Story telling with Code – Archaeology of Climate Modelling

Ethnography of Coding, workshop, University of Lancaster, 30.-31.3.2006

Gabriele Gramelsberger, Freie Universität Berlin, Institut für Philosophie,
Habelschwerdter Allee 30, D-14195 Berlin, http://www.userpage.fu-berlin.de/~gab
gab@zedat.fu-berlin.de

Exposing the roots of recent earth system models reveals layers of 'ancient' code starting with Jule Charneys programming of ENIAC in 1950 - one of the first scientific computer based simulations. The story which had been told about the planet earth in those years was simple but it is still the core of recent climate modelling (Bjerknes system of equations). During the decades the complexity of the story increases corresponding to the growing performance of computers. Amazingly, the recent climate models fit conveniently into the pockets of the researchers, in spite of their growing complexity and their need of computer power. However, they do account for considerable scientific arguments, political interest and large amounts of research funding. Computer based simulations not only describe (theory) but also enact (laboratory) their objects of research, e.g. the planet earth. This twofold epistemological function of computational science characterises the transformations in science towards "story telling", triggering a vigorous debate about problems of evaluation, validation and credibility. This paper outlines the practise of coding within the field of climate modelling and proposes to use the coded climate models as a source of applied scientific knowledge for science researchers.

Keywords: Computer based simulation, climate research, forecasts, FORTRAN, modelling

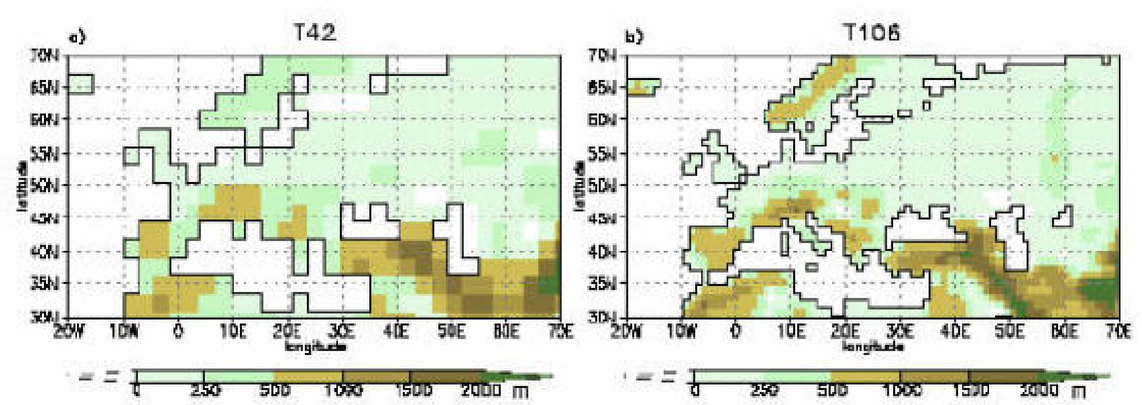
1. Instructive Surprises in the World of Climate Modellers

As a science philosopher one usually theoretises about scientific methods of knowledge production and world making and as a science researcher one usually analyses the social interactions between researchers and their ways of knowledge production. But after years of studying the influence of computer based simulation on scientific knowledge production the object of desire became an irresistible force of attraction. One wants to see “the model” with ones own eyes, wants to risk a view into the deep space of code.

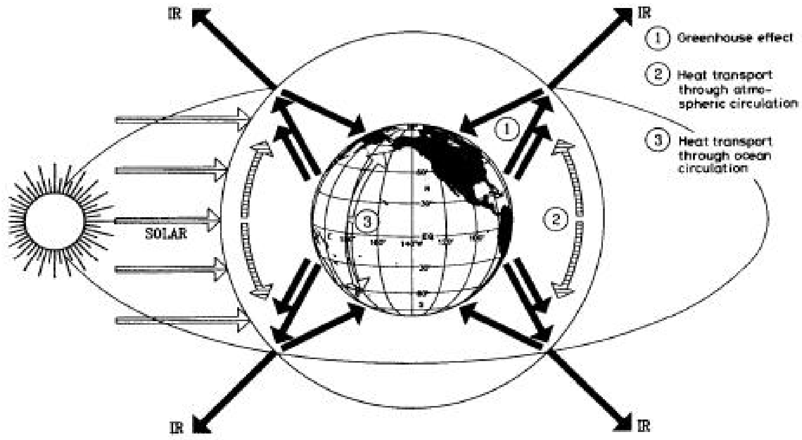
When I first came in touch with climate models there were three things which unexpectedly intrigued me immediately. The first surprise was that these models - which predict global climate changes, man's imminent decline or, at least, his uncomfortable future - fit conveniently into my pocket: A complete ocean with fishes, the atmosphere with clouds and the ability for turbulent behaviour, countrysides with vegetation, lakes, land and sea ice, the anthroposphere and a lot more existed in my handbag when I left the institute of meteorology. Of course this world was a completely semiotical one enclosed in the memory of my USB stick, like the world in a nutshell. But, provided that I had a supercomputer and the necessary batch files, I had developed scenarios of global disasters and hopes.

