Climate Change Research: Evaluation and Policy Implications
S. Zwerver, R.S.A.R. van Rompaey, M.T.J. Kok and M.M. Berk, (Eds.)

Uncertainty management in integrated modelling, the IMAGE case

Jeroen P. van der Sluijs

Department of Science Technology and Society, Utrecht University, Padualaan 14, NL-3584 CH Utrecht, The Netherlands (E-mail: sluys@chem.ruu.nl)

Abstract

Integrated assessment models of global environmental problems play an increasingly important role in decision making. This use demands a good insight regarding the reliability of these models. In this paper we analyze uncertainty management in the IMAGE-project (Integrated Model to Assess the Greenhouse Effect). We use a classification scheme comprising type and source of uncertainty. Our analysis shows reliability analysis as main area for improvement. We briefly review a recently developed methodology, NUSAP, that systematically addresses the strength of data in terms of spread, reliability and scientific status (pedigree) of information. This approach is being tested through interviews with model builders.

1. INTRODUCTION

Facilitated by developments in computer technology, "integrated modelling" emerges in the mid eighties as a new approach to interface science and policy concerning complex environmental issues. For instance, in the 1986 RIVM annual report vice director general of the ministry of VROM Kees Zoeteman argues that RIVM can give optimal shape to its role as interface between science, policy and monitoring by means of integrated modelling. To give an other example: the RAINS model (Regional Acidification INformation and Simulation), developed in the eighties at the International Institute for Applied System Analysis (IIASA), was used in the international acid deposition negotiations and became an annex of the SO_2-protocol. For the climate problem there are at least 16 different integrated assessment models in active use or under active development (Weyant, 1994).

In this paper we focus on the IMAGE model, developed at RIVM, which started as a pioneer in this field. IMAGE has been used for scenario calculations in the influential policy document "Zorgen voor Morgen" (Concern for Tomorrow) (Langeweg, 1988) and for the development of emission scenario's for IPCC working groups II and III, the latter in combination with the Atmospheric Stabilization Framework of the US Environmental Protection Agency (Swart, 1994a). The revised version of the model, IMAGE 2, is currently being used in all three working groups of IPCC. For the future it is likely that integrated models will be used to support the negotiations during the United Nations Conference of Parties to the Climate Convention in Berlin, March, 1995, and follow up (Swart, 1994a; Alcamo, 1994b).

The inherent uncertain character of the knowledge on future climate and the poor scientific understanding of the geosphere biosphere system combined with the high decision stakes associated with policy choices supported by these models, demand a good insight in and a high awareness of the quality, the reliability and the limitations of the model. According to Swart (1994a), the uncertainties in the knowledge of climate change are so large that climate scientists sometimes claim that an integrated approach is not meaningful in the case of climate change. The model builders themselves oppose this view by stating that "*while there is a great uncertainty regarding our future, we have a certain responsibility to take our best scientific understanding, and use that to develop reasonable policies*" (Alcamo, 1994b). Such controversies imply that the question of uncertainty management is becoming increasingly important. In this study we investigate for the IMAGE case how scientific uncertainties are being managed and we analyze the scope of the current practice of uncertainty management, using a two dimensional classification scheme comprising type and source of uncertainty.

2. UNCERTAINTY MANAGEMENT IN THE IMAGE PROJECT

On the basis of literature study and interviews with the model builders, we have made an inventory of the way questions of uncertainty and quality have been and are being addressed in the IMAGE project. The first version of IMAGE was developed in the period 1985-1990 (De Boois and Rotmans, 1986; Rotmans 1990). In this period climate change started being signalized as a policy issue. IMAGE was even used to put the issue on the policy agenda and the model was developed despite initial lack of interest in such a model of policy makers (Rotmans, 1994; Swart, 1994b). IMAGE is designed as a deterministic model. The treatment of uncertainty has not been explicitly considered in the design of the model (Weyant, 1994). This might be because the issue of uncertainty management was less urgent at that time than it is nowadays. Despite these circumstances, sensitivity analysis and uncertainty analysis have been carried out from the very beginning.

