. While this path towards complexity is a serious trend in climate modeling, to which the striving for realism is often linked, it is but one driving force of modeling, and it is not without problems. First, global Earth System models are not the only models used by researchers; they are assisted by simpler models, “idealized” models, etc. In addition, there are scientific debates on the best means of improving models. Certain scientists think that we will better understand and project the climate by improving parametrizations of

¹³ That’s still recalls Latour’s analysis about pedology in Amazonia: « But these acts of reference are all the more assured since they rely not so much on resemblance as on a regulated series of transformations, transmutations and translations » (Latour 1999, p. 58) « It is not realistic ; ot does not resemble anything. It does *more* than resemble. It *takes the place of the original situation*, which we can retrace... » (Ibid, p. 67)

essential atmospheric phenomena as opposed to describing in detail various types of vegetation, while others think that modeling will benefit most from taking into account new factors. Other internal debates exist on the limits of GCMs, for example, on the question of to what point can these models be used on the regional scale to respond to growing social demand. Illustrating the problematic character of the notion of realism, climate scientists use this term to designate a number of different things: they speak of the realism of a description of a process – of a parametrization, for instance - and of the realism of a simulation. Now, those two realisms are often initially incompatible (I evoked the difficulty of implementing new parametrizations earlier) – the choice to privilege one or the other depends on the objective of modeling, as well as on the culture and institutional framework of the laboratory (Shackley et al., 1998). The researchers' work consists of defining what counts, that cannot be reduced to an ineluctable advance towards complexity or maximum resolution. However, this scientific work, and the question of realism in particular, cannot be isolated from the political stakes involved in climate modeling. Climate projections rely on models that are expected to provide to policymakers increasingly better representations of all climatic elements in order to produce a higher-quality and more precise projection, notably on the local scale. It is necessary to discuss the naive vision of models, to cast doubt on their realism, to underscore their limits and to challenge that conception of science-policy relationship based on the « linear model » (Sarewitz 2000, Pielke 2002), even if it can pose the risk being confused with (or used by) the opponents of climate change policy.

3- Modeling and field science versus theorico-experimental sciences. Similarities and differences

To what extent are the preceding epistemological remarks specific to the relation between numerical models and observations? Could the same analysis be carried out in regards to the relationship between theory and experimental results? That data are reconstructed, that validation norms are defined at the same time as that which is to be validated, that sciences do not produce mimetic images of reality, all of this results from very general findings. On the level of field observations, there don't seem to be differences in principle between the older classical sciences and climate modeling. Of course, differences in size and scale exist, since the climate sciences depend on a veritable (technical, meteorological, administrative...) "mobilization of the world" (Latour 1989, p. 539), which has allowed for the development and maintenance of a dense and complex measurement network. Numbers representing objects and properties involved in climate (meteorological measurements on the ground, cloud measurements from satellites, water measurements from ships, characteristics of vegetation...) circulate, are concentrated and combined, and this information coming from all over the world is made available to models. In order to master the avalanche of data inside this gigantic "center of calculations," it is still necessary to classify, superimpose, invent new transformations that will reduce the number of these objects, summarize their relationships and transcribe them to our scale. Computation permits these manipulations at a grand scale, and allows for virtually transforming the climate and experimenting on observations. The change is by degree: the computer prolongs an old history, offering increased computational power and superimposition capacity (Latour, 1989). The power of computer science endows models and data with unprecedented plasticity (Pickering 1995) and allows for new connection between them.

As we have seen, numerical computation has made possible the modeling of a system containing many heterogeneous elements and processes that interact nonlinearly, but in doing this it has also made validation particularly arduous. A similar underdetermination is found in observations: climate scientists have at their disposal and abundance of data collected in massive quantities, which are at the same time more global, plural and incomplete than experimental measurements. If experiments are carefully tailored to testing a theoretical hypothesis, observations in the climate domain are rather “ready to wear”: it is necessary to choose, select, and even adjust if need be. How is it possible to move past this double underdetermination of models and observational data? How can the infinite potentialities of the model be confronted with the infinite realities of the field? The few examples we have seen have given an idea of it : by choosing a domain, a definition, a series of questions, and a mode of analysis through which the model’s and the real climate’s functioning can be studied at once.

