Integrated Assessment Models and the Management of Uncertainties

Jeroen van der Sluijs

WP-96-119

October 1996

International Institute for Applied Systems Analysis

Laxenburg, Austria

ABOUT THE AUTHOR

Jeroen van der Sluijs is Ph.D.-candidate at the Department of Science Technology and Society of Utrecht University, Padualaan 14, 3584 CH Utrecht, The Netherlands (E-mail: J.vanderSluijs@nwsmail.chem.ruu.nl). His forthcoming thesis will be a critical analysis of the assessment practice of the climate issue, with a special focus on the question of uncertainty management in science for policy. He participated in the 1996 IIASA Young Scientists Summer Program.

TABLE OF CONTENTS

LIST OF ACRONYMS

AIM	Asian-Pacific Integrated Model
ASF	Atmospheric Stabilization Framework
ASTM	American Society for Testing and Materials
CETA	Carbon Emission Trajectory Assessment
CSERGE	Center for Social and Economic Research into the Global Environment
CFCs	Chlorofluorocarbons
CH_3CCl_3	Methychloroform
CH_4	Methane
CH_2Cl_2	Dichloromethane
$CHCl_3$	Chloroform
CO	Carbon monoxide
CO_2	Carbon dioxide
COP	UN Conference of Parties to the Climate Convention
CSERGE	Centre for Social and Economic Research into the Global Environment
DGM	Directorate General for Environmental Protection
DICE	Dynamic Integrated Climate Economy
DMS	Di Methyl Sulphide
ENSO	El Niño Southern Oscillation
ESCAPE	Evaluation of Strategies to address Climate change by Adapting to and Preventing Emissions
EPA	Environmental Protection Agency
EPCEM	European Postgraduate Course in Environmental Management
ESM	Earth System Model
FCCC	see UNFCCC
FUND	The Climate Framework for Uncertainty, Negotiation, and Distribution
GCAM	Global Change Assessment Model
GCM	General Circulation Model (also: Global Circulation Model)
GDP	Gross Domestic Product
GHGs	GreenHouse Gases
GNP	Gross National Product
IAM	Integrated Assessment Model

ICAM	Integrated Climate Assessment Model
IEA	Integrated Environmental Assessment
IGBP	International Geosphere Biosphere Programme
IIASA	International Institute for Applied Systems Analysis
IMAGE	Integrated Model to Assess the Greenhouse Effect
IPCC	Intergovernmental Panel on Climate Change
IPCC WGI	Intergovernmental Panel on Climate Change, Working Group I
ISM	Integrated Science Model for assessment of climate change
LDC	Lesser Developed Country
LOS	Norwegian Research Centre in Organisation and Management
MAGICC	Model for the Assessment of Greenhouse gas Induced Climate Change
MARIA	Multiregional Approach for Resource and Industry Allocation
MARKAL	Market Allocation
MBIS	Mackenzie Basin Impact Study
MCS	Monte Carlo Simulation
MCW	Model of Global Warming Commitment
MERGE	Model for Evaluating Regional and Global Effects of GHG Reductions Policies
MIT	Massachusetts Institute of Technology
MiniCAM	Mini Climate Assessment Model
N_2O	Nitrous Oxide
NAPAP	US National Acid Precipitation Assessment Program
NH_3	Ammonia
NO_x	The sum of Nitric oxide (NO) and Nitrogen dioxide (NO_2)
NRP	Netherlands National Research Programme on Global Air Pollution and Climate Change
NUSAP	Numeral Unit Spread Assessment Pedigree notational system
NWO	Netherlands Organization for Scientific Research
O_3	Ozone
OECD-GREEN	Organization for Economic Co-operation and Development GREEN model
PAGE	Policy Analysis of the Greenhouse Effect
PAGES	PAst Global changeES (part of IGBP)
PC	Personal Computer
PEF	Policy Evaluation Framework

ppmv	Parts per million by volume (a measure of concentration)
ppbv	Parts per billion by volume (a measure of concentration)
ProCAM	Process Oriented Global Change Assessment Model
RAINS	Regional Acidification INformation and Simulation
RICE	Regional Integrated Climate Economy (regionalized version of the Nordhaus' DICE model)
RIVM	NL National Institute of Public Health and Environmental Protection
RMNO	Advisory Council for Research on Nature and Environment
SO_2	Sulfur dioxide
SF_6	Sulfur Hexafluoride
TARGETS	Tool to Assess Regional and Global Environmental and Health Targets for Sustainability
ULYSSES	Urban LifestYles, SuStainability, and Environmental aSsessment
UN	United Nations
UNFCCC	United Nations Framework Convention on Climate Change
US	United States of America
VEC	Valued Environmental Component
VOC	Volatile Organic Compound
VROM	Netherlands Ministry of Housing, Physical Planning and the Environment
WCRP	World Climate Research Programme
WRI	World Resources Institute
YSSP	IIASA Young Scientist Summer Program

1. Introduction

"The key point to remember is that without thorough and systematic modeling and analysis of the uncertainty of the problem, we can not be sure that the results of a model, especially a very large and complex one, mean anything at all." (Morgan and Henrion, 1990)

In the mid-eighties Integrated Assessment Models (IAMs) emerged as devices to interface science with policy. A perfect IAM would model the complete causal chain, including all the feedbacks within this chain. The causal chain starts with socio-economic drivers leading to economic activity and other practices, leading to emissions and other pressure on the environment leading to environmental changes, leading to physical impacts on societies and ecosystems, leading to socio-economic impacts, eventually returning to cause changes in the socio-economic drivers. There is a controversy regarding the usefulness of IAMs for addressing climate change, in the light of the huge uncertainties and unresolved scientific puzzles in this field. In our view, this controversy, combined with the current policy-use of IAMs and the aspirations of model developers to operationalize Article 2 of the UNFCCC, urgently demands standards for good scientific practice for uncertainty management in IAMs. Therefore, better insight into IAM practice, uncertainties and their management is a prerequisite.

Initially, the climate research programs aimed at the reduction of the uncertainties (e.g. WCRP, 1979, IGBP, 1992). Nowadays, it is recognized by prominent climate researchers that more research does not necessarily reduce the overall uncertainties regarding future climate. For some elements, uncertainties are reduced. However, ongoing research also reveals unforeseen complexities in the climate system, which add to the perceived uncertainty. Furthermore, the insight is growing that there are unresolvable limits to the reduction of scientific uncertainties, due to the epistemological limits of science, our limited capacity to know and understand, limits to our capacity to handle complexity, computer limitations, and the inherent unpredictability of the (partly chaotic) climate system. This growing insight has given mankind an increased stake in understanding the limits of science. If we can't reduce uncertainties, we will have to learn to live with them.

Twelve years ago, Keepin and Wynne (1984) noted that the identification of 'objective policy truths' respectively objective answers to policy questions, is an unrealistic aim in science for policy. This is also stressed by other authors. For instance, Giarini and Stahel (1993) observe "...*the starting of a new era of challenges and opportunities in the evolution of human society; an era in which an unrealistic quest for certainty will be replaced by an understanding of its limits.*".

What is needed in the assessment community is a radical change in focus from "reducing uncertainties" to "managing uncertainties and complexities" and a better understanding of the limits of science in relation to its policy use in IAMs. As a contribution to this objective, answers are sought to the following key questions:

i) What are the possibilities and limitations of IAMs in relation to the aspirations of the modelers to model the complete cause-effect chain and to guide and inform the policy process?

ii) What are the main areas of improvement in uncertainty management in IAMs?

iii) How can we distinguish between uncertainty due to lack of quality and inherent uncertainty of the system, and how can we attribute a part of the overall lack of quality in a model to its individual constituents? In other words, how can we disentangle the uncertainty problem in such a way that we can identify the weakest parts of the model. The weakest parts are those constituents whose individual lack of quality (that is the potentially reducible part of its individual uncertainty) contributes the most to the overall lack of quality in the model outcome. We will also explain why such a identification is needed for the development of adequate response strategies that take the uncertainties into account, and for the setting of research priorities to reduce uncertainties.

In this paper, the main focus is on the IMAGE model (Integrated Model to Assess the Greenhouse Effect, see Figure 1), which served as a case study. IMAGE has a leading role in process-oriented integrated modeling. According to an independent evaluation of the Dutch NRP (National Research Programme on Global Air Pollution and Climate Change), *"IMAGE is an outstanding undertaking, the first attempt to produce a comprehensive integrated assessment framework for the climate change problem and still at the forefront of the field internationally."* and *"IMAGE has established a niche as a world leader in integrated systems modeling."* (Science and Policy Associates, Inc., 1995).

In the following, we first describe the rise of IAMs in the mid-eighties as devices to interface science with policy. We describe what IAMs are, in the detail that is necessary for a good understanding of the next sections. Then, we discuss the key uncertainties and limitations in each stage of the causal chain of the climate issue. Next, we discuss the controversy on the (policy) usefulness of IAMs for the climate issue. In a next step, we analyze the mismatch between the types and sources of uncertainty that should be addressed on the one hand and the current practice of uncertainty management in IAMs and available methodologies to address different types and sources of uncertainty in models on the other hand. Furthermore, we look at the reasons for the mismatch and identify areas for the improvement of uncertainty management. In the final section, we discuss the question of how to proceed from the current problems in uncertainty management and, building further on the work by Funtowicz and

Ravetz (1990), we propose a method for disentangling the uncertainty problem in complex Integrated Assessment Models.

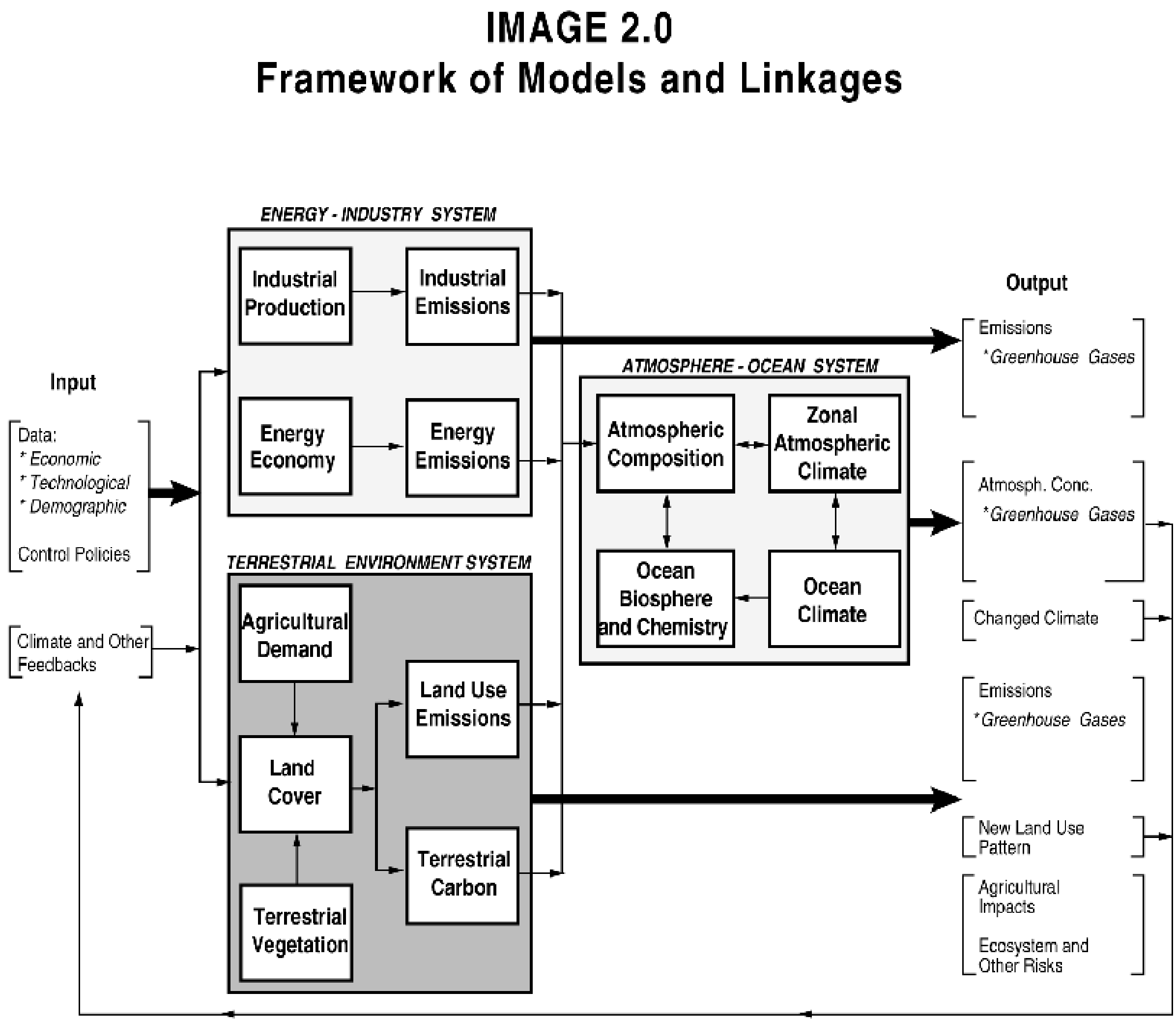

Figure 1 The IMAGE 2.0 model (Alcamo, 1994)