The second surprise was that these models seemed to be living organisms, proliferating for decades and outliving more than two generations of climate researchers. Starting with Jule Charneys simple model written for ENIAC in 1950, directly plugged into the calculator, these models have evolved today into complex and powerful earth systems around the core of Vilhem Bjerknes equations from 1904. They fit on my USB stick but they need the vastest machines ever built if one wants them to perform.

The third surprise was that these models running on the vastest machines like the Japanese earth simulator unveil a perforated world with holes as big as Luxemburg and time gaps of hours and days. Imagine living in such a discrete world, it must be a strange feeling not only experiencing a discrete existence, but also an average valued one.



[1] Europe grid resolution T42 for long-run global climate models and T106 for regional climate models or short-run global climate models and weather prediction models. (T42 = 12 288 volume elements per layer)



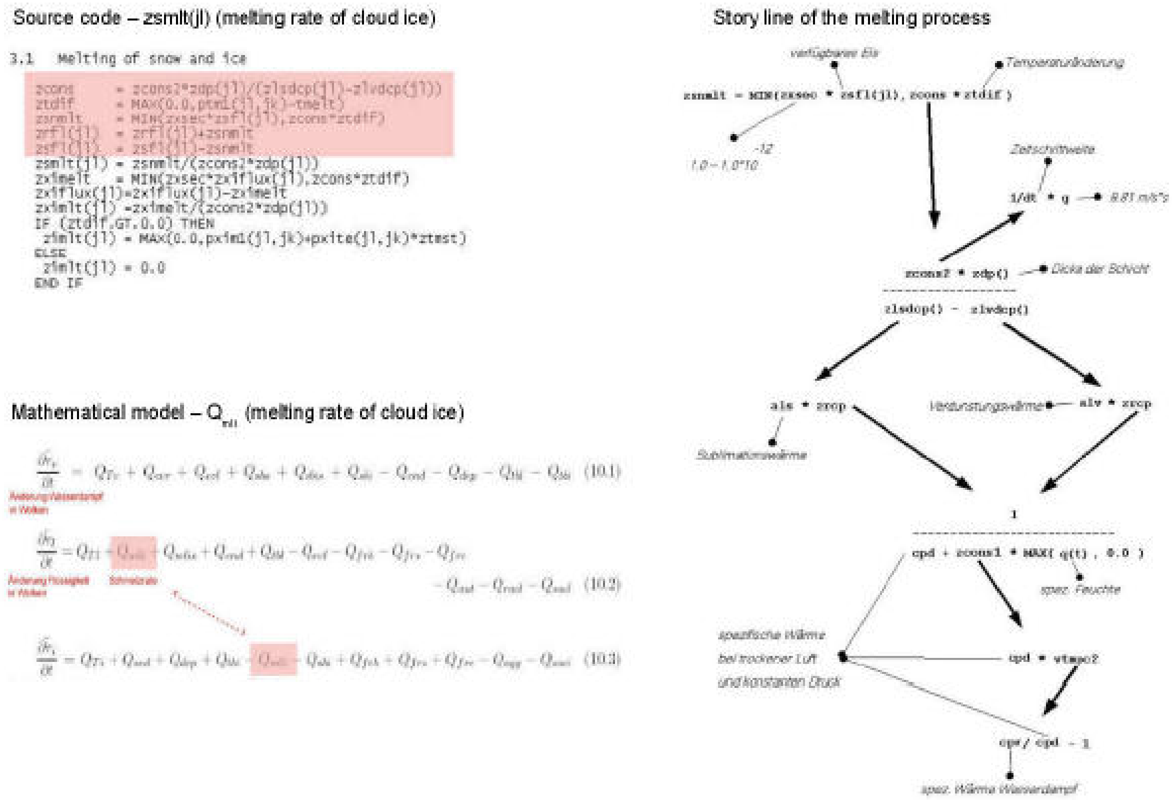
[2] Scheme of the energy fluxes determining the energy balance of the global climate system