In the beginning, IMAGE suffered a lot of criticism from scientists, who thought that the approach was far too simplified (Rotmans, 1994). The scientific status of IMAGE has increased significantly since the IMAGE team started to publish their work in high impact scientific journals (e.g. Science Citation Index, 1991) such as *Climatic Change* (Rotmans, 1994). In 1992 Joe Alcamo became project leader of IMAGE, bringing the experience from the development of the RAINS model. Since then several changes have become visible: a complete revision of the model; improved communication with policy makers and scientists and improved exposure to peer review (e.g. the complete IMAGE 2.0 model was published as a special issue of the journal "Water Air and Soil Pollution" (Alcamo, 1994a)).

Thorough sensitivity analyses and uncertainty analyses have been carried out for several sub-models of IMAGE 2 (e.g. Krol and van der Woerd, 1994). The technique used is Latin Hypercube Sampling, which is an advanced Monte Carlo based tool to map the relative contribution of specified uncertainty sources to the spread in model output and to assess the propagation of uncertainties through the model. A special software package for this purpose, UNCSAM (UNCertainty analysis by Monte Carlo SAMpling techniques) was developed (Janssen *et al.*, 1994). The complete inexactness-uncertainty of the model has not been computed yet. A major problem in determining total inexactness uncertainty is the identification of the spread and distribution functions of all input data and model parameters (Rotmans, 1994). Variables that are highly uncertain (such as population growth) are managed by making them scenario variables.

The completeness and the quality of the IMAGE 2 model have been addressed by two "international review meetings" (Hordijk, 1993). An other strategy to improve quality is to include the best science available (Alcamo, 1994b). This means that the model is often being revised in dialogue with scientists. This is also reflected by the attempts to improve the integration of IMAGE and the NRP-research (National Research Program on Global Air Pollution and Climate Change) (Berk, 1993).

Other sources of error, however, such as numerical artefacts in model calculations and errors due to numeral approximation are not assessed. According to Hordijk (1994), who agrees that this is an omission, even the NRP is not interested in the mathematical tour de force required to assess these errors.

3. A CLASSIFICATION SCHEME FOR UNCERTAINTY

According to its *source*, scientific uncertainty can be classified as: (1) *data uncertainties* that arise from the quality or appropriateness of the data used as inputs to models; (2) *modelling uncertainties* that arise from: (a) incomplete understanding of the modelled phenomena or (b) numeral approximations used in mathematical representation; and (3) *completeness uncertainties* covering all omissions due to lack of knowledge (Vesely and Rasmuson, 1984). Funtowicz and Ravetz (1990) have given a classification of *types* of uncertainty: (i) *inexactness* (significant digits/error bars); (ii) *unreliability*; (iii) *border with ignorance.* The combination of both classifications produces a two dimensional classification scheme defining areas to be addressed in uncertainty management in integrated models. We have applied this scheme to the findings of the previous section to indicate the scope of the current practice of uncertainty management in IMAGE (Table 1). The table shows that the best covered area are inexactness-uncertainties in input data and parameters. The reliability of the input data and model structure are not systematically addressed. Other omissions are that possible numerical artefacts in model calculations and errors due to numeral approximation are not assessed. This area can be addressed by a

Table 1
Two dimensional classification scheme for uncertainty, showing the scope of uncertainty analysis efforts on IMAGE

source / type	input data	model structure		model completeness
		parameters	relations	
inexactness	systematically addressed for some sub models	systematically addressed for some sub models	not addressed	Explicitly addressed by advisory board
unreliability	(passively via peer review of publications)	(passively via peer review of publications)	Explicitly addressed by advisory board; (passively via peer review of publications)	Explicitly addressed by advisory board
ignorance	(passively via "include best science available")	(passively via "include best science available")	(passively via "include best science available")	(passively via "include best science available")

mathematical tour de force which lies beyond the scope of this paper. Uncertainty resulting from ignorance is not addressed too. Ignorance and completeness uncertainties are the most difficult to address. In fact they can only be addressed indirectly via quality control procedures, such as peer review, of the production process of the scientific information used in the model.