2. The emergence of IAMs as a science-policy interface

Facilitated by developments in computer technology, *integrated modeling* emerged in the mid-eighties as a new paradigm for interfacing science and policy concerning complex environmental issues. In the second half of the eighties, it was believed that integrated modeling would be the optimal way to interface science with policy (Zoeteman, 1987). Parson (1994) claims that: *"To make rational, informed social decisions on such complex, long-term, uncertain issues as global climate change, the capacity to integrate, reconcile, organize, and communicate knowledge across domains - to do integrated assessment - is essential."* Integrated assessment models are believed to produce insights that cannot be easily derived from the individual natural or social science component models that have been developed in the past (Weyant, 1994).

The first generation of integrated models focused on acid rain. The RAINS model (Regional Acidification INformation and Simulation) is the most obvious example (Alcamo *et al.*, 1986; Alcamo *et al.*, 1990; Hordijk, 1991a). RAINS was developed in the eighties at the International Institute for Applied Systems Analysis (IIASA). The RAINS model played a major role in the international acid deposition negotiations in the framework of the United Nations Convention on Long-Range Transboundary Air Pollution and became an annex to the SO_2-protocol (Hordijk, 1991b, 1994). In the second half of the eighties, RIVM developed the IMAGE model (Integrated Model to Assess the Greenhouse Effect), which was a pioneer in the field (De Boois and Rotmans, 1986; Rotmans, 1990). IMAGE has been used for scenario calculations in the influential Netherlands policy document "Zorgen voor Morgen" (Concern for Tomorrow; Langeweg, 1988) and for the development of emission scenarios 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 substantially revised version of the model, IMAGE 2 (Alcamo, 1994a), has been used by all three working groups of IPCC, mainly for developing reference and policy emission scenarios (Swart, 1994b; Alcamo, 1994b, Science and Policy Associates, 1995). Results produced by IMAGE 2.0 and IMAGE 2.1 were presented to the negotiators at the United Nations Conference of Parties to the Climate Convention (COP-1) in Berlin, March, 1995 (Alcamo *et al.*, 1995) and follow up (Alcamo and Kreileman, 1996), and at the second meeting of the COP in Geneva, July 1996 (Swart *et al.*, 1996).

Over the past few years, the number of integrated assessment models (hereafter referred to as *IAMs*) of the climate problem has grown significantly. In 1990, there were 3 IAMs: RIVM's IMAGE model (De Boois and Rotmans, 1986; Rotmans, 1990), WRI's Model of Warming Commitment (MWC, Mintzer, 1987), and Nordhaus' model (1989, 1990). The current number of IAMs for the climate issue is at least 40 (see Appendix 1 for a list).

At present, one of the main policy questions being addressed by IAMs concerns the operationalization of Article 2 of the United Nations Framework Convention on Climate Change (FCCC): *"The ultimate objective of this Convention and any related legal instruments that the Conference of the Parties may adopt is to achieve, in accordance with the relevant provisions of the Convention, stabilization of greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system. Such a level should be achieved within a time frame sufficient to allow ecosystems to adapt naturally to climate change, to ensure that food production is not threatened and to enable economic development to proceed in a sustainable manner."*

Article 2 is operationalized in recent IMAGE modeling studies by defining 'safe landing zones' and corresponding 'safe emission corridors'. The 'safe landing concept' is a metaphor. If a plane goes down too slowly, it may miss the runway and crash beyond it. If a plane goes down too quickly it may crash before it reaches the runway. In climate terms: if policies are too few and too late, serious climate impacts may be unavoidable. If measures are too strict and too early, economic effects may be unacceptable (Swart *et al.*, 1996). The starting point is a set of climate goals that are considered safe. Such goals are based on scientific assessment but the selection of their level is a political process. Among the attempts to provide scientific rationales for the level of such goals are those by Krause (1989), Rijsberman and Swart (1990) and Jäger (1990). A typical set of such climate goals is:

i. change in global surface temperature relative to 1990 ≤ 2.0°C;

ii. rate of temperature change ≤ 0.2°C/decade;

iii. sea level rise relative to 1990 ≤ 40 cm;

iv. maximum rate of emission reduction ≤ 2 %/yr.

In terms of Article 2, goals (i) and (iii) relate to the 'prevent dangerous anthropogenic interference with the climate system', goal (ii) relates to the 'time frame sufficient to..', whereas goal (iv) relates to the 'enable economic development to proceed in a sustainable manner'. For the above given set of goals, IMAGE calculates that the 'safe emission corridor' or allowable global emissions of all greenhouse gases together in 2010 range from 6.2 to 14.1 Gt C/yr CO_2 equivalents. Current emissions are about 9.6 Gt C/yr CO_2 equivalents (Alcamo and Kreileman, 1996). As a means of graphical representation, these IMAGE studies use the three colors from a traffic light: If all criteria stay at least 20% below their target value, the line is green, if at some moment in the simulated future at least one indicator is exceeded by more than 20% the line turns red, if at least one criterion is in the 20% interval around the target value, the line is amber. This method allows the evaluation of scenarios and the testing of the extent to which the simulated climate stays within the safe landing zones defined by the four climate goals, given the total set of assumptions that constitute the model. Whether this really indicates that given a scenario we will stay in that zone depends on the validity of the model, and whether such really is safe depends on the validity of the operationalization of 'safe' with the four climate goals. Unfortunately, the answer to both questions is 'we don't know and we won't know' (see also Ravetz, 1986).

3. What are IAMs?

With regard to the climate issue, a perfect IAM would analyze the full cycle of socio-economic drivers, economic activity, anthropogenic emissions of greenhouse gases -although many of them look at CO_2 only-, their concentrations in the atmosphere, the resulting climate forcing, climate change, sea level change etc., and finally the impacts of these on the economy, food production systems, water supply systems, ecosystems as well as other human activities. On the one hand, 'integrated' refers to both the completeness that is aimed at and the inclusion of the feedback loops in and between the represented coupled cause effect chains (Rotmans, 1994). On the other hand, 'integrated' refers to the notion that these models bring together information and analysis from disparate disciplines (Parson, 1994, Weyant, 1994).

3.1. Definitions of IAMs

The literature shows a variety of definitions of IAMs, which we have summarized in Table 1.

Table 1. Definitions of Integrated Assessment and Integrated Assessment Models.

Author	Definition of IAM
Swart, 1994a p.128	"Integrated models are defined here as interdisciplinary models that capture the full cause-impact chain of the climate change problem and all sectors involved. Thus, in this definition, no integration across environmental themes is implied."
Parson, 1994	"The two defining characteristics are a) that it seek to provide information of use to some significant decision-maker rather than merely advancing understanding for its own sake, and b) that it bring together a

	broader set of areas, methods, styles of study, or degrees of confidence, than would typically characterize a study of the same issue within the bounds of a single research discipline."
Weyant, 1994 p.3	"One definition of an integrated assessment model is a model that projects future economic activity and the links between it and greenhouse gas emissions, between emissions and atmospheric concentrations and climate change, between climate change and the physical impacts on economies and ecosystems that result, between those physical impacts and their economic valuation. Following Parson (1994), however, for our purpose here we define any climate change focused model that links together information and analysis from disciplines that have not traditionally been combined as an "integrated assessment" model."
Rotmans and Dowlatabadi, 1995	"Integrated assessment can be defined as an interdisciplinary process of combining, interpreting and communicating knowledge from diverse scientific disciplines in such a way that the whole cause-effect chain of a problem can be evaluated from a synoptic perspective with two characteristics: (i) it should have added value compared to single disciplinary assessment; and (ii) it should provide useful information to decision makers"
Toth, 1995 p257	"In this paper, and throughout the collection that follows, the terms 'integrated model' and 'integrated assessment' refer to a set of formal models or studies without modeling support that are combined into a consistent framework to address one or more issues in the problem of global climate change."
Kasemir and Jaeger, 1996; Baily *et al.*, 1996	"As integrated environmental assessments we understand procedures to arrive at an informed judgement on different courses of action with regard to environmental problems. The information required refers to physical, chemical, biological, psychological, socio-economic and institutional phenomena, including the relevant decision-making processes."

In this paper we define an *integrated assessment model* as a mathematical representation of a coupled natural system and a socio-economic system, modeling one or more *cause-effect chains* including feedback loops, and explicitly designed for the purpose of addressing policy questions, mostly by means of scenario analysis. This explicit policy purpose defines the difference between IAMs and Earth System Models (ESMs) such as Atmosphere Ocean General Circulation Models (GCMs) and geochemical models, which are designed primarily for scientific purposes. It should however be noted that ESMs such as GCMs could also be used (and in fact they are) to look at policy questions.

The integration across disciplines, which is in Swart's, Parson's, Weyant's, Rotmans and Dowlatabadi's and Kasemir and Jaeger's definitions (see Table 1) is the consequence of including both the natural and the social system and representing complete cause-effect chains. In the third row of Table 2, we give examples of disciplines that have competence with respect to elements of each of the stages of the causal chain. Multidisciplinarity is indeed a general characteristic of IAMs, but not unique to IAMs. In fact, many ESMs also are multidisciplinary. The two unique characteristics of IAMs are that they (1) integrate the natural system and the socio-economic system in one model and (2) have an explicit mission to address policy questions.

A computer and a - user-friendly - software implementation of the IAM, including a user interface, are used to numerically integrate the model through time with a user-definable set of presumed future time series of input variables (e.g. population growth). Such a set of presumed future time series is called a *scenario*. Integrated models also permit the inclusion of a presumed portfolio of measures in a scenario, such as the introduction of carbon tax, a switch to renewables, reforestation, etc.

Janssen and Rotmans (1995) distinguish two different approaches in current IAM-practice: the macro-economic parameterized approach and the biosphere-climate process-oriented approach. The macro-economic models aim at deriving cost effective strategies to cope with the climate problem, whereas the biosphere-climate process oriented models aim at analyzing the consequences of human activities. In this paper we focus on the biosphere-climate process oriented IAMs. Most of our examples and illustrations from the current modeling practice come from the IMAGE model, which served as a case study (see also Van der Sluijs, 1995).