2. Story Telling with Coded Models

Reflecting on the unexpected insights, the concept of science as a story telling business became increasingly important in my theory of simulation based sciences. Science always has this highly abstract and objective notion in our imagination. Scientists explore the truth, the reality, and they use cryptic and unbribable tools like complex mathematical equations and transformations. The stories they tell establish a specific class of narration: Enormously successful, totally believed by us and with an audience eager to hear more of these stories. The closer view of one of these stories – the schematic energy balance of the global climate systems, shown in figure 2 – unhides the narrator's position: Listening to a meteorologist mediates the irritating impression that he somehow sojourns at an extraterrestrial place like an astronaut. He is watching the climate processes from the satellite view unlike a weather forecast modeller who always stands on the ground looking into the sky and trying to grasp some information from the clouds. Besides, this quixotic perspective of climate theory shown in the above scheme unveils the typical narration form of a mathematised theory: a symmetric and schematic view of great simpliticity. No people, no skirmish, no horse-trading disturb this meta-positional, schematic and eternal perspective e.g. of Bjerknes theory displayed within a few equations.

Simulation models are based on such narrations but they add a mass of realistic details to the abstract core of Bjerknes equations. They garnish the story of global climate circulation with subscale parameters about fishes in the ocean, clouds in the atmosphere, turbulent processes of air and water, vegetation and topography, land and sea ice, and last but not least with climate relevant aspects of green house gas producing people, environmental skirmishes and horse-trading effects with CO2 emissions. The abstract story of global climate circulation has inappreciably transformed into a story of global climate change containing all the references of a good story: drama, actors, interesting locations and an enormous catastrophic potential.[1] Within the scientific field of computer based simulation a basic equation (1) rules the form of narration:

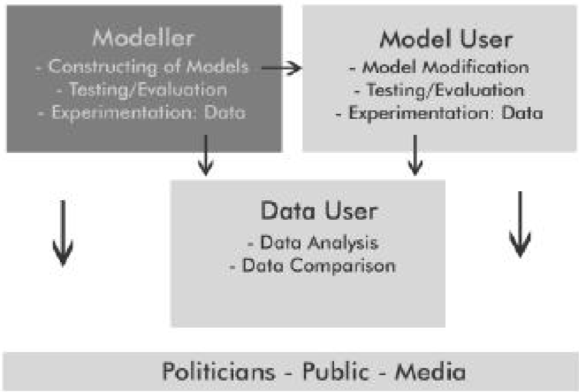
```
Theory = Mathematics = Code (f90) = Story (1)
```



[3] Story „Melting of snow and ice“ of the atmospheric model ECHAM5 (cloud.f90 file)

[1] Scientific simulation performs a concrete story, e.g. about the climate changes of the planet earth. In difference to pieces of theory, published in books and articles and sometimes contradicting and interfering which each other, a simulation model has to tell a coherent and logically consistent story if it wants to run. In the case of climate modelling thousands of equations have to be processed in an accurate way. Diagnostic tools are being developed to check the accuracy and to uncover mistakes which are – within the narrative terminology – illusive: They give a false story about the world. They tell a lie.