More fundamentally, there is a difference between the relations that link theories and experiments on the one hand, and modeling and field observations on the other. All scientific disciplines share the requirement of testing that permit to decide between statements. An experimental statement benefits from its capacity to resist controversy by demonstrating that it is not a simple fiction, the instrument having been constructed precisely so that “the experimental fact would be explained by the answer to the question posed” (Stengers, 1993). This is not the case for numerical modeling because models have to do precisely with fiction. Like fictions, and unlike theories, models have the capacity to “express content not immediately situated in the dimensions of true or false” (Barberousse and Ludwig, 2001, p. 23) they can “at once contradict accepted hypotheses and tell us something about the world” (*ibid.*, p.36). No model has exclusivity on representing a phenomenon, since many representations coexist that meet different needs and that are sometimes linked to choices made by the author¹⁴. As far as field observations are concerned, it is impossible to prove with certainty the stability of established relationships to their subjects (this is one of the difficulties of validating simulations of future climate and interpreting data from the past). A valid explanation for one terrain will not necessarily hold in another (it often happens that a parametrization validated for one region of the globe is not suitable for another because the dominant processes were not identical). Situations cannot be purified, “no single cause has the general power to cause, each one is part of a history and it is from this history that it gets its identity as a cause” (Stengers, 1993, p.159). Any validation of a model with data from the field has to take into account these multiple causes that vary in time and space.

Conclusion

The central importance of observation for modeling cannot be overemphasized. There has been a parallel evolution between progress made in modeling and the collection of larger and larger data quantities, by satellites in particular. If climate modeling has advanced to a degree unparalleled in the environmental sciences, this is due to the exponential growth in computational power and to resolution of equations governing atmospheric circulation, but also due to a unique and dense instrument network that is highly developed and organized on an international scale.

¹⁴ In climate models, parametrizations are often identified by their author’s name - we hear of “convection pattern from X” or « boundary mayer scheme od Y »

Let's return to the the questions posed by the philosophers evoked at the beginning of this article. If models do not directly represent the real world, if their numerical experiments are not about the world, how can they bring knowledge about reality? What gives models their credential? What guarantees knowledge obtained from simulations? We have looked for the responses in researchers' practices. What establishes scientists' trust in their own models, what guarantees the value of the knowledge they produce, is validation or evaluation by observational data, much more so than their relation to theories or experiments.

Relationships to theories undoubtedly constitute one of the originalities of climate modeling, and distinguish it from other modes of knowledge production. Its capacity to integrate heterogeneous theories, milieus and elements, to span scales of time, space and complexity, to combine the singular and the contingent, along with its faculty for simulation, taking into account historical evolution and fictional scenarios – all of this constitutes new ways of understanding scientific knowledge peculiar to numerical modeling. On the other hand, in its relations to the real (instrumented) world, climate modeling resembles other scientific disciplines, thereby establishing its credentials and reliability. It is these relationships between models and observations, these languages invented by scientists and common to simulations and data, these multiple original and robust links, that fasten climate modeling to the material world, and that contribute to the improvement of models and scientists' understanding of climatic phenomena.