Biosphere-climate process oriented IAMs have a modular structure of sub-models. For instance, the IMAGE 2 model consists of three sub-systems consisting of - in total - 13 sub-models (see Figure 1). Each sub-model has its roots in a different discipline of science. The individual sub-models are more or less radically simplified and aggregated input-output models that are usually derived from comprehensive models. While the comprehensive models usually have high process detail and consist of mathematical equations that directly reflect the processes as we think they occur in reality, the simplified modules in IAMs are more like calibrated black- or gray-boxes. This simplification is the inevitable consequence of computer limitations and the mission of IAMs to address policy questions: To be of use to the policy-making process, IAMs should facilitate the comparison of many different user-definable scenarios in a reasonable time frame. If the most comprehensive model available were to be used for each sub-model, the calculation of one scenario would take several years of calculation time, even on the fastest super-computer. The ideal of IAM-modelers is to produce a model that can be used interactively by a policy-maker on her or his own desk-top PC and that gives results that do not differ significantly from the hypothetical IAM that would result from choosing the most comprehensive models available for each module.[1]

3.2. Variability in IAM-modeling practice

In this section, we discuss inter-IAM variability in its approach and its major sources. To start with the first step, there are several techniques to draft aggregated simple input-output models from comprehensive high process-detail models. Meta-models can be generated by fitting simple mathematical equations to input-output data from the comprehensive model, but also by aggregating inputs and outputs by sensitivity analysis (Hordijk, 1994). The degree of simplification depends on the state of knowledge, the level of aggregation required and the complexities of the systems. All modules in an IAM should have about the same level of detail. Differences in aggregation level across modules are often handled by (statistical) interpolation and aggregation techniques. These techniques are partly based on somewhat arbitrary (though educated) untested assumptions.

Differences in aggregation level of sub-models are not the only problem that is encountered when sub-models from different disciplines are coupled in an IAM. In the IIASA Integrated Assessment project, the geographical regions used in the population model were different from the geographical regions in the energy model. Other problems are that the sub-models can have differences in the basic assumptions or different values for the same paramter when this parameter is used in more than one sub-model. When coupling the models, one has to go back to the basic assumptions to make sure that they are consistent troughout the IAM. Sometimes, technical fixes are introduced to handle these problems.

Climate IAMs differ in a range of aspects. Weyant (1994) classifies the IAMs as being (1) either more or less complex in representing the natural science process associated with climate change, (2) either more or less complex in representing the economic process associated with climate change (3)

---

[1] Not all modelers share this ideal, because direct use of IAMs by policy makers runs the risk of manipulative use of the model. The classic example is the case of US-president Bush' adviser Sununu, who misused a climate IAM to support the stance that measures were not necessary.

either explicitly incorporating uncertainty or not explicitly incorporating uncertainty. Toth (1995) emphasizes other aspects: They can be partly or fully integrated, depending on how much of the cause-effect chain is covered. They differ in the level of integration. The most advanced models integrate equations in one single system. Other models use different techniques of hard and soft linking to transmit data between individual modules. Further they differ in comprehensiveness in terms of the degree to which models include sources and sinks of all greenhouse gases.

Another important aspect is the extent to which feedbacks between different stages are taken into account dynamically during the simulation. Dynamically means that all feedbacks are evaluated and take effect after each single time step of numerical integration of the differential equations that constitute the system (or after each discrete time step in discrete models), before the next numerical integration step (or discrete time step in discrete models) is performed (see e.g. Jacoby and Kowalik, 1980). According to Weyant (1994): *"very few of the operational models include interactions and feedbacks between modules other than a straight pass through of information from one module to the next. The IMAGE 2.0 and the AIM models are notable exceptions".* The TARGETS model and the MIT model also include complex interactions but were not yet operational when Weyant wrote his review.

Further, IAMs differ in level of aggregation and disaggregation. Aggregation is defined as the joining of more or less equivalent elements that exhibit mutual interaction (Goudriaan, 1993). Aggregation can take place across different scale-dimensions, leading to different resolutions on these scales. The most relevant scale dimensions in IAMs are: temporal scale (e.g. diurnal; seasonal; annual), spatial scale (e.g. local; regional; continental; global), systemic scales (e.g. individual plants; ecosystems; terrestrial biosphere), and conditional scales (e.g. ecosystem internal variability; inter-ecosystem variability; global variability).

According to Parson (1994), there is no consensus regarding the best approach: *"Perhaps the most serious consequence of the immaturity of the field is that there is no shared body of knowledge and standards of 'best practice' for integrated assessment. Such knowledge is likely to develop with more thought and practice, but its present absence makes it ill-advised to pursue a single, authoritative vision of integrated assessment. On both intellectual and managerial dimensions, there are many plausible ways of addressing the most basic challenges of integrated assessment. There is no single right way to do it."* Parson's view is consistent with Weyants' (1994) notion that: *"It is also possible that further disaggregation and more explicit treatment of uncertainty would not lead to different insights than produced by these simpler deterministic models."* However, the reverse is also possible.

4. Key uncertainties and limitations faced by IAMs of the climate issue

The ambition of the IAM community is to model the entire cause-effect chain of anthropogenic climate change in one integrated model. In this section we explore the possibilities and limitations of IAMs in relation to this ambition. In Table 2, we show key uncertainties and limitations in modeling future behavior for each stage of the causal chain, following the causal taxonomy developed by Norberg-Bohm *et al.* (1990) (first row in Table 2), to which we have added 'Culture and Values'[1]. The key uncertainties listed in the Table are gathered mainly from a review of the literature and from personal communication with experts combined with our own expertise. The Table does not pretend to be complete, but rather provides illustrative examples of the various kinds of limitations and uncertainties that are encountered in the assessment of possible future behavior of key constituents in each stage of the causal chain.

[1] This modification of Nordberg-Bohm's classification was inspired by the preliminary draft "ULYSSES Research Protocols GeMID: Generic Model Iterative Dialogue for Urban Lifestyles And Sustainablity" dd 16/08/1996 that was distributed via the ULYSSES E-mail mailing list.

Causal chain	Culture and values	Demands for goods and services	Choice of technologies and practices	Flux of materials	Valued Environmental Components	Exposure	Consequences
Examples of variables	-Values -Attribution of responsibility -Culture -Risk perception -Rationality -Ethical attitude -Driving value -Myth of nature -Preferences -Religion -Laws/Legislation	-Population size -GNP -Consumption per capita	-Energy efficiency -Land use -Life style -Share of coal, natural gas, nuclear, renewables, biomass -Energy price	-Emissions of greenhouse gases (CO_2, CH_4, N_2O, CFCs etc.) -Emissions of ozone precursors (CO, NO_x, VOCs etc.) -Emissions of aerosols	-Temperature -Precipitation -Soil moisture -Sea level -Storm frequency -River runoff -Tidal amplitude -Ocean circulation patterns	-Welfare -Healthiness/fitness of population -Sensitivity -Adaptability -Vulnerability -Damage thresholds of buildings and infrastructure for storms and floodings	-Agricultural production -Biodiversity -Storm damage -Flood damage -Migration patterns -Loss of property -Land loss -Health -Loss of water supply
Corresponding sub-system or sub-model	-Society -Culture -Values -Laws	-Economy -Population -Mobility -Agricultural demand -Values -Culture -Behavior	-Energy economy -Transport economy -Industrial production -Households -Agriculture -Land use -Life style	-Industrial emissions -Energy emissions -Land use emissions -Global Carbon Cycle -Other Global Biogeochemical Cycles -Atmospheric chemistry	-Terrestrial biosphere -Marine biosphere -Atmosphere Ocean Climate system	-Vulnerability assessment models -Values -Culture	-Impact assessment -Cost effectiveness -Cost benefit analysis -Values -Culture
Examples of competent disciplines	-Cultural theorists -Psychologists -(Moral) Philosophers -Sociologists -Historians -Jurists	-Demographers -Economists -Statistical agencies -Mathematicians -Psychologists	-Energy-experts -Economists -Engineers -SSST-ers (Social Studies of Science and Technology) -Cultural theorists -Psychologists -Philosophers -Sociologists	-Statistical agencies -Ecologists -Geo(bio)chemists -Biologists -Atmospheric chemists -Engineers	-Meteorologists -Oceanographists -Atmospheric chemists -Ecologists -Geo(bio)chemists -Biologists -Social psychologists	-Engineers -Hydrologists -Geographers -Biogeographers -Epidemiologists -Philosophers	-Economists -Hydrologists -Epidemiologists -Mathematicians -Psychologists -Sociologists -Cultural theorists
Key uncertainties	-Incomplete understanding -Undeterministic elements	-Demographic uncertainties -Behavioral uncertainties -Undeterministic elements	-Unpredictability of technological innovations -Incomplete understanding of implementation barriers -Undeterministic elements	-Incomplete information	-Incomplete understanding - Biogenic feedbacks -Chaotic behavior -Multiple equilibria/Non-smooth behavior -Modelability of surprise -Linkages with other anthropogenic environmental changes	-Uncertainty about vulnerability of future societies subjected to climate change	-Fundamental limits to predictability of future regional climate -Fundamental uncertainties in the attribution of monetary values

Table 2 Examples of variables, sub-systems, competent disciplines and key uncertainties in each stage of the causal chain.

We will discuss the uncertainties in each step of the causal chain in more detail.

4.1. Culture and Values

Culture and values are at the basis of the causal chain. The future development of key variables such as risk perception, attribution of responsibility, life-attitudes (e.g. soberness), ethical attitude (ecocentrism or anthropocentrism), driving value (growth, equity, stability), myth of nature perceived-most-plausible (robust, fragile, or robust within bounds), laws and legislation, valuation of consequences for future generations, and other factors that can influence how many and what goods and services we demand and what life-style we develop, are poorly understood and are to a certain extent open ended and hence unpredictable.

4.2. Demands for goods and services

A given ensemble of culture and values that constitutes a human society, gives rise to demands for goods and services, such as energy, transport, housing, heating, food etc. The resulting total demands are further a function of population size and composition, behavior, life style etc. The future behavior of these entities is very hard to forecast reasonably reliably on time scales longer than roughly 10 years. IAMs for the climate issue usually use a 100-year time horizon. For instance, demographic uncertainties concern uncertainties about what factors trigger structural change in fertility behavior, which still is poorly understood (Van Asselt, Beusen and Hilderink, to be published; Rotmans and De Vries, to be published). Keilman (to be published) observed that errors in the prediction of fertility are much higher than those in mortality and that behaviorally determined variables are difficult to forecast. In comparisons of past population projections with the actual data, large errors have been found for both the young and the old after a forecast period of 15 years (up to +30 percent for the age group 0-4, and 15 percent or lower for women aged 85+ are not uncommon). According to Keilman, this suggests that those old forecasts supplied useful information for perhaps up to 10-15 years ahead, but certainly not longer. He also concludes that detailed studies for a few countries have found only modest systematic improvements in the accuracy of forecasting over time (if at all) when series of population forecasts produced by statistical agencies over a long period were analyzed.

An interesting development in the field of population forecasting are the probabilistic population projections, obtained by Monte Carlo Analysis using subjective distribution functions[1], developed in the IIASA population project (Lutz *et al.*, 1996).

[1] Monte Carlo Analysis and Subjective distribution functions will be explained and discussed in section 7.1 of this paper.

4.3. Choice of technologies and practices

There are a number of constraints on the prediction of the future choice and availability of technologies and practices to fulfill demands for goods and services. A fundamental constraint here is the unpredictability of technological innovation. In 50 years, there might be an inexpensive clean energy technology available. However, the availability of this technology does not mean that it will be used. Several studies have shown that many negative-cost energy saving measures can be taken in different sectors of industry, but that they are not taken because there are essentially unknown or poorly understood barriers to their implementation (e.g. Worrell, 1994), and because lock-in effects affect technology choice (Arthur, 1988). A way of endogenizing some aspects of technological innovation in the models is the use of learning-curves to represent trends in energy intensities of different countries. This approach is taken *inter alia* in the IIASA-WEC scenario's and in the energy sub-model (TIME) of IMAGE and Targets (WEC/IIASA, 1995; De Vries and Van den Wijngaart, 1995).