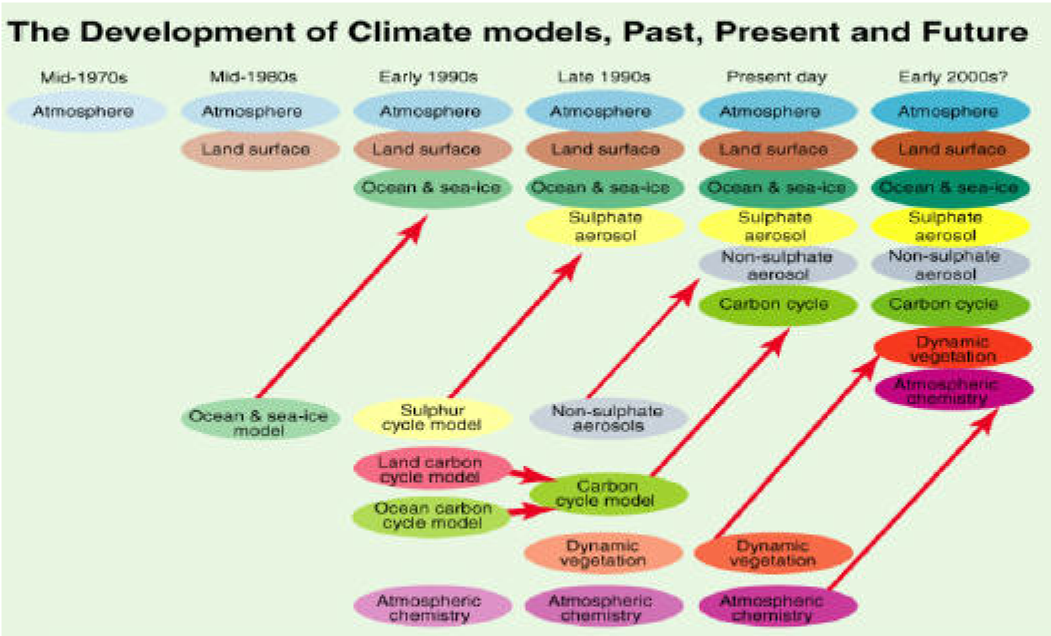
The "melting of snow and ice" – story (fig. 3) is told as follows: 'If there is ice in a cloud and the temperature rises, the density of the ice particles will decrease. When density decreases the heat of evaporation will be discharged and this will increase the humidity of the cloud. Once the barometric pressure of the air reaches a certain threshold, it will rain.' This is one of a thousand little stories within the modern poem of global climate change.[2] The recent problem of the summation of all these tiny stories within the model is: they don't proclaim a happy end for mankind.



[4] Climate research community

The practise of scientific climate story telling follows its own rules and rhythms. Synchronised with the IPCC report publication cycle of five to six years,[3] the international – that is Europe, US and Japan - modeller community creates a new version of the story, new models. To sum up, they need about two years to develop a new version, up to one to two years to evaluate the new version and about two years to run the scenarios for the upcoming IPCC report, short break for recreation, loop. Before IPCC, the development cycle was up to the research teams, dating from the first global climate model written by Syukuro Manabe and Richard T. Wetherald in the early 1970s.[4] Over the years climate research has hoarded miles of code.

[2] The term „poem" does not imply the fiction of climate change or climate research. In fact, 0.6 degree Celsius global warming is proved by measurement and the forecasted 2.0 to 3.5 degree Celsius global warming until 2100 by climate simulation seemed to be rather realistic. The narration terminology has been chosen for a philosophical purpose to understand the strategies and impact of today's computer based science in a better way.
[3] IPCC Intergovernmental Panel on Climate Change: 1990: „Climate Change"; 1995: „Climate Change 1995: The Science of Climate Change"; 2001: „Climate Change 2001: The Science of Climate Basis"; 2007: „Climate Change 2007". http://www.ipcc.ch/pub/reports.htm
[4] Manabe, S./Wetherald, R.T.: The Effects of Doubeling the CO2 Concentration on the Climate of a General Circulation Model, in: Journal of Atmospheric Sciences 32/1975, 3-15



[5] The development of climate models since mid-1970s

3. Code as a Source for Science Researchers – Archaeology of Climate Modelling

One of the most intriguing aspects of the scientific poem of global climate change is its exuberant growth over the years. Looking back, this development can be investigated within the archaeology of climate modelling. Model archaeology has yet to be unearthed. Exposing the roots of recent earth system models unveils layers of 'ancient' code starting with Jule Charneys programming of ENIAC in 1950, one of the first scientific computer based simulations. The story which had been told about the planet earth in those years was simple but it is still the core of recent climate modelling (Bjerknes system of equations). During the decades the complexity of the story increases corresponding to the growing performance of computers. But one can still find thirty years old pieces of code in today's models.

Studying the source code of simulation models will become important to science researchers. Simulation models contain applied knowledge in a different way than written theory addresses knowledge. Simulation models are coded theory and they create an experimental system for performing theory. Science researchers might have a twofold interest, in code and coding: primarily in the development of new knowledge articulated within the code, secondarily in coding simulation models as a new way of knowledge production and a new form of publication.