4. NUSAP AND THE PEDIGREE MATRIX

Funtowicz and Ravetz (1990) have designed an innovative tool to address the issue of uncertainty and reliability: the NUSAP (Numeral Unit Spread Assessment Pedigree) notational scheme for scientific information. NUSAP is designed to act as a heuristic for good scientific practice and as a system for expressing and communicating uncertainties. It consists of five qualifiers: **N**umeral, **U**nit, **S**pread, **A**ssessment and **P**edigree. The last three qualifiers address the various aspects of uncertainty. The *spread* qualifier conveys an indication on the inexactness of the numeral and unit places. The *assessment* qualifier should express a judgement on the reliability of the three previous qualifiers, it is a measure for the strength of the data. *Pedigree* conveys an evaluative account of the production process of the information, and can be seen as a measure for scientific status of the associated knowledge. Because many aspects are relevant in evaluating the production process, a matrix is used to represent pedigree. Depending on its application a pedigree matrix consists of a set of suitable evaluation criteria (e.g. peer acceptance), and defines modes of these criteria (e.g. low, high) which are coded hierarchically. A pedigree matrix for research is given in Table 2. For instance the CO_2-fertilization parameter in the IMAGE 1 model was based on a theoretically based model, experimental data, faced medium peer acceptance and its value was subject to competing schools, yielding a research pedigree (3,4,2,2). The Second Law in thermodynamics has a pedigree (4,4,4,4). The pedigree matrix enables to compare strength of data in terms of the applied criteria and brings to light the weak parts of the model. This also helps in priority setting for model improvement.

Table 2
The pedigree matrix for research as designed by Funtowicz and Ravetz (1990)

Code	Theoretical Structure	Data-input	Peer acceptance	Colleague consensus
4	Established theory	Experimental data	Total	All but cranks
3	Theor. based model	Historic/field data	High	All but rebels
2	Computational model	Calculated data	Medium	Competing schools
1	Statistical processing	Educated guesses	Low	Embryonic field
0	Definitions	Uneducated guesses	None	No opinion

Funtowicz and Ravetz also proposed a pedigree matrix for environmental models (Table 3). When tested in our interviews, the hierarchy in the columns of this pedigree matrix proved to be controversial: according to Alcamo (1994b), there is no 'good' or 'bad' in model structure. His alternative ranking of the modes for *data input* and *testing* is given between brackets in Table 3. With respect to model structure it is obvious that a black box model has a lower scientific status (if any at all) than a model that is completely governed by established physical laws and has a high process detail. However, if a very simple meta model is derived from, and secured by, a complex model with high process detail, the simple meta model can be equally good to model the process. For these cases we propose to apply the pedigree matrix to the mother model while taking into account the consequences of the simplifications in the meta model.

It is not surprising that model builders use other criteria to evaluate model quality than Funtowicz and Ravetz do. They have different critical roles (compare Clark and Majone, 1985). The model builder has to fulfil the needs of policy makers without compromising too much the scientific credibility of the model. This results in usefulness as the main quality criterium (Mermet and Hordijk, 1989; Swart, 1994a).

Table 3
The pedigree matrix for environmental models as designed by Funtowicz and Ravetz (1990). Between brackets: alternative ranking codes, attributed by a model-builder.

Code	Model structure	Data input		Testing	
4	Comprehensive	Review	(1)	Corroboration	(2)
3	Finite-element approximation	Historic/field	(1)	Comparison	(1)
2	Transfer function	Experimental	(1)	Uncertainty analysis	(1)
1	Statistical processing	Calculated	(1)	Sensitivity analysis	(1)
0	Definitions	Expert guess	(0)	None	(0)

5. CONCLUSIONS

The analysis of uncertainty management in the IMAGE project shows as main area of improvement the assessment of reliability of the input data and model structure. The NUSAP notational scheme for scientific information as designed by Funtowicz and Ravetz (1990) is a tool to systematically address the issues of reliability and quality. Model builders use other evaluation criteria for model quality than Funtowicz and Ravetz. These differences can be understood from their differing critical role. They are however no obstacle to using the NUSAP methodology as a guiding checklist to identify weak parts of the model in terms of the applied criteria.