Large uncertainties are also associated with the estimates of future costs of technologies, which are based largely on informal guess work regarding the quantification of poorly known cost-factors, as we know from our own experience in this field (Van der Sluijs *et al.*, 1992).

4.4. Fluxes of material in the environment

The estimation of emissions is the next step in the causal chain. For a given future energy demand and technology mix to fulfill this demand, you need to know the emission characteristics of the technologies and practices in order to determine the associated emissions. A problem here can be incomplete information on e.g. the emission characteristics of the technologies and life styles. Much more difficult to estimate are the changes in fluxes of greenhouse gases related to land-use changes and changes in practices, especially with regard to the non-CO_2 greenhouse gases. There are already significant uncertainties in estimating current emissions of all greenhouse gases due to incomplete information. (e.g. Ebert and Karmali, 1992). The sources of CH_4 and N_2O are not well quantified (Schimel *et al.*, 1996 in IPCC'95) and emissions of substances such as CH_3CCl_3 and CO that affect atmospheric chemistry pathways and life time of greenhouse gases are not well quantified either. The uncertainties in sources and sinks of CO_2 are still large and still allow for the view that there is a missing (unidentified) carbon sink. For instance, the combined CO_2 fertilization, N-fertilization and other climate effects on global plant growth for the period 1980-1988 is estimated by IPCC'95 to be 1.3 ± 1.5 GtC/yr (estimated 90% confidence interval), indicating an error in the numeral that is of the same order of magnitude as the numeral itself, and indicating that even the sign of that flux is unknown as it can also be negative.

Man's knowledge about atmospheric chemistry of the non-CO_2 greenhouse gases is incomplete. The estimates for the atmospheric lifetimes of non-CO_2 greenhouse gases are surrounded by huge uncertainties. For instance, the uncertainties in estimated life time amount to 300% for $CHCl_3$, 200% for CH_2Cl_2, 25% for CH_4 whereas the uncertainties in the lifetimes of the various CFCs range from 20% to 300% (all figures from Schimel *et al.*, 1996 in IPCC'95). As a consequence, the resulting future greenhouse gas concentrations calculated with Global Carbon Cycle models and Atmospheric Chemistry models are surrounded by significant uncertainties. These uncertainties however can be reduced when further research succeeds in quantifying sources and sinks more accurately and in unraveling the complexities of interrelated atmospheric chemistry pathways.

4.5. Valued Environmental Components

Changes in material fluxes lead to changes in Valued Environmental Components (VECs), which are those attributes of the environment which humans value. In general, we value those components not in themselves (although this would be debated by *inter alia* the deep ecologists), but because changes in them lead to undesired consequences. The environmental components valued by human societies are apt to change with changes in culture, values, perception, and technology; all are unpredictable processes. As an implicit working hypothesis, the current IAMs assume that the set of currently valued environmental components will be the same as the set of components valued in the future. However, the current record falsifies the validity of this assumption: Looking over the currently available time series of successive climate risk assessments from the early-seventies until the present, many new VECs have shown up over time that were not valued before; these have emerged due to innovations in climate research (Jäger *et al.*, to be published, Van der Sluijs and Van Eijndhoven, to be published). Examples are the concentrations of newly discovered greenhouse gases such as CFCs, N_2O and CH_4; growing interest in storm frequencies and intensities triggered by the dramatic increase in storm damage in the USA caused by super storms in the past decade; growing interest in local soil moisture, due to the '*ensure that food production is not threatened*' objective in Article 2 of the FCCC; growing interest in ocean circulation patterns since the discovery of the 'conveyor belt' and the recent notion that it might be switched off via climate-change-induced changes in the hydrological cycle.

The future states and temporal and spatial distribution of Valued Environmental Components such as temperature, precipitation, soil moisture, sea level, tidal amplitude, storm intensity, frequency and duration, ocean circulation patterns etc. are evaluated using Earth System Models such as coupled Atmosphere Ocean General Circulation Models (GCMs, also known as: Global Climate Models), ocean circulation models, models of ice-dynamics and sea-level, biosphere models, etc.

The most advanced models used for the assessment of climate change are the GCMs. The simplified climate models used in IAMs are calibrated to GCM results that are supposed to back them scientifically. The usefulness of GCMs (and hence of the simpler climate models in IAMs) for formulating policy advice has been debated. As two well-known modelers put it, the common wisdom is that feedbacks can "*be predicted credibly only by physically based models that include the essential dynamics and thermodynamics of the feedback processes. Such physically based models are the general circulation models*" (Schlesinger and Mitchell, 1987). The IPCC takes the same position on the usefulness of Atmosphere Ocean GCMs: *"... it is generally believed that it is only through such models that we can gain a scientific understanding (and hence a reliable predictive capability) of climate and climate change."* and *"This faith in the fundamental soundness of the modeling approach does not deny the presence of significant errors in current models nor the utility of models known to be incomplete, but does provide confidence that these errors can and will be reduced through continuing modeling research."* (Gates *et al.*, 1996 in IPCC-WGI'95)

Ann Henderson-Sellers, a prominent GCM modeler involved in several GCM inter-comparison projects and co-author of the chapter from which we took the previous IPCC quote, says elsewhere about these models: *"Today's climate models are essentially useless for virtually all forms of policy advice related to climate change. They are useful for some forms of short-term forecasting and medium range climate advice (e.g. El Niño projections, ...), but for long-term advice related to the enhanced greenhouse effect the value is minimal at best. The key conclusions of the models are driven by the assumptions and the various structures and devices used to simplify the calculations to make the models computable with today's technology. This massive problem is an important feature of the difficulty in linking the science and the policy."* (Henderson-Sellers, 1996a, p.43/44). We think Henderson-Sellers touches a very important point here which holds for essentially all Earth System Models, namely that they are not designed for answering policy questions, but rather for gaining insight into the modeled system. They key-word for gaining insight is simplification. According to Leo Schrattenholzer, a model is a simplified representation of a system, where simplification is the goal and not the restriction (IIASA Seminar on "Comparing different philosophies or practical approaches to modeling", 9 July 1996). What happens now is that models designed for scientific purposes are used directly in integrated assessment models to address policy questions. It might well be that addressing policy questions requires differently designed models in which simplified, idealized, smoothed deterministic representation is not adequate.

We will discuss some key problems currently encountered in the practice of modeling the natural system. These are (i) incomplete understanding, (ii) biogenic feedbacks, (iii) chaotic behavior, (iv) multiple equilibria and (v) linkages with other global environmental changes.

i. Incomplete understanding of the modeled system

In a recent ranking exercise in which 16 prominent climate modelers from the US were involved, the top five sources of incomplete-understanding-uncertainty in climate modeling were identified from a list of 25 candidates (Morgan and Keith, 1995). These are:

- Cloud distribution and optical properties (including aerosol effect);
- Convection-water vapor feedback (all processes transport water vertically, except in the planetary boundary layer);
- Carbon dioxide exchange with terrestrial biota
- Carbon dioxide exchange with the oceans (including the ocean biota)
- Oceanic convection (e.g. high latitude production of deep water that is believed to drive the so-called "conveyer belt")

The participating experts expect a very significant reduction of uncertainty regarding climate sensitivity if complete understanding of these 5 elements is be reached. Uncertainties about non-CO_2 greenhouse gases were ranked low.

These findings are consistent with MIT's claim that "*MIT concludes that the most gain in reducing overall uncertainty in climate behavior would be achieved by better understanding three processes: convection, cloud formation, and ocean circulation. They also conclude that progress in these fields will be slow during the next decade.*" (E-lab January-March 1995).

ii. Biogenic feedbacks

A number of feedback mechanisms, especially those in which the biota play a key role, are left out of the models because of a reluctance on the part of meteorologists and oceanographers to quantify these processes (Margulis and Lovelock, 1974; Schneider, 1989; Leggett, 1990; Turkenburg and van Wijk 1990; Ambio, Febr. 1994). Until recently, very few attempts had been made to include biological processes in GCMs, beyond highly simplistic representations of the land-surface. Examples of biogenic mechanisms that might play a significant role in climate feedback loops include:

- The role of vegetation in surface properties (Melillo *et al.*, 1996, in IPCC'95);
- CO_2-induced reductions in stomatal conductance, resulting in lower evapotranspiration which affects both soil moisture and latent heat transport (Melillo *et al.*, 1996, in IPCC'95);

- The formation of biogenic substances that form a molecular top-layer upon ocean water which inhibits ocean evaporation (Personal communication P. Westbroek);
- The role of Di-Methyl Sulphide (DMS) produced by marine algae in modulating cloud formation and cloud optical properties (Charlson *et al.*, 1987);
- Enhanced aerobic respiration and large-scale oxidation by erosion and fire of high latitude peats (these peats are estimated to contain 450 GtC) caused by eventual drying and warming of these regions. This has also consequences for the CH_4 balance, because if the aerobic top layer of the soil becomes thicker through drying, its CH_4 uptake will increase. (Melillo *et al.*, 1996, in IPCC'95);
- The role of biota in the carbonate-silicate geochemical cycle: In a CO_2 rich world the weathering rates of silicate might increase through enhanced vascular plant growth and enhanced CO_2 concentrations in soils, and by the intensification of the hydrological cycle via increased wash-out of carbonate. Carbonate is the product of the silicate weathering reaction, so increased wash-out will speed up the weathering rate. This feedback can be inferred from Van der Sluijs *et al.* (1996). A possible consequence is that the carbon sink caused by silicate weathering will no longer be negligible for the shorter time scales in a CO_2 rich world. This hypothesis is supported by tentative unpublished work by Jan Goudriaan, who did experiments with a modification of his carbon cycle model that included the weathering equations from the Van der Sluijs *et al.* 1996 model (personal communication with Jan Goudriaan, 1994);
- Albedo changes due to land-use change and vegetation changes;
- The influence of climate change on algae-blooms of Emiliania Huxley and the influence of these blooms on the albedo (personal communication P. Westbroek).

For a more comprehensive discussion of the biospheric modulation of the climate we refer to Margulis and Lovelock (1974), Lovelock (1988) and Westbroek, (1991). For a recent review of biogenic feedbacks we refer to Woodwell and Mackenzie (1995).

More detailed investigation of biogenic feedbacks requires the coupling of biosphere models to GCMs. This task is compounded because this means that the GCMs will have to resolve the vertical structure of the planetary boundary layer, which present GCMs don't do. Further it requires a physiologically based representation of the processes controlling canopy conductance. As long as the biophysical key processes controlling soil moisture are not taken into account in GCMs, their simulations of soil moisture in a high CO_2 world are highly questionable (Melillo *et al.*, 1996 in IPCC'95). Further, given that the importance of these, and other feedbacks, is currently not known, but could conceivably be significant, the estimates of climate sensitivity using current GCMs might well be inaccurate.

iii. Chaotic behavior

Chaotic behavior usually refers to the phenomenon that very small changes in the system parameters or initial state can have a disproportionally large impact on the system behavior of non-linear systems, which makes the system practically unpredictable. Since Lorenz, it has been widely believed that weather is chaotic, with a loss of coherence (for neighboring initial conditions) in one or two weeks (Abarbanel *et al.*, 1991). This notion puts limits to the long-term predictability of weather and climate (see also Tennekes, 1992, 1994). The IPCC assumes that elements of the climate system are chaotic, while other elements are stable. *"The existence of these stable components allows prediction of global change despite the existence of the chaotic elements"* (Houghton *et al.*, 1990 p.80). IPCC calls their own assumption of smooth response of the climate system to forcing a *"reasonable working hypothesis, which receives some support from the smooth transient response simulated by coupled ocean atmosphere models."* (Houghton *et al.*, 1990, p.80).