- *Code*

In coding a simulation model scientists have to formulate numerous assumptions. E.g. the theory tells that plankton stores CO_2. CO_2 is stored long-term when plankton sinks to the oceans bed, but when fish consume plankton it will soon lead back to the CO_2 cycle of the ocean and the atmosphere. Measurement can give selective information about the proportion of long-term and short-term storage in relation to the concentration of plankton and fish. But the simulation code has to be formulated as a definite fish-plankton-relationship valid in a mean form for the complete ocean. The code has to tell where the CO_2 is coming from, how many fishes and which kind of plankton populate the ocean, and so forth. In communication with biologists modellers have to identify the relevant one among thousands of different plankton species. They construct two groups of species: one group absorbing CO_2, the other ignoring it. To initialise the simulation assumptions of the various concentrations, proportions and sources have to be done. The simulation needs parameter values and quantified proportions. This level of concretisation can not be found in written theory published in books and articles. Therefore code is an important source for science researchers if they want to explore the computer based scenarios of computational sciences. Model archaeology will expose layers of ancient code, dated theory, programming languages and software concepts, and old stories.

- *Coding*

Creating a simulation model is a long-term project produced by generations of climate modellers. Climate models are collaborative texts. Software tools help science researcher to discover the working processes. E.g. the graphical diagrams of CVS software allow one to comprehend the changes produced by the modellers. The documentation within the software gives information about authors and references to published theory. The procedure of software releases gives insight in coded theory as new forms of publication with their own platforms and rules. Ethnographic methods – interviews, video documentations, observation - uncover the processes of creating code and theory at the same time. Concisely, studying the coding processes of simulation models gives insights into a new way of scientific knowledge production and a new form of publishing theory.

4. The performative dimension of code

There is another intriguing aspect to coded theory. Science usually proves its stories with experimentation. But doing experiments with the actual climate isn't conceivable. The simulated model not only unveils the story, it is also the test bed to prove it. From an epistemical point of view this is an extraordinarily interesting aspect: Computer based simulations describe (theory) and enact (laboratory) their objects of research, e.g. the planet earth. This twofold epistemological function of computational science characterises the transformations in science towards "story telling", triggering a vigorous debate about problems of evaluation, validation and credibility. And it is not only the planet earth which functions simultaneously as an applied theory and a laboratory for virtual experiments, there are black holes, molecules, new materials and a lot more.

The origin of this new concept of scientific world making - the "new" scientific world explained by theory, produced by code and explored by virtual experiments – is based on the performative dimension of code. Algorithms describe and enact their objects. Dealing with code introduces a creative attitude to science in the manner of telling stories and playing around with possible scenarios, but also of extrapolating theory under virtual circumstances and creating new objects, technologies and realities.

[1] Europe grid resolution: Sillman, J.: Regionale Klimamodellierung – Wetterlagenklassifikation auf Basis des globalen Atmosphärenmodells ECHAM, Diplomarbeit TU Bergakademie Freiberg, 2003

[2] Scheme of the energy fluxes determining the energy balance of the global climate system: Siebert, H. (ed.): Environmental Scarity: The International Dimension, Tübingen 1991

[3] Story „Melting of snow and ice" of the atmospheric model ECHAM5 (cloud.f90 file): Gramelsberger, G.: Computerexperimente in der Klimaforschung, research report, Freie Universität Berlin/BBAW Berlin, upcoming 2006

[4] Climate research community: Gramelsberger, G., upcoming 2006

[5] The development of climate models since mid-1970s: Climate Change 2001: The Science of Climate Basis, IPCC Report 2001

Gramelsberger, G.: Computer Based Simulations – New Tools of Scientific Knowledge Production, Freie Universität Berlin/Science Policy Studies Initiative: Berlin-Brandenburgische Akademie der Wissenschaften (research project, 2004 - 2007), http://www.sciencepolicystudies.de/projekt/computersimulationen/index.htm

Röckner, E. et al.: ECHAM 5 Modellreport, MPI Rep. 349, Max Planck Institute of Meteorology, Hamburg 2003
http://www.mpimet.mpg.de/en/extra/models/echam/echam5.php