THE CONTROVERSIAL STATUS OF SIMULATIONS

Günter Küppers and Johannes Lenhard
Institute for Science and Technology Research (IWT)
University of Bielefeld
PB 100 131
33501 Bielefeld, Germany
E-mail: johannes.lenhard@uni-bielefeld.de

KEYWORDS

Epistemology, modelling, imitation, simulation, validation.

ABSTRACT

It is claimed that simulations maintain an independent status in knowledge production – they combine traits both of experiment and theory. In particular, simulation procedures differ fundamentally from numerical calculations in the strict sense, rather they can be seen as imitations of complex dynamics by a suitable generative mechanism. To argue for this claim results from a case study of climate research are employed. The simulation models of the general circulation of the atmosphere are based on theoretical models (PDEs), but they cannot be derived from the theory. The case study will show that simulation models have to take certain liberties. Namely, the performance on the computer is more important than the model's derivation and its accuracy of calculation.

INTRODUCTION

Computer simulations are employed in rapidly increasing fields of science and practice. In many cases, there is no other method of investigation at hand, be it that analytical methods break down because of the non-linear behaviour of complex systems, or that experimental methods are too risky or simply too expensive. To mention just some examples out of the broad spectrum where computer simulation are applied: How does a plasma behave under the influence of a magnetic field, how is a car body deformed during an accident, i.e. in a virtual crash test? Or how will the global climate change over the next fifty years? These are only three examples out of a large list of research questions tackled by simulation methods.

Simulations form an important part in the analysis of complex dynamics. This suggests the question: what is the nature of the knowledge produced by simulations? In particular, what can be said about the validity and certainty of simulation-based knowledge? Can questions of that kind be considered in a general manner at all? How can simulations be characterized in epistemic respect?

These questions belong to what can be called the philosophy of simulation. In the literature of the philosophy of science, one can find a controversial discussion. Are simulations merely numerical calculations of mathematical models, supplemented perhaps by more or less instructive visualizations? Or, to the contrary, do simulations assume an original position in the production of new knowledge?

Some philosophical accounts try to give a definition of computer simulations by considering the specific capacities of simulation models. E.g., Paul Humphreys derives his definition from the fact that traditional mathematical methods fail:

”*Working Definition*. A computer simulation is any computer-implemented method for exploring the properties of mathematical models where analytic methods are unavailable.“(Humphreys 1991, 500f.)

Various other accounts focus on further properties of simulation models, like their specific affinities to dynamic models (cf. Hartmann 1996; Brennan in McLeod 1968), or the visualisation of simulation results (cf. Rohrlich 1991, or Hughes 1999).

The dominating opinion, considered by Winsberg (2003) to be the "common view", holds that simulation methods open up new possibilities to calculation. In philosophical respect, however, they would be rather uninteresting. The success story is impressive, but essentially simulations deliver the solutions of mathematical equations by the *brute force* of the electronic computer. For instance, M. Stöckler takes this view. He writes:

”Computer simulations enable tremendous progress on a pragmatic level, as a matter of degree in terms of speed and quantity, but not a revolution in the principles of methodology.“ (Stöckler 2000, 356)

In the following, we should like to defend the opposite view. We shall argue that simulations constitute a new kind of synthesis between experimental and theoretical approaches.

Proceedings 18th European Simulation Multiconference
Graham Horton (c) SCS Europe, 2004
ISBN 3-936150-35-4 (book) / ISBN 3-936150-36-2 (CD)

SECOND ORDER MODELLING

We intend to show that computer simulation rely on procedures that differ fundamentally from numerical calculation in the strict sense, that is, they do not simply replace variables and parameters by numerical values in mathematical formulas. They do not solve complex systems of equations. Rather, simulations are numerical *imitations* of the unknown solution of differential equations, or the imitation of complex dynamics by a suitable generative mechanism.

The Monte Carlo method may serve as an illustrative example. This method goes back to the joint effort of Stanislaw Ulam and John von Neumann in Los Alamos, where they cooperated during the Manhattan project. Monte Carlo may count as the first simulation method (cf. Richtmyer and von Neumann 1946). Imagine that one intends to determine the volume of a certain body via Monte Carlo. One can embed the body into a cube of which the volume is known. The idea is to replace the (unknown) primitive by a ratio that can be determined "empirically", or quasi-empirically, by iterating computer runs. The computer determines a point out of the cube at random, if this point belongs to the body, the trial is said to be successful. By re-iterating this random choice one can determine the unknown volume as the ratio of successful trials out of a great number of trials. To put it in other words: the integration is imitated by a generative mechanism.