Bibliography

- Armatte M., Dahan A. (2004). Modèles et modélisations (1950-2000) : nouvelles pratiques, nouveaux enjeux, *Revue d'histoire des sciences*, 57/2
- Atten M., Pestre D. (2002). *Heinrich Hertz. L'administration de la preuve*. Paris : PUF
- Barberousse A., Ludwig P. (2001). Les modèles comme fiction, *Philosophie*, n°68, p.16-43
- Bengtsson, L. *et al.* (2004). Can climate Trends be Calculated from Re-Analysis Data?. *J. Geophys. Res.*, V109, doi:10.1029/2004JD004536
- Bony S. *et al.* (2004). « On dynamic and thermodynamic compônents of cloud changes », *Clim. Dyn.*, 22, 71-86
- Cartwright N. (1983). *How the laws of physics lie*, Clarendon Press
- Cartwright N. (1999). Models and the limits of theory : quantum Hamiltonians and the BCS model of superconductivity, in Morgan M. S. and Morrison M. ed. (1999). *Models as mediators : perspective on natural and social science*. Cambridge : Cambridge University Press
- Chaboureau J.-P., and P. Bechtold, (2002). A simple cloud parameterization derived from cloud resolving model data: Diagnostic and prognostic applications. *J. Atmos. Sci.*, 59, 2362-2372.
- Dahan Dalmedico A. (2001). History and epistemology of models : meteorology as a case study (1946-1963), *Arch. Hist. Exact Sci.* 55, pp. 395-422.
- Dahan Dalmedico A., Guillemot H. (2008), *Climate Change: scientific dynamics, expertise and political challenges*, in *Science and Sovereignty*, C.Paradeise (ed.) Routledge, New York.
- Edwards P. (1996). Global comprehensive models in politics and policy making, *Climatic Change*, 32 (2), 149-161
- Edwards P. (1999). Global climate science, uncertainty and politics : data-laden models, models-filtered data, *Science as Culture*, vol 8, n°4
- Edwards P. (2006) : « Meteorology as Infrastructure Globalism », *Osiris* 21 : 229-50
- Fleming J. (1998). *Historical perspectives on climate change*. Oxford : Oxford University Press
- Forsyth T. (2003), *Critical Political Ecology: The Politics of Environmental Science*. Routledge, New York
- Galison P. (1987). *How Experiments End*. Chicago : University of Chicago Press
- Galison P. (1996). Computer simulations and the trading zone, in Galison P. and Stump D. (dir.), *The disunity of science : boundaries, context and power*. Stanford : Stanford University Press
- Guillemot H. (2007). Les modèles numériques de climat, in Dahan A. (dir.). *Les Modèles du futur*. Paris : La Découverte
- Hacking I. (1989). *Representing and intervening*. New York : Free Press
- Hacking I. (1992). The self-vindication of the Laboratory sciences, in Pickering A. (dir.), *Science as practice and culture*. Chicago : University of Chicago Press.
- Heymann M., (2006) : Modeling reality. Practice, knowledge, and uncertainty in atmospheric transport simulation, *Historical Studies in the Physical and Biological Sciences*, September 2006, Vol. 37, No. 1, p. 49–85
- Kandel R. (2002). Les modèles météorologiques et climatiques, in Nouvel P. (dir.), *Enquête sur le concept de modèle*. Paris : PUF
- Knorr-Cetina K., (1992). The couch, the cathedral and the laboratory : on the relationship between experiment and laboratory in science, in Pickering A. (dir.), *Science as practice and culture*. Chicago : University of Chicago Press
- Knuuttila T. *et al.* (2006). Computer models and simulations in scientific practice (Editorial), *Science Studies*, Vol. 19 n°1, 3-11
- Latour B. (1989). *La science en action*. (réédition Gallimard 1995)