The basic assumption behind the IPCC view is that there is a sharp distinction between fast elements of the atmosphere ocean systems and the slow elements (Hasselman, 1976). The fast components are partly chaotic, but the slow ones are assumed to be non-chaotic. This view has been criticized *inter alia* by Abarbanel *et al.* (1991) on the grounds that such a strict separation between fast and slow elements is, in principle, invalid because it should be treated as a coupled system. They claim that the real issue in the predictability of climate is whether the atmosphere-ocean systems constitute a chaotic dynamical system on all time scales, and provide evidence (but no proof) from a 134 years global mean temperature record that chaotic behavior exists on all time scales.

Recent findings of the IGBP-PAGES (International Geosphere Biosphere Programme, PAst Global changES) also indicate that the climate system on long time scales is not as smooth as was assumed hitherto: *"High resolution records (from ice core and lake sediments) reveal rapid climate changes by several degrees within a decade or so"* (Lorius and Oeschger, 1994). This makes them conclude that *"global change science faces a new great question: can climate ever be predictable?"*

iv. Multiple equilibria/non-smooth behavior

The issue of transitivity is another key uncertainty in modeling future states of valued environmental components. A transitive system is one which has only one equilibrium state, an intransitive one has at least two equally acceptable states. In an almost intransitive system, on the other hand, it is impossible to determine what is the 'normal state', *"since either of the two states can continue for a long period of time, to be followed by a quite rapid and unpredictable change to the other"* (Henderson-Sellers & Robinson 1986:361).

The notion of multiple equilibria in the earth's climate is not new. Already in 1978 Oerlemans and Van den Dool showed, with a zonally averaged climate model of the energy balance and satellite measurements from that time, that for the actual solar constant, both the present climate and an ice-covered earth are stable solutions of the model. They investigated the effect of variation in the solar constant in detail, and found that if the solar-constant is decreased by 9-10% the warm solution jumps to the cold one. Transition from the cold to the warm solution requires an increase of the solar constant to 109-110 % of its present value. One of their conclusions was that our climate is more stable for solar variations than was previously assumed, but also that the model is more sensitive to changes in the greenhouse effect than to solar variations.

Clearly, the 'do-ability' of climate prediction depends on the system's transitivity. Yet, at present, geological and historical data are not detailed enough to determine which of these system types is typical for several sub-systems of the geosphere-biosphere system and the resulting coupled earth system. It is easy to see that should the climate turn out to be almost intransitive it will be extremely difficult to model (Henderson-Sellers & Robinson 1986:361). According to Bengtsson (1992:716) ".... *there are no means at present of determining whether the atmosphere-ocean-earth system is transitive or intransitive. Perhaps the most intriguing and one of the most challenging problems in climate research is the suggestion that the thermohaline circulation in the Atlantic Ocean is intransitive.*". Recently, Rahmstorf (1995) showed, with a global ocean circulation model coupled to a simplified climate model, that moderate changes in fresh water input in the North Atlantic thermo-haline circulation (the so-called "conveyor belt", of which the Gulf Stream is a component) can induce transitions between different equilibrium states, leading to substantial changes in regional climate of several degrees in time scales of only a few years.

In the earlier mentioned survey among 16 prominent US climate experts, 14 of them gave as their expert judgement on the question *"Are there multiple stable climate states?"* the answer *"yes"*, one answered *"no"*, and one answered that he/she views the climate system as a non-equilibrium system wandering through phase space. The latter view is supported by Van der Sluijs, Westbroek and De Bruyn (1996), who speculate that the Pleistocenic glacial interglacial cycles, through which the current climate is believed still to be looping, might be understood as stable limit cycles in the long term carbon cycle.

A closely related issue is non-smooth behavior. Non smoothness means that there are discontinuities (jumps) or discontinuities in the first derivate (sudden changes in trends) in the behavior of a system over time. Non-smoothness introduces the problem that trends identified from the past trajectory are bad predictors for the future behavior in the immediate environment of a discontinuity in the modeled phenomenon. The geological record of the earth system suggests the existence of non-smoothness in the natural system. We will discuss this issue later in this paper in the section on the modelability of surprise.

v. Linkages with other anthropogenic environmental changes

There are links between acidification and climate change. Sulphate aerosols affect the radiation balance by reflecting incoming solar radiation. They also have an indirect effect via their role in cloud formation (Tailor and Penner, 1994; IPCC 94; 95). A further link between acidification and climate change is nitrogen fertilization of the terrestrial biosphere (Melillo *et al.*, 1996 in IPCC'95). There also are links between stratospheric ozone depletion and climate change. The halocarbons that cause ozone depletion also are strong greenhouse gases. The substances that are developped to replace CFC-11 and CFC-12 since the Montreal Protocol, are much less effective in depleting the ozone layer, but they still are strong greenhouse gases. Further, ozone is by itself a greenhouse gas. Another link is the fact that the enhanced greenhouse effect cools the stratosphere, leading to increased formation of stratospheric clouds that catalyze ozone depletion (Austin *et al.*, 1992). Finally, increased UV-B radiation (caused by ozone depletion) has an influence on the marine biota (Denman *et al.*, 1996 in IPCC-I, 1995).

There are several other global and local environmental changes going on that are not yet considered in the IAMs although they do have indirect links with the climate system, the carbon cycle and the atmospheric chemistry of the other greenhouse gases. Examples are pollution of river-systems and oceans which might affect shelf sea and ocean biota, which in turn affect carbon fluxes in these systems. Further, changes in the geochemical cycles of phosphate and nitrogen might be significant by their eutrophication of continental shelf areas and also via the fertilisation effects on the terrestrial eco-systems, which in turn affect the carbon cycle (e.g Denman *et al.*, 1996; Melillo *et al.*, 1996 both in IPCC'95).

4.6. Exposure

Exposure is the next step in the causal chain. The changes in valued environmental components are linked to consequences via different exposure pathways. In their contribution to the IPCC Second Assessment Report, IPCC Working Group II distinguishes between sensitivity, adaptability, and vulnerability (IPCC-II, 1995). Following their definitions: Sensitivity is the degree to which a system will respond to a change in climate conditions. Adaptability refers to the degree to which adjustments are possible in practices, processes or structures of systems, to projected or actual changes in climate. Adaptation can be spontaneous or planned and can be carried out in response to or in anticipation of changes in conditions. Vulnerability defines the extent to which climate change may damage or harm a system. It depends not only on a system's sensitivity, but also on its ability to adapt to new climatic conditions. For instance, the consequences of sea-level rise or changes in distribution of malaria vectors or other climate-zone related diseases are highly determined by the sensitivity, adaptability, and vulnerability of the local systems. These are in turn affected by parameters such as welfare of a local society or fitness of a local population. An industrialized country that has enough money for a good coastal defence system will have less exposure to future sea level rise than an LDC country that has no money for adequate coastal defence. Regarding increased storm intensities, it may well be that future building technology results in super-storm-resistant houses, so that future storm damage stays within proportions. Future breeding (or genetic engineering) of drought-resistant crops may change the consequences of droughts on agriculture. Outdoor agriculture might be rare in a hundred years from now; so the vulnerability of future food production to climate change might change dramatically. A serious omission in current Integrated Assessment practice is that the question of what the future world on which the climate change will be imposed might look like, has not yet been addressed (personal communication Thomas Schelling, July 1996). Instead, the modeled climate change is imposed upon the current world. Examples of this are the RIVM studies on the possible future distributions of malaria (Martens, Rotmans and Niessen, 1994; Janssen and Martens, 1995) and of schistosomiasis (Martens, 1995). They themselves recognize major shortcomings of their studies: *"The extent of an increase in malaria risk will be superimposed upon change in malaria transmission associated with socio-economic development and the (in)effectiveness of control measures."* (Martens, Rotmans and Niessen, 1994) and *"among others, two of the limitations of the present model version are the non-inclusion of the impacts of socio-economic developments and land-use changes on the occurrence of malaria."* (Janssen and Martens, 1995). Martens (1995) adds another important constraint, namely "*in this study, one specific health impact of an anthropogenic-induced climate change is being investigated separately, although in many instances interactions between the various health effects of a climate change are possible if not probable (e.g. synergism between infectious diseases and levels of undernutrition) and they may accumulate in vulnerable populations."* To this, we can add uncertainties about future methods of coping with malaria as a result of technological and medical innovations. It cannot be excluded that malaria will have ceased to exist in a hundred years time.

4.7. Consequences

The last link in the causal chain is formed by the consequences such as floodings, shifting climate zones, changes in agricultural production, extinction of species, changes in ecosystems, changes in migration patterns, storm damage, effects on water supply, etc. Just as we discussed above with regard to exposure, it is uncertain whether (if not unlikely that) the set of consequences that is being valued to day will be the same in the future. Again, because of the fundamental unpredictability regarding what consequences will be valued by future societies, the implicit -but probably invalid- working hypothesis of current IAMs is that this set remains unchanged.

A more serious problem is that the prediction of the consequences of exposure to future changes in valued environmental components requires regional-specific information on future states of valued environmental components. However, regional prediction is not yet possible with current state-of-the-art of climate models, and the inherent predictability of climate diminishes as geographical scale is reduced, as can be seen from the following quotes from IPCC: *"Confidence is higher in the hemispheric-to-continental scale projections of coupled atmospheric ocean climate models than in the regional projections, where confidence remains low. There is more confidence in temperature projections than hydrological changes".* and *"Considering all models, at the 10^4 - 10^6 km^2 scale, temperature changes due to CO_2 doubling varied between +0.6°C and +7°C and precipitation changes varied between -35% and +50% of control run values, with a marked interregional variability. Thus, the inherent predictability of climate diminishes with reduction in geographical scale."* (Kattenberg *et al.*, 1996).

We further illustrate this massive problem with a few quotes from Henderson-Sellers (1996b): *"The urgent issue is the mismatch between the predictions of global climate change and the need for information on local to regional change, in order to develop adaptation strategies.*"; "*The dilemma facing policymakers is that scientists have considerable confidence in likely global climatic changes but virtually zero confidence in regional changes.*"; "*Unfortunately, climate models cannot yet deliver this type of regionally and locationally specific prediction and some aspects of current research even seem to indicate increased uncertainty.*"

Regional prediction of future soil moisture changes is essential in assessing the risks that climate change will bring to the food production system. This issue is topical in the light of Article 2 of the UNFCCC and the current climate negotiations on that convention. Several inter-comparison projects of GCMs show that the uncertainties in geographical distributions of future soil moisture are huge, and that after these GCM inter-comparison exercises the uncertainties turned out to be much higher than was assumed by the IPCC 1990 report (Henderson-Sellers, 1996b).

The economic impacts have to be modeled by economic models and cost-benefit analyses which also have large uncertainties. Such models are usually based on the concept of economic equilibrium. However, linked to the proof of economic equilibrium is the result that there are usually multiple equilibria (Sonneschein, 1974). Jaeger and Kasemir (1996) have argued that the existence of multiple equilibria casts severe doubts on the possibility of meaningful cost-benefit analysis concerning different climate policies. Also, the insight is growing that real economies are essentially in a far-from-equilibrium state (Giarini and Stahel, 1993, p.221). This adds to the perceived uncertainty since it makes our previously gained understanding of the dynamics of equilibrium economic systems inadequate.