In the following, we intend to foster this thesis of simulation as imitation, and we shall argue that simulation models cannot be derived directly from theoretical models like systems of differential equations. To find a convincing argument requires to avoid the easy way, and to choose for example a field where the theoretical approach via partial differential equations (PDEs) seems to determine the simulation model. Let us choose climate research as a case study, and have a closer look at the so-called general circulation models. Even here, simulations can be characterized as imitations of complex processes by generative mechanisms. Admittedly, these mechanism are guided by theory, but not determined by theory. At the end, we shall argue that the performance of the model on the computer is more important than its theoretical derivation and its accuracy of calculation.

Imitation of a Solution: the Dynamics of the Atmosphere

In 1955, Norman Phillips, working at Princeton's Institute for Advanced Studies, succeeded with the so-called *first experiment* in simulating the dynamics of the atmosphere, i.e. in reproducing the patterns of wind and pressure of the entire atmosphere in a computer model. (Phillips 1956. For more details of the experiment, cf. Lewis 1998, for a broader history of ideas of the modelling of the general circulation of the atmosphere, cf. Lorenz 1967.) This development of a simulation model of the general circulation of the atmosphere was celebrated as a major breakthrough. It surprised the experts, because it was generally accepted that a theoretical modelling approach would hardly be possible – the complexity of the processes that determine atmospheric circulation was judged insurmountable for an approach via a simple model.

After a little while, the simulation-based approach of such general circulation models (GCMs) became the silver bullet of climate research. In 1960 already, the *Geophysical Fluid Dynamics Laboratory* (GFDL) was founded in Princeton to follow up on this approach. This was the first institution with the official task to simulate in climate research. (The National Center for Atmospheric Research (NCAR) in Boulder, Colorado was founded in the same year.) Today, GCMs dominate climate research, being developed further at a handful of research institutions worldwide.

This first attempt to build a simulation model of the entire atmosphere was considered an "experiment". This underlines how uncertain the success of this project was. At the same time, the conception of "experiment" expresses an important aspect for the methodology: in simulations, scientists use their models like an experimental set-up. Hence, the results of simulations acquire a quasi-empirical character.

The simulation model of the "first experiment" worked with a very coarse spatial discretization of the atmosphere. In vertical direction, it exhibited only two layers, and horizontally the grid cells covered more than 200.000 km^2. In the experiment, the initial state was an atmosphere at rest, with no differences in temperature and no flow. Phillips had to introduce the physical laws that govern the dynamics of the atmosphere. He used only six basic equations (PDEs) which since then are called the "primitive equations". They are generally conceived of as the physical basis of climatology. These equations express well-known laws of hydrodynamics – the surprising thing was that only six PDEs were sufficient to reproduce the complex behaviour, and Phillips had the skill and luck of making an adequate choice.

The physical basis had to be adapted to the grid. The construction of a discrete model is a typical task of simulation modelling. The global and continuous equations of hydrodynamics had to be reformulated in order to calculate the time evolution of the relevant variables – pressure, temperature, wind speed – at the grid nodes step by step.

In the second step of the experiment, the dynamics was started, i.e. the radiation of the sun and the rotation of the earth were added. The atmosphere settled down in a so-called *steady state* that corresponded to stable flow patterns.

The tantalizing question was whether the model was able to reproduce the global flow patterns of the real atmosphere well-known from observations. For instance, one criterion was the complex pattern of the so-called surface westerlies, winds blowing continuously north the equator. The result was positive – everyone was impressed by the degree of correspondence. As we mentioned before, the experts were sceptical about the possibility of a global (and not far too complicated) model, but the empirical success was convincing. The decisive criterion for success was the adequate imitation of the phenomena, i.e. the flow patterns, not the derivation from theoretical principles.