- Latour B. (2001). *L'espoir de Pandore*. Paris : La Découverte. English version (1999): *Pandora's hope. Essays on the reality of science studies*. Harvard University Press
- Latour B. (2004). Why Has Critique Run Out of Steam ? From Matters of Fact to Matters of Concern, *Critical Inquiry* Vol 30 n° 2 pp.25-248
- Küppers G. et Lenhard J. (2007). « From hierarchical to network-like integration : a revolution of modeling style in computer simulation », in Lenhard J., G. Küppers and T. Shinn (ed.): *Simulation: Pragmatic Construction of Reality, Sociology of the Sciences*, vol. 25, Dordrecht: Springer
- Lenhard J. and Winsberg E. (2008). Holism and Entrenchment in Climate Model Validation, (*draft ??*)
- Le Treut H. (1999). « Le défi de la modélisation du climat », in *INSU, 30 ans de recherche en sciences de l'Univers 1967-1997, 1997-1999, Lettre du Programme International Géosphère-Biosphère- Programme Mondial de Recherche sur le Climat (IGBP-WCRP) n°9*
- Lewis J.-M. (1998). Clarifying the dynamics of the general circulation : Phillips's 1956 experiment », *Bulletin of American Meteorological Society*, 79, 39-60
- Miller C. (2001). Scientific Internationalism in American Foreign Policy : The Case of Meteorology, 1947-1958, in *Changing the atmosphere : expert knowledge and environmental governance*, ed. Clark Miller and P. Edwards, MIT Press
- Morgan M. S., Morrison M. (eds) (1999). *Models as mediators : perspective on natural and social science*, Cambridge : Cambridge University Press
- Nebeker F. (1995). *Calculating the weather : meteorology in the 20th century*. International Geophysics Series Vol 60, Academic Press
- Norton S., Suppe F. (2001). Why Atmospheric Modeling Is Good Science, in Miller C. and Edwards P. (eds). *Changing the atmosphere : expert knowledge and environmental governance*, Cambridge : MIT Press
- Oreskes N. et al. (1994). Verification, validation, and confirmation of numerical models in the earth sciences, *Science*, vol. 263, p.641-646
- Pickering A. (1995). *The Mangle of Practice. Time, Agency and Science*. University of Chicago Press
- Pielke R. Jr. (2002). Policy, politics and perspective, *Nature* 416, 368
- Randall, D.A. et al. (2007). Climate Models and Their Evaluation. In Solomon et al. (eds) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC*. Cambridge : Cambridge University Press..
- Rochas M. and Javelle J.-P. (1993). *La météorologie*. Paris : éditions Syros
- Sarewitz D., (2000). « Science and environmental policy : an excess of objectivity » in R. Frodeman (ed.), *Earth matters : the Earth Sciences, Philosophy and the Claims of Community*, pp. 79-98
- Shackley S., Wynne B. (1996). Representing uncertainty in global climate change and policy : Boundary ordering devices and authority, *Science, Technology and Human Values*, 21, 3
- Shackley S. et al. (1999). Adjusting to policy expectations in climate change modeling : an interdisciplinary study of flux adjustments in coupled atmosphere-ocean general circulation models, *MIT joint program on the science and policy of climate change*, report n°48, mai 1999
- Shackley S. (2001). Epistemic Lifestyles in Climate Change Modeling, in Edwards P., *Changing the atmosphere : expert knowledge and environmental governance*, ed. Clark Miller and P. Edwards, MIT Press
- Sismondo S. (1999). Models, simulations and their objects, *Science in context*, 12, 2, 247-260
- Stengers I. (1993). *L'invention des sciences modernes*. Paris : La Découverte

- Sundberg M. (2007). Parameterizations as Boundary Objects on the Climate Arena, *Social Studies of Science* ; 37; 473-488
- Van der Sluijs J. *et al.* (1998). Anchoring devices in science for policy : the case of consensus around climate sensitivity, *Social Studies of Science*, vol 28 n°2
- Varenne F. (2003). La simulation conçue comme expérience concrète, in Müller J.P. (éd.), *Le statut épistémologique de la simulation*, éditions de l'ENST, 299-313
- Varenne F. (2007). *Du modèle à la simulation informatique*, Paris : Vrin
- Winsberg E. (1999). Sanctioning models : the epistemology of simulation, *Science in context*, 12, 2
- Winsberg E. (2003). Simulated experiments : methodology for a virtual world, *Philosophy of science*, 70 : 105-125