It is generally perceived that consequences matter only in terms of what people value (although this is not tenable from the philosophical position of *inter alia* the deep ecologists). A widely acknowledged limitation of utility theory and welfare theory is the inherent impossibility to rank and aggregate utility and preferences in an objective way and the impossibility to objectively attribute value to consequences. Despite these notions, working group III of IPCC has attached monetary values to the costs and benefits of human intervention in the system. The issue of monetary valuation using methods such as "willingness to pay" is currently subject to vehement controversy. Monetary valuation has been criticized for its unfairness (the willingness to pay to prevent loss of life of one US citizen is much higher than the willingness to pay to prevent loss of life of one Bangladesh citizen, in other words, willingness to pay biases its conclusions in favor of projects that harm the poor); for the irrelevance of values based on preferences (willingness to pay cannot fully address the importance to human society of large scale ecosystem integrity); for the existence of better measures (e.g. indicators of ecological integrity); for its encouragement to exclude effects that are hard to measure; and for its distortion of non-market consequences to which it draws attention. For a more comprehensive discussion on the debate on monetary valuation and its uncertainties we refer to Toman (1996) and Adams (1995, 1996).

Several of the observations in the above analysis were also recognized at a 1993 IIASA workshop on the state-of-the-art in Integrated Assessment Models. As a result of this workshop Toth (1995) lists some major problems in uncertainty management faced by the IAM modelers community:

i) Ignorance: *"Our ignorance is vast at every single step in this process as we go from emissions to concentrations to climate forcing to changes in temperature and other climatic attributes over to the impact assessments and damage estimates."*

iii) Uncertainties don't reduce: *"Moreover, recent results in science that have been presented at the workshop do not provide much assistance in reducing these uncertainties. On the contrary they even add to the uncertainty of integrated models."*

iii) The climate system might be unpredictable: *"As presented at the workshop by Schotterer and Oeschger (1994), recent results from ice-core analyses suggests that the global climate system was rather unstable in various geological periods and that major climatic changes occurred in very short periods of a few decades in the complete absence of anthropogenic influence. This raises serious doubts about the predictability (and modelability) of the climate system in the first place."*

iv) Increase in detail means increase in unreliability: *"The clear danger in integrated assessments is that the more detailed and the more specific they get, the more unreliable the modeling results will be."*

v) Diversity of links makes uncertainty analysis increasingly difficult: *"Moreover, uncertainty analysis in large scale modeling is becoming increasingly difficult as a result of the modular structures and the diversity of linkages."*

In conclusion, a major problem with climate IAMs is that our present-day knowledge and understanding of the modeled system of cause-effect chains and the feedbacks in between is incomplete and is characterized by large uncertainties and limits to predictability. A closely related problem is that the state of science that backs the mono-disciplinary sub-models differs across sub-models. This implies that given the present state of knowledge, climate IAMs consist of a mixture of elements covering the whole spectrum from educated guesses to well established knowledge. It is also uncertain to what extent the model is complete.

The implications of these problems depend strongly on the application of the IAM. A policy maker who uses the model as a tool to help choose policy options suitable for achieving a given goal definitely needs good insight into the reliability of the outcomes. A scientist who uses the model to assess the relative importance of the uncertainties is less dependent on the overall reliability of the outcomes. The importance of uncertainty management is a function of the context of use and the objective of use of IAMs (compare Clark and Majone, 1985; Mermet and Hordijk, 1989; Beck *et al.*, 1996a, 1996b). In practice, IAMs are used with a mix of policy-oriented and knowledge-acquisition-oriented objectives and in a variety of contexts.

5. The usefulness and use of IAMs for the climate issue

In this section, we investigate the possibilities and limitations of climate IAMs to guide and inform the policy process. In the first part of this section we sketch the range of opinion on the (policy) usefulness, obtained from the literature and interviews with two modelers. Within the controversy on the policy-usefulness of climate IAMs, we seek to identify what applications have been agreed upon as valid. In the second part of this section we discuss the context in which IAMs are used.

5.1. The policy-usefulness of climate IAMs

In the previous section we found that climate IAMs are currently based on a mixture of knowledge that covers the whole spectrum ranging from educated guesses to well established knowledge. The uncertainties are large at every stage of the causal chain, and many scientific puzzles still have to be solved. Precisely because of these circumstances, some members of the international climate community feel that we might not be ready to link different aspects of the climate change issue together. The concern is that given the uncertainties in each of the individual components, linking it together would multiply uncertainties. This concern is widely heard in the debate on integrated assessment. This is what Henderson-Sellers (1996a) calls "uncertainties explode". Swart (1994) observed a similar concern among the social science community: "*Because of the large gaps in knowledge in the social sciences, in a recent report by social scientists of the US National Research Council, a strong warning is given against too much emphasis on the development and application of integrative models, encompassing both natural sciences and the human dimensions of global change (Stern et al., 1992). Climatologists dispute the inclusion of impacts in integrated models. This would misleadingly suggest a deterministic linkage between causes, physical effects and socio-economic impacts.*" Swart adds the observation that there is an increasing desire for integrating as many aspects of global change as possible and doing collaborative research, even if pertinent uncertainties continue to exist.

Leen Hordijk[1] (1994) sees a danger in the circumstance that parts of the IMAGE model anticipate scientific developments, instead of running parallel to or lagging behind scientific developments. *"For the climate case, the uncertainties are more numerous, bigger and more complex than with*

[1] In the eighties Leen Hordijk used to be project leader of the RAINS model at IIASA. In 1987 Hordijk moved to RIVM and from the sidelines became involved in the IMAGE project. Later Hordijk was appointed as a professor at Wageningen Agricultural University where he leads a research group on Environment and Climate (WIMEK). In 1992 he set up the IMAGE advisory board.

acidification. This can make climate IAMs ineffective in the policy process. With the RAINS model the situation was different. Acidification was an acknowledged problem in a large part of Europe and the major scientific puzzles concerning trans-boundary transport had been solved. The effects were already visible. So, you had a good starting point. Good models did not yet exist, but they were being developed in parallel with the scientific developments."

Hordijk sees three different strategies for doing integrated modeling: *"One strategy is to be ahead of the scientific development, this is what the IMAGE model did. IMAGE 1 was used with a signalizing function, and that worked. IMAGE 2 is getting more and more detailed, and then you have to be extremely careful not to be too far ahead of the scientific developments. A second strategy is to run parallel to scientific development. That is what RAINS did. The third strategy is to wait until science has matured and then develop an IAM. That is what happened in the US national acidification program (NAPAP). There, only after more than ten years of research, and that was done deliberately, one did start to construct an IAM. One of the directors of NAPAP absolutely disagreed with the way the RAINS model was realized and was used. He saw it as a mixing up of policy and research. The approach that he advocated was to wait until the science is complete, and only then do integrated modeling."*

Joe Alcamo[1] is a proponent of the use of climate IAMs to inform the policy process. He argues: *"We have global agreements to act on climate change and other environmental problems. 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. Our best understanding can be expressed in an IAM like IMAGE 2, and it can then be used to analyze policies. We recognize that our best current scientific knowledge may not be the best knowledge in the future."*

Alcamo thinks there is nothing wrong with an IAM being ahead of the scientific developments. He thinks it can stimulate discussion and speed up the process of model-improvement. In a response to an earlier version of our paper, Alcamo did not completely agree with Hordijks claim that RAINS ran parallel to scientific development. Alcamo thinks that both RAINS and IMAGE 2 were ahead of the accepted science. He provided the inclusion of nitrogen-transport in the RAINS model as an example: *"when hydrologists and soil scientists told us that is was imperative to include nitrogen transport and deposition into RAINS, atmospheric scientists told us that these calculations "were not ready".*

[1] Joe Alcamo was prominently involved in the development of both the RAINS model and the IMAGE 2 model. He used to be deputy project leader of the RAINS model at IIASA and until recently he was project leader of the IMAGE 2.0 project. Nowadays, Alcamo is a professor at the Center for Environmental Systems Research at the University of Kassel.

Nevertheless, a Polish scientist (Jerzy Bartnicki) and I went ahead and built a simple European-scale model of nitrogen deposition in Europe, published it in a journal, and included it as a submodel in RAINS. In my opinion this action stimulated other researchers with a better model to give us the nitrogen calculations we needed in RAINS." (E-mail message from Joe Alcamo to Jeroen van der Sluijs, 16-10-1996).

Alcamo's response to the concerns about the multiplication of uncertainties is: *"I argue that it is not always the case that uncertainty from one component, propagates to the next. It is also possible that uncertainty is dampened from one component to the next. I can give you an example. In the literature,* CO_2 *emission estimates for the year 2100 have a range of a factor of 50. But if a global carbon cycle model, is run with a factor of 50 difference in* CO_2 *emissions, it only produces a factor of 3 or 4 difference in atmospheric* CO_2 *concentrations. Furthermore, this factor of 3 or 4 variation in atmospheric concentration produces a difference of a factor of two or less in computed global temperature increase. So to a degree there is theoretical evidence that the global system integrates the variations found in its different components, and can actually dampen out fluctuations in different parts of the global system. In the end, this dampening process allows us to build integrated, multi-component models without having them suffer from unacceptably high uncertainties."* This view is however not widely accepted in the scientific community, because there is no *a priori* reason to assume that it would more likely for uncertainties to cancel out than to add up.

Parson (1994) stresses how useful IAMs can be to make rational informed social decisions, while he further claims that IAMs assist in the structuring of uncertainties: *"First, integrated assessment can help (indeed is necessary) to answer the broadest bounding question, how important is climate change. Second, IA can help assess potential responses to climate change, either with a benefit-cost framing that compares costs of responses to the impacts they prevent, or with a cost-effectiveness framing that assesses relative effectiveness and cost of different response measures to meet a specified target. Third, IA can provide a framework in which to structure present knowledge, providing several benefits. Perhaps the most important contribution is structuring of uncertainty and sensitivity: how well quantities and relationships are known, and how strongly valued outputs depend on them. Finally, integrated assessment can serve the longer-term goal of capacity building."* (Parson, 1994)

The usefulness of integrated modeling to assess uncertainties and to guide research is also stressed in the evaluation of the Dutch NRP: *"A second general issue is the role of this theme* [that is integration] *in guiding the research area, and the role of IMAGE in particular. IMAGE (through uncertainty assessment) can provide information on the relative importance of uncertainties on different areas which may be useful (used in conjunction with other information) in guiding the* *programme. However, issues like the scope for and research cost of uncertainty reduction and the policy-relevance of uncertainties (at different levels, national, international) need to be addressed. Notwithstanding this, the importance of IMAGE should not be overestimated."*

Having evaluated the state of the art of IAMs at an IIASA workshop, Toth (1995, p265) says on the subject of their usefulness *"If the building blocks are so shabby, is it worthwhile building integrated models at all? The answer is clearly yes, despite the present weaknesses of the models. The reason is that modeling forces us to reveal our assumptions and changing those assumptions shows how important they are with respect to the outcome."*

Morgan and Henrion (1990) also support integrated modeling: *"There are legitimate reasons for building large and complex models. Such models are justified when the details of the system are well understood* ***and*** *the inclusion of these details in the model is essential to the insight or answer that is sought."* Then, they notice that these criteria are not met for problems such as climatic change and conclude: *"Modeling any of these problems involves complex systems of coupled differential equations and large amounts of data to establish initial conditions. These models cannot be used to produce precise predictions like those of engineering design models. Rather they provide a vehicle for research on systems we do not yet fully understand."*. At present Morgan is deeply involved in a major climate IAM effort at The Center for Integrated Study of the Human Dimensions of Global Change at Carnegie Mellon University (NSF, 1996).

Hellström (1996) arrives at almost the same conclusion: *"The primary significance of models seems to be one of heuristics; once we dispense with the assumption that models are true depictors of the world 'out there', their value becomes that of guidance for researchers and policy makers. They become more of a policy instrument useful for the furthering of a science-policy dialogue than traditional scientific artefacts."*

Janssen and Rotmans[1] (1995) say something similar about the usefulness of climate IAMs: "*The models are meant to have an interpretive and instructive value rather than being prediction or "truth" machines.*". To stress the latter point, Rotmans and De Vries have chosen "*Insights, no answers*" as the title for their forthcoming book on IAMs.