The continuous primitive equations of the atmosphere were not solved in the strict sense during Phillips' experiment, rather, the phenomena of the atmosphere were imitated by the generative mechanism of the discrete difference equations. The success of the imitation was judged by its correspondence to the flow patterns observed. Insofar, the validation of simulation results relies on a quasi-empirical strategy.

There is a certain condition that simulations have to fulfil, a condition that you will certainly be well-acquainted with, namely stability. The generative mechanism chosen that is to imitate a certain dynamics, must "run" on the computer. It must not become instable because of discretization errors and truncation errors building up. Numerical instabilities are a severe and fundamental problem of simulation modelling.

The simulations of climate were caught up in this problem as well. In the course of the story of climate simulations, we will hit on a remarkable fact: The orientation to simulation as the numerical calculation of a solution led climate research to a dead end. Further progress became possible only after the insight had been accepted that in simulation modelling something more was allowed and required than strict dependency from the given system of equations to be solved. In particular, the discrete model cannot be deduced from the theoretical model (the primitive equations)! To justify this claim will be the strongest argument for our viewpoint that simulations are imitations and can be described as modelling of second order.

Arakawa's trick

Phillips simulation experiment was a tremendous success, but it exhibited also an important failure of the simulation model: the dynamics of the atmosphere were stable only for a few weeks. After about four weeks, the internal energy blew up, and the system "exploded" – the stable flow patterns dissolved into chaos. „After 26 days, the field … became very irregular owing to large truncation errors, and is therefore not shown." (Phillips 1956, 145)

Nevertheless, the experiment was seen in general as a success. The possibility of simulating the atmospheric circulation was not doubted. Instead, it was acknowledged as a challenge for further research to achieve a stable model. Questions of stability were important for climate research, as it was interested in long-time predictions. And they were of equal importance for the simulation method in general, because it is a rather general and typical task of replacing the "natural" dyamics of a system of PDEs by the "artificial" dynamics of a discrete system in a stable way. Phillips was well aware of the superior importance of stability issues and he suggested the truncation errors as the cause of instability.

Years of intensive and highly competitive research followed. The solution of the problem was assumed to consist in adequate smoothing procedures to cancel out the errors before they could blow up. This strategy was obviously oriented at the ideal of calculating as correctly as possible. Instabilities were seen as resulting from errors, inaccurate deviations of the discrete model from the true solution of the continuous s ystem.

The decisive breakthrough, however, was achieved by a different approach, one pursued by Akio Arakawa, a mathematically highly gifted meteorologist who was developing a GCM at the University of California in Los Angeles (UCLA). For him, imitating the dynamics was of prime importance, and less the precise calculation of a solution.

At heart, Arakawa had realised that one could set aside a strict solution of the basic equations. One even *should* do so! If the time development of the simulation was reproducing the patterns of the atmosphere in a sufficiently adequate manner, and if the simulation was stable, then it was not obliged to be a solution of the basic equations – not even in the limit! To put the whole thing in a nutshell, the point of Arakawa's approach was: imitation of the phenomena beats solution of the equations.

Of course, this does not imply that one can simulate a given dynamics by completely arbitrary mechanisms. Far from that. And Arakawa very wisely adhered to the given equations. BUT he applied what later was to be known as "Arakawa's computational trick". The basic equations define a generative mechanism whose development over time is formally described by the Jacobi operator. Arakawa replaced the Jacobian by another operator he himself had constructed, later called Arakawa-Jacobian. The construction of the Arakawa operator is full of highly sophisticated mathematical arguments. The details do not matter in our context here (cf. Arakawa 1966, and the reconstruction in Arakawa 2000.) The pivotal fact is that the Arakawa operator permitted a stable long-time integration because he avoided the non-linear instability Phillips had had to face. Arakawa was able to prove this property mathematically.

To guarantee the stability of the simulation procedure, Arakawa had to introduce further assumptions, partly contradicting experience and the physical theory. He had to assume that the kinetic energy in the atmosphere would be preserved. This is definitely not the case in reality, where part of this energy is transformed into heat by friction. Moreover, dissipation is presumably an important factor for the stability of the real atmosphere. So we can summarize that Arakawa, in assuming the preservation of kinetic energy, "artificially" limited the blow-up of instabilities. In the real atmosphere, friction is responsible for that effect.