6. FURTHER RESEARCH

Further research is needed to identify suitable sets of evaluation criteria for pedigree matrices accommodating different critical roles. An other important area for research is the development of a sensible way to represent and communicate the information on uncertainty, reliability and limitations of a model in a form comprehensible to the users and minimizing ambiguity regarding its interpretation.

7. REFERENCES

J. Alcamo (ed), IMAGE 2.0: Integrated Modeling of Global Climate Change, in: *Water, Air, and Soil Pollution*, **76**, Nos. 1/2, 1994a.

J. Alcamo, *personal communication* (interview) 21 november 1994b.

M. Berk, *Integratie van NOP-onderzoek en het IMAGE model*, Report of a NRP Study-day, Utrecht, 22 January 1993.

W.C. Clark and G. Majone, The Critical Appraisal of Scientific Inquiries with Policy Implications, in: *Science Technology and Human Values*, **10** (3), 6-19, 1985.

H. de Boois and J. Rotmans, Overzichtsmodel van de CO_2-Problematiek, in: *Berichten uit het RIVM 1985*, Bilthoven, 1986.

S.O. Funtowicz and J.R. Ravetz, *Uncertainty and Quality in Science for Policy*, Kluwer, Dordrecht, 1990.

L. Hordijk (ed.), *Report International Review Meeting IMAGE 2.0 11-13 January 1993 Amsterdam*, NRP report 00-09, Bilthoven, 1993.

L. Hordijk, *personal communication* (interview), 15 november 1994.

P.H.M. Janssen, P.S.C. Heuberger and R. Sanders, UNCSAM: a Tool for Automating Sensitivity and Uncertainty Analysis, in: *Environmental Software*, **9**, 1-11, 1994.

M.S. Krol and H.J. van der Woerd, Uncertainty Analysis for the Computation of Greenhouse Gas Concentrations in IMAGE, in: J. Grasman and G. van Straten (eds), *Predictability and Nonlinear Modelling in Natural Sciences and Economics*, Kluwer, Dordrecht, 1994.

F. Langeweg, *Zorgen voor Morgen, Nationale Milieuverkenningen 1985-2010*, Samson H.D. Tjeenk Willink, Alphen aan den Rijn, 1988.

L. Mermet and L. Hordijk, On Getting Simulation Models Used in International Negotiations: a Debriefing Exercise, in: F. Mautner-Markhof (ed), *Processes of International Negotiations*, Westview Press, Boulder CO, 205-237, 1989.

J. Rotmans, *IMAGE An Integrated Model to Assess the Greenhouse Effect*, (Thesis), Rijksuniversiteit Limburg, 1990.

J. Rotmans, *personal communication* (interview), 9 november 1994.

Science Citation Index, *Journal Citation Reports, A Bibliometric Analysis of Science Journals in the ISI database*, 1991.

R. Swart, *Climate Change: Managing the Risks*, (Thesis), Free University Amsterdam, 1994a.

R. Swart, *personal communication* (interview), 23 november 1994b.

W.E. Vesely and D.M. Rasmuson, Uncertainties in Nuclear Probabilistic Risk Analyses, in: *Risk Analysis*, **4**, 313-322, 1984.

J.P. Weyant, *Integrated Assessment of Climate Change: A Overview and Comparison of Modeling Approaches*, paper for the writing team 6/7 of working III Intergovernmental Panel on Climate Change Lead Authors Meeting, Geneva, September 8-10 1994.

B.C.J. Zoeteman, Verslag van de Directeur van Hoofdsector III: Chemie en Fysica, in: *Berichten uit het RIVM 1986*, Bilthoven, 1987.