In summary, the positions in the debate vary from "We are not ready to do integrated modeling, we have to wait until all science used in the model has the status of well established knowledge", to

[1] Jan Rotmans from RIVM developed the IMAGE 1 model in the period 1986-1990; he was prominently involved in the development of the ESCAPE model for the Commission of European Communities, which was the precursor of IMAGE 2, and he is currently project leader of the TARGETS model at RIVM.

"We have the responsibility to use our best scientific understanding to develop reasonable policies. Integrated modeling is the optimal way to combine our knowledge in such a way that we can evaluate the consequences of different policy scenarios, do cost-benefit framing or optimize cost effectiveness to reach a target." Apparently there is agreement that IAMs are no truth-machines and cannot reliably predict the future, but rather are heuristic tools. IAMs are capable of testing sensitivity, answering 'what if' questions (although each answer has to be followed by ", given the total set of assumptions of this model'), ranking uncertainties, ranking policy options, assessing the relative importance of uncertainties, identifying research priorities and providing insights that cannot easily be derived from the individual natural or social science component models that have been developed in the past.

Despite the fact that some experts maintain that we are not ready for integrated assessment, the models are being used at present to directly address policy questions, for instance by identifying 'safe emission corridors', which are presented to negotiators as answers rather than as insights. These cirumstances imply an urgent need for uncertainty management, quality assurance, high standards of IAM practice, and a high awareness of the limitations of models.

5.2. The context of use of IAMs

Mermet and Hordijk (1989) have presented a framework that correlates the role of assessment models to different kinds of policy contexts in which they are used and to the level of use. Although they inferred the framework as a result of a debriefing exercise concerning the use of the RAINS model in international negotiations, the framework is applicable to assessment in general.

Role of model	User level: Individual	User level: Collective (joint use by all parties)
Model as motor of the process	A party promotes the model as an active basis for its position	All parties agree to use the model as reference framework for the process
Model as a source of information	A party uses the model to complement the argumentation of its position	The model is considered by all parties as one source of information used in the process
Model indifferent because marginal or useless	A party is reluctant to move from the political to a more technical ground	The negotiation is so adversarial that "rational" analysis of the problem plays little role
Model undesirable	A party disagrees on the science or fights the model as a tactic in the negotiations politics	Prospects for the use of the models are terrible

Table 3. Types and levels of use of assessment models in the negotiation process (Mermet and Hordijk, 1989).

A realistic image of science should take into account its limitations and intrinsic uncertainties and the notion that scientific data do not necessarily correspond intrinsically to expert interpretation and policy conclusions, because these are 'underdetermined' by any scientific knowledge because of the repertoire of interpretive possibilities existing at each link in the argumentative chains (Van Eijndhoven and Groenewegen, 1991; Van der Sluijs *et al.*, 1997; Van der Sluijs and Van Eijndhoven, forthcoming).

This means that we have to go beyond the technocratic view in which science can provide an uncontestable scientific foundation for the policy process, which would correspond to the upper right corner of Table 3. On the other hand, the lower right corner of Table 3 ends at relativism in which science is of no use in decision making. This position is not compatible with what present day science **can** do, namely provide conclusive (though tentative) inter-subjective guidance for the ranking of plausibility and validity of theories on 'the world out there'. The position that is most compatible with both the potential and the limitations of science is in our view somewhere between '*the IAM serves as **one** source of information used in the process*' and '*the IAM serves as **the** reference framework for the process*'.

The position of models in integrated assessment is subject to debate. As we said before, in the second half of the eighties RIVM promoted integrated modeling as the optimal way of interfacing science with policy (Zoeteman, 1987). In the interim evaluation of the Dutch NRP, it was observed that *"The NRP does not yet have a well-developed approach to performing the assessment function (currently called integration). This function should involve synthesis, integration and communication of the program's research results for use by decision makers and also provide guidance to the other themes on emerging priority research needs."* and *"Despite its value as an assessment tool, exclusive emphasis on the IMAGE model is unlikely to enable the overall integration of the NRP. Additional assessment tools and approaches are needed."* (Science Policy Associates, Inc. and Holland Consulting Group, 1992). Three years later, in the final evaluation of the first phase of the Dutch NRP it was observed that integration and assessment within the NRP are still *"dominated by IMAGE"*, although the toolbox starts to broaden: *"In 1992, IMAGE was its only component; since then projects have been added under the categories "risk analysis" and "political context"."* (Science and Policy Associates, Inc., 1995).

The many limitations of climate IAMs are one of the reasons why Kasimir and Jaeger (1996) and Bailey *et al.* (1996) have recently redefined Integrated Environmental Assessment, incorporating a much wider coverage of activities than modeling only (see their definition in Table 1). They argue that *"IAM is not a complete IEA methodology. Integrated Assessment Modeling is an important activity within the boundaries of IEA but it is only part of the assessment, not the whole."* Parson (1996) makes a comparable case for broadening the toolbox with unconventional assessment methodologies. The second phase of the Netherlands NRP also has a broader interpretation of integrated assessment. They define it as a process in which a cluster of activities aims at optimizing the use of scientific knowledge for policy purposes. The activities include integration, risk analysis, policy analysis, and dialogue with policy and society (NOP-MLK, 1994). These recent developments marginalize the role of IAMs compared to the high hopes for IAMs in the past decade in which IAMs were propagated as **the** method for interfacing science and policy. One should keep in minde this changing position of IAMs in integrated environmental assessment when discussing the problems of uncertainty management in these models.

6. Uncertainty management in IAMs

In this section, we turn to a more general discussion of uncertainty management in IAMs. We analyze the mismatch between the types and sources of uncertainty that should be addressed on the one hand and the current practice of uncertainty management in IAMs and the available methodologies for addressing different types and sources of uncertainty in models on the other hand. Further we will look at the reasons for the mismatch and identify areas for improvement.

6.1 Stochastic versus deterministic modeling

Sub-models of IAMs can be either deterministic or stochastic. In deterministic models all parameters and variables of the model have point values at any given time. In stochastic models the parameters and variables are represented by probability distribution functions. In the case of Gaussian assumptions the vector of all parameters in the model can be represented stochastically by its first two statistical moments, a vector of mean values and an associated covariance matrix. Usually a vector of standard deviations of the parameters is used instead of a covariance matrix, due to lack of information on the nature of the co-variance matrix. Such a simplification conveys the implicit assumption that the parameters are stochastically independent and do not arise from a joint distribution, which might be unrealistic and must be taken into account when the results of subsequent uncertainty analysis are being evaluated (Young and Parkinson, to be published). There are intermediate models in which only some of the parameters are stochastic, or in which parameters and variables are represented by a two-fold range or a three-fold range rather than a probability distribution function.

Monte Carlo Simulation is the most powerful technique for stochastic modeling. In its simplest form, Monte Carlo traces out the structure of the distributions of model output by calculating the deterministic results (realizations) for a large number of random draws from the distribution functions of input data and parameters. This takes a lot of computer time. For that reason more advanced sampling methods have been designed that reduce the required number of model runs needed to get sufficient information on the distribution in the outcome. Latin Hyper Cube sampling is the most efficient method currently available. It makes use of stratification in the sampling of individual parameters and preexisting information about correlations between input variables (McKay *et al.*, 1979; Janssen *et al.*, 1991; Janssen *et al.*, 1994; Schimmelpfennig, 1996).

When Rotmans started with IMAGE 1 in 1986, he chose a deterministic approach, for several reasons. Information available on the distribution functions of most parameters and variables was, and still is, insufficient. Rotmans doesn't see any advantage in assuming normal distributions or uniform distributions for unknown distributions. To do this would suggest exactness which does not correspond to our current state of knowledge. Rotmans also stresses that distribution functions based on the best available knowledge change over time due to the progress of research. The dynamics in perceived uncertainties is tremendous (Rotmans, 1994). Alcamo (1994b) is also skeptical about stochastic IAMs: *"It is hard enough to build deterministic models. Making the model stochastic does not solve your uncertainty problem. It makes it more explicit for better or for worse."* To him it is far from clear whether Monte Carlo or any kind of stochastic simulation will work when analysing the uncertainty of a multi-component integrated model of the global environment. Further, he sees no reason why stochastical modeling should be *a priori* better than deterministic modeling followed by uncertainty analysis.

Weyant (1994) is more positive about stochastical modeling: *"This is a formidable task, but one that can be immensely valuable if completed successfully. Successful completion of this enterprise will require model integration and management that has rarely been achieved in the past."*

Young and Parkinson (to be published) strongly criticize the current practice in which stochastic dynamic models are the exception rather than the rule in environmental science research. According to them, "*simulation models of the environment based on deterministic concepts are more extensions of our mental models and perceptions of the real world than necessarily accurate representations of the world itself.*" They make a case for a new modeling approach that combines stochastically defined simulation models with data-based mechanistic models obtained from time series of observations. They proved the feasibility of this new modeling approach by applying it to the Enting-Lassey Global Carbon Cycle Model which is used by the IPCC. According to them: *"Until comparatively recently, MCS* [Monte Carlo Simulation] *was difficult to justify because it absorbed large amount of computing power and occupied unacceptable large computing time. Unless the model is very large, as in the case of the GCMs, this is no longer the case, however, since the rapid improvements in computer power have even made it possible to carry out MCS studies for most moderately sized simulation models on desk top personal computers. Larger models, such as the 26th order, non-linear global carbon cycle model considered in section 4, still presents problems for desktop machines or workstations but, as we shall* *see, the analysis can be carried out quite straight forwardly on parallel processing computers which are ideally suited to the generation of the multiple realizations required by MCS."*

Schimmelpfennig (1996) analyzed the representation of uncertainty in economic models of climate impacts and reviewed the methodologies available for characterizing uncertainty. He found that uncertainty is poorly represented in existing studies of climate impacts. He criticizes this from the notion that *"when only mean values are presented as results, most of the information about the underlying distributions of random variables has been discarded."* For the economic models, he came to the same conclusion as Young and Parkinson for the natural system models: *"What is needed are Monte Carlo type simulations".*

Another problem with stochastic models is the interpretation of the results. According to Rotmans, there is no means to represent the outcomes of stochastic models without introducing confusion. This also enhances the chance that results will be misused (Rotmans, 1994). Alcamo also stresses the problems that will arise in interpreting the outcomes from stochastic models: *"we haven't even figured out how to use that new explicit information"*. The result of Monte Carlo Simulation is a bundle of trajectories representing the distribution function of the outcome for a given input scenario. In theory, that information is valuable for identifying for instance to what extent the bundle of trajectories remains within a pre-defined 'safe corridor'. Then we can perform a goal-searching procedure by adjusting the scenario according to the discrepancy between the calculated bundle and the safe corridor. We can repeat this until we have e.g. the cheapest scenario that is safe enough (e.g. 95% of the bundle lies inside the corridor).