Incidentally, John von Neumann had employed a very similar strategy. He introduced an "artificial viscosity" to bring about a realistic behaviour while simulating the propagation of "shock waves" (cf. Winsberg 2003).

Utilizing conservation assumptions that obviously went against theory and experience was taken up very sceptical by the community. As Arakawa remembered, the tenor was: "Why require conservation while nature does not conserve?" (Arakawa 2000, 16) While most researchers were convinced that the promising path was to find a solution of the primitive equations as accurate as possible, Arakawa had made an additional modelling step. This step was not derived from the theoretical basis, and was justified only by the results of simulation runs that showed quasi-empirically that the Arakawa operator led to a successful imitation. The success of his approach was eventually generally accepted, and his initially controversial approach, is today conceived of as a "computational trick" (cf. Weart 2001).

It is an illustrative fact that Arakawas approach proved to be superior in the course of a simulation experiment. At first, in the 1960ies, Arakawa had a couple of scientific rivals that tried to develop different smoothing procedures to avoid the problem of non-linear instabilities. Again, an experiment brought about a decision. In 1978, Jule Charney conducted a simulation experiment that consisted in the competition of different GCMs. They ran in parallel, starting with the same initial conditions. Three GCMs took part: Leith/ Lawrence Livermore Lab; Smagorinsky/ GFDL Princeton und Arakawa/ UCLA. The first had implemented smoothing procedures, while the UCLA-model was based on the Arakawa-operator. Phillips describes the outcome:

"„...three general circulation models (...) were used for parallel integrations of several weeks to determine the growth of small initial errors. Only Arakawa's model had the aperiodic behavior typical of the real atmosphere in extratropical latitudes, and his results were therefore used as a guide to predictability of the real atmosphere. This aperiodic behavior was possible because Arakawa's numerical system did not require the significant smoothing required by the other models, and it realistically represented the nonlinear transport of kinetic energy and vorticity in wave number space." (Phillips 2000, xxix)

This strengthens our argument: The traditional view interpreted simulation as numerical calculation of the solution of a mathematically or physically defined system. In the face of instabilities, this view had come to a dead end. The desired stability could be obtained only by smoothing procedures. These, in turn, resulted in unrealistic long-term behaviour. The competing approach of Arakawa attained a more "realistic" picture of the dynamics of the atmosphere, because it was based on counter-intuitive, artificial and physically unmotivated assumptions. In other words: he utilised the partial autonomy of simulation modelling as a modelling of second order.

CONCLUSION

Let us take stock of the case study. We have seen that there is no relation that would permit to derive a simulation model from a theoretical, or from a mathematical model. Simulation modelling can– and must – take certain liberties. In short: performance beats theoretical accuracy.

Because of this partial autonomy of the simulations from the theoretical basis, simulations are not merely numerical calculations. Rather, they are models of second order in the sense of iterated model construction. This result does not seem to be restricted to simulation models that stem from a system of continuous PDEs. They constituted the extreme case where simulation as calculation seemed to be most plausible. Other simulation procedures employ generative mechanisms in a similar way to imitate a certain dynamic.

Therefore, the adequacy of a simulation model cannot be theoretically deduced, nor derived from general principles. Simulation results have to be judged by experience. The quasi-empirical approach, that permits to tune the models in face of theoretical model experiments, is a necessary methodological condition for simulations of that kind. Theoretical alternatives can be compared empirically. Hence, it is justified to speak of computer simulations as of experimenting with theories. This provides simulations an independent status in knowledge production, they combine traits both of experiment and theory. (For similar theses concerning the autonomous status of simulations cf. Rohrlich 1991, Humphreys 1991, Galison 1996.)

This seems to be a result that forms (at least) an important part of an epistemic characterization of simulations, and of the knowledge they provide. What does this imply for our initial question concerning the validity of simulation results?

There, however, one evil seems to have been replaced by another. Simulation methods can achieve a good performance without strict theoretical justification. But then the results have to be validated by experience. Does this not fundamentally weaken the potential for predictions?

We maintain that no easy answer is possible. One may observe interesting new strategies of validation, e.g. in climate research. There the models are validated by comparing their predictions in retrospect with known data. One has to trust, however, in the stability of the dynamics to do so. In general, this is a challenge for further research: to make a survey of the pragmatic solutions that can be found in science to cope with the problem of validating this new kind of simulation knowledge.

REFERENCES

Arakawa, A. 1966. "Computational Design for Long-Term Numerical Integration of the Equations of Fluid Motion: Two-Dimensional Incompressible Flow. Part I." *Journal of Computational Physics* 1, 119-143.

Arakawa, A. 2000. "A Personal Perspective on the Early Years of General Circulation Modeling at UCLA." In *General Circulation Model Development,* D. A. Randall (ed.). Academic Press, San Diego, 1-66.

Brennan, R. D. 1968. « Simulation is Wha-a-t? Part II." In *Simulation, The Dynamic Modeling of Ideas and Systems with Computers,* J. McLeod (ed.). McGraw-Hill, New York, 5-12.

Galison, P. 1996. "Computer Simulations and the Trading Zone". In *The Disunity of Science: Boundaries, Contexts, and Power,* P. Galison and D. J. Stump (eds.). Stanford University Press, Stanford, California, 118-157.

Hartmann, S. 1996. "The World as a Process. Simulations in the Natural and Social Sciences." In *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View*, R. Hegselmann (ed.). Kluwer, Dordrecht, 77-100.

Hughes, R. I. G. 1999. "The Ising Model, Computer Simulation, and Universal Physics." In *Models as Mediators*, M. S. Morgan and M. Morrison (eds.). Cambridge University Press, Cambridge, 97-145.

Humphreys, P. 1991. "Computer Simulations". In *PSA 1990*, Vol. 2, F. Fine and R. Wessels (eds.). Philosophy of Science Association, East Lansing, 497-506.

Lewis, J. M. 1998. "Clarifying the Dynamics of the General Circulation: Phillips's 1956 Experiment." *Bulletin of the American Meteorological Society* 79, No. 1, 39-60.

Lorenz, E. 1967. "The Nature of the Theory of the General circulation of the Atmosphere". Technical Paper No. 218, World Meteorological Organization WMO, Geneva, 115-161.

Phillips, N. 1956. "The General circulation of the Atmosphere: a Numerical Experiment. *Quaterly Journal of the Royal Meteorological Society* 82, No. 352, 123-164.

Phillips, N. 2000. "Foreword". In *General Circulation Model Development*, D. A. Randall (ed.). Academic Press, San Diego, xxvii-xxix.

Rohrlich, F. 1991. „Computer Simulation in the Physical Sciences In *PSA 1990*, Vol. 2, F. Fine and R. Wessels (eds.). Philosophy of Science Association, East Lansing, 507-518.

Stöckler, M. 2000. "On Modeling and Simulations as Instruments for the Study of Complex Systems". In *Science at Century's End*, M. Carrier, G. J. Massey and L. Ruetsche (eds.). University of Pittsburgh Press, Pittsburgh, 355-373.

Weart, S. 2001. "Arakawa's Computation Trick", American Institute of Physics. www.aip.org/history/climate/arakawa.htm.

Winsberg, E. 2003. "Simulated Experiments: Methodology for a Virtual World." *Philosophy of Science* 70, 105-125.

AUTHOR BIOGRAPHIES

GÜNTER KÜPPERS was educated in physics and did his dissertation in theoretical physics on problems of pattern formation in hydrodynamic convection flows 1969, University of Munich. Habilitation in Science Studies 1991, University of Vienna. 1966 – 1974 research fellow at the Max-Planck-Institute for Plasma-Physics, Garching. From 1974 – 1994 Director of the University Centre for Science Studies. Since 1994 member of the board of directors of the Institute for Science and Technology Studies (IWT) at the Bielefeld University. His fields of specialisation are: systems-theory and theory of self-organisation, dynamics of social systems, innovation networks and knowledge society.

JOHANNES LENHARD was educated in mathematics and philosophy at the universities of Heidelberg and Frankfurt. 1993–1998 assistant professor at the University of Frankfurt, in 1998, he did his dissertation in probability theory (population genetics). After an intermezzo as mathematical consultant he changed to the University of Bielefeld. Currently he is conducting a research project on the method of computer simulations (together with Günter Küppers) that is funded by the VW-Stiftung. His fields of interest include: philosophy of science, science studies, epistemology of mathematics.