There are two major problems in applying such an approach. The first is that if the 'safe corridor' has more than two dimensions, graphical representation of the bundles of trajectories and the corridor is no longer possible. This implies that model-specific aggregated performance indicators have to be constructed to reduce the number of dimensions of the corridor to two or less. Such constructs give rise to a loss of information and rely upon the assumptions that have to be made to construct them. In other words, such indicators are tricky, and you have to do a lot of thinking about the consequences of your assumptions when interpreting your results. The second problem is that, unfortunately, computer resource limitations make the above-described procedure impossible in practice, because the computing time required for such stochastic goal-searching is almost infinite, especially if the 'safe corridor' is multi-dimensional (as is the case in the example of the IMAGE Safe Landing approach, which is a four-dimensional safe corridor). There are advanced stochastic goal-searching techniques which make it unnecessary to calculate the whole bundle of trajectories before an adjustment is made (see e.g. Ermoliev and Wets, 1988). Instead, the algorithm makes specific adjustments in the scenario after each single realization. The number of iterations to reach convergence is hardly larger than the number required to calculate one bundle. Yuri Ermoliev (Personal communication, 27-8-1996) is a strong advocate of stochastic search techniques. He criticizes scenario analysis for its endlessness (there is an infinite number of possible scenarios) and arbitrariness, and sees the stochastic search technique as a way to identify scenarios that meet a certain goal. He thinks the models should be used for generating solutions, not for making predictions and endless explorations of possible future developments. The stochastic search technique allows comparative analysis of solutions such as optimal structures of future societies (in terms of mix of technologies and practices to fulfill demands for goods and services), policy variables and decisions that are robust against the uncertainties. Further it can be used in hedging studies (Manne and Richels, 1991).

6.2. The modeling of surprise

"Much of the work to date has been based, implicitly or explicitly, on an evolutionary paradigm - the gradual, incremental unfolding of the world system in a manner that can be described by surprise-free models, with parameters derived from a combination of time series and cross-sectional analysis of the existing system. ... The focus on surprise-free models and projections is not the result of ignorance or reductionism so much as of the lack of practically usable methodologies to deal with discontinuities and random events. The multiplicity of conceivable surprises is so large and heterogeneous that the analyst despairs of deciding where to begin, and instead proceeds in the hope that in the longer sweep of history surprises and discontinuities will average out, leaving smoother long-term trends that can be identified in retrospect and can provide a basis for reasonable approximations in the future." (Brooks, 1986).

Surprise can play a role in every step of the causal chain (Table 2). Examples from the past are discrete events such as the oil shocks of 1973 and 1979; discontinuities in long-term trends, such as the acceleration of USA oil imports between 1966 and 1973; but also events that turn out to trigger or accelerate the policy process such as the 1988 US heat wave, the unprecedented damage (US$ 15,500,000,000) caused by super storm Andrew in 1992 (Property Claim Services, 1996) etc. The natural system also has surprises such as the volcanic eruption of Mt. Pinatubo in June 1991 which is believed to be responsible for the observed discontinuity in the trends in atmospheric concentrations of CO_2, CO and CH_4 and in temperature (McCormick *et al.*, 1995).

A further issue is that non-linear stochastic systems might have contra-intuitive future states which are missed in the model evaluation if the system representation is inadequate. Such an inadequacy can be the neglect of covariance and dependencies of parameters in stochastic models. In real-world stochastic complex systems, the variable probability values are constantly in flux. Further, the natural stochasticity in nature constantly alters the relationships between system components, and new external variables are added regularly, which change the natural conditions for the overall system. For instance, the introduction of human-made substances, such as CFCs, into the atmosphere has dramatically changed stratospheric chemistry. Furthermore, the emission of one compound changes atmospheric chemistry pathways of other compounds.

The simplifications made to model complex systems despite our limited understanding might well rule out certain characteristics of system dynamics such as the existence and nature of attractors in the system, which might be crucial in the evaluation of future behavior of the system. This was demonstrated by Van der Sluijs, De Bruyn and Westbroek (1996) who showed that scientifically tenable moderate changes in the assumptions of the BLAG model of the long-term carbon cycle can have a dramatic influence on the resulting qualitative model behavior, by introducing the existence of a transition of the system from point value equilibrium solutions of the system to a stable limit cycle solution when the system is exposed to external forcing.

The question of the possibility of modeling surprise was an important discussion point at the recent IIASA and LOS Centre Meeting "Climatic Change: Cataclysmic Risk and Fairness". Some of the participants argued that because most of the Earth Systems Models use smoothened, idealized and deterministic functional relations, a part of the potentially identifiable surprise is ruled out by the way the model is constructed, namely idealized smooth curves are used to represent relations between variables, whereas nature contains noise and proves time and time again to be much more capricious and erratic. For instance, the temperature record of the global mean temperature obtained from aggregated measurements and advanced reconstructions of past climates is non-smooth and is understood as a mixture of cyclical behavior on virtually all time scales (such as the diurnal cycle, the seasonal cycle, the El Niño Southern Oscillation, the 11 year and 22 year solar cycles, the 80-90 year solar cycle, the Milankovitch cycles of 22, 43 and 100 thousand years), trends, and irregular fluctuations. These fluctuations are usually called 'natural variability' of the climate. The trends and the cyclic behavior can be modeled with smooth assumptions, the irregular part cannot.

Recently, a new bottom-up modeling technique for complex adaptive (social) systems has been developed, called agent-based modeling, that can to a certain extent be used to model some of the aspects of surprise. The method has been demonstrated with a traffic model for the city Albuquerque. Travel behavior and decision-rules of every single inhabitant of the city (the agents) are modeled, together with the road network. The resulting aggregated traffic patterns over the time of a day show the build-up of morning rush-hour traffic and resulting traffic jams. The hope is that a deeper understanding of how complex adaptive systems work will suggest the right type of mathematical structures and lead to a decent theory of these processes, which this new school of modeling believes will ultimately lead to the increased predictability of surprises, such as traffic jams (Casti, 1996a; Casti, 1996b). A somewhat comparable bottom-up approach for modeling the biosphere has been proposed by Westbroek and Muyser (1992), but their approach has not yet been demonstrated with an operational model.

The solution of unknowns cannot simply be "let's put every single detail that we know in the computer and let's hope that something shows up". In our view, what is needed for scientific progress is to ask the right questions, not to produce an endless stream of 'what-if answers'. In that sense computers are useless, because they cannot ask questions, they only produce answers. We think that at least in the current state of the art, the usefulness of agent-based modeling and bottom-up biosphere modeling for IAMs is questionable. The agent-based approach can be of some use in a broader process of integrated environmental assessment to tentatively model decision systems or future technology choice in order to obtain some kind of feeling for its possible dynamics.

Given the absence of adequate methodology to model surprise, a systematic search for examples of non-linearities from the past might be the prelude to a search for possible future surprises (Brooks, 1986). Other strategies that can help us to understand surprise are focusing on the underlying principles of surprise, which is what happens in surprise theory (Holling, 1986) and systematic 'thinking the unthinkable' through imagining unlikely future events followed by the construction of plausible scenarios by which it might be realized (Kates and Clark, 1996).

Non-smoothness introduces a problem in sensitivity and uncertainty analysis because classic uncertainty analysis is based on smooth systems. Sensitivity analysis of non-smooth systems is a special topic that deserves more attention. Such analysis should focus on the identification of (thresholds in) indicators which could be used to predict jumps in the system and discontinuities in trends.

6.3. Addressing various sources of uncertainty

In the literature a variety of different classifications of uncertainty can be found which we have summarized in Table 4.

Table 4. Classifications of uncertainty

Vesely and Rasmuson, 1984	1. Data uncertainties (arise from the quality or appropriateness of the data used as inputs to models); 2. Modeling uncertainties

	a. incomplete understanding of the modeled phenomena; b. numeral approximations used in mathematical representation; 3. Completeness uncertainties (all omissions due to lack knowledge).
Environmental Resources Limited, 1985	Errors in modeling: a. process error (due to model simplification); b. functional error (uncertainty about the nature of the functional relations); c. resolution error; d. numerical error.
Hall, 1985	1. Process uncertainty; 2. Model uncertainty 3. Statistical uncertainty 4. Forcing uncertainty (involved in predictions which presuppose values that are unknowable)
Morgan and Henrion, 1990	1. Sources of uncertainty in empirical quantities a. Statistical variation and random error b. Subjective judgement and systematic error c. Linguistic imprecision d. Variability e. Inherent randomness and unpredictability f. Disagreement g. Approximation 2. Uncertainty about model form
Funtowicz and Ravetz, 1990	1. inexactness (significant digits/error bars) 2. unreliability 3. border with ignorance
Wallsten, 1990	1. Ambiguity (confusion in communication, avoidable) 2. Vagueness (imprecision in meaning)

	3. Precise uncertainties (objective and subjective probability)
Wynne, 1992	1. Risk (know the odds); 2. Uncertainty (don't know the odds); 3. Ignorance (don't know what we don't know) 4. Indeterminacy (open-ended causal chains or networks)
Helton, 1994	1. Stochastic uncertainty (arises because the system under study can behave in many different ways. It is a property of the system.) 2. Subjective uncertainty (arises from a lack of knowledge about the system. It is a property of the analysts performing the study.)
Hoffman and Hammonds, 1994	1. Uncertainty due to lack of knowledge 2. Uncertainty due to variability
Rowe, 1994	1. Four dimensions of uncertainty: a. Temporal (uncertainty in future states/ past states) b. Structural (uncertainty due to complexity) c. Metrical (uncertainty in measurement) d. Translational (uncertainty in explaining uncertain results) 2. Variability is a contributor to uncertainty in all dimensions. Sources of variability: a. Underlying variants - inherent to nature - that contribute to the spread of parameter values i. apparent inherent randomness of nature ii. inconsistent human behavior iii. nonlinear dynamic systems (chaotic) behavior b. Collective / individual membership assignment c. Value diversity

On the basis of on Vesely and Rasmuson's (1984) classification for sources of uncertainty and Funtowicz and Ravetz' (1990) classification for types of uncertainty, Van der Sluijs (1995) has proposed a two-dimensional classification scheme defining areas to be addressed in uncertainty management in IAMs. This scheme is presented in Table 5. Inexactness refers to error-bars, probability distribution functions, multiple tenable model structures etc. Unreliability refers to the level of confidence, quality, soundness, scientific status etc. of the knowledge on inexactness. Ignorance refers to all 'don't know what we don't know'. Funtowicz and Ravetz talk about the border with ignorance rather than ignorance because by definition we cannot say anything useful about that of which we are ignorant, *"but the boundless sea of ignorance has shores which we can stand on and map."*

Table 5
Two-dimensional classification scheme for uncertainties, defining areas to be addressed in uncertainty management in IAMs.

source / type	input data	model structure		model completeness
		parameters	relations	
inexactness				
unreliability				
ignorance				

We will now briefly discuss methodologies and strategies for uncertainty analysis and uncertainty management in IAMs, using the classification scheme of Table 5.

6.3.1. Dealing with uncertainties in input data and model parameters

There are very sophisticated tools available for sensitivity analysis and uncertainty analysis in models (Janssen *et al.*, 1990, Janssen *et al.*, 1994). A widely used technique is Latin Hypercube Sampling, which is an advanced Monte-Carlo-based tool to map the relative contribution of specified uncertainty in model input and model parameters to the spread in key model outputs. If applied to the entire model, these techniques also assess the propagation of uncertainties through the model. In terms of Table 5, these techniques map the uncertainty in model outcome resulting from inexactness in data input and parameters. A major problem in determining total inexactness uncertainty by Monte Carlo Simulation is the identification of the spread and distribution functions of all input data and model parameters (Rotmans, 1994). A substitute for the lack of knowledge on the distribution functions is the use of subjective probability functions which are obtained by combining expert judgments. Titus and Narayanan (1996) have demonstrated this technique to identify the distribution functions in input data and parameters for their Monte Carlo analysis of a sea level rise model. The IIASA population project has also used this technique to develop probabilistic world-population projections (Lutz *et al.*, 1996).

Although subjective probability is an imperfect substitute for established knowledge, and despite the problems of aggregation of expert judgement (see e.g. Keith, 1996), if nothing better is available it is better to use subjective probability distributions than deterministic point-values so that you have at least a first approximation of the uncertainty. In our view, Titus and Narayanans 'Delphic Monte Carlo' approach is a shining example of uncertainty assessment which deserves to be followed.

Another problem is that sensitivity analysis and Monte Carlo modeling are resource consuming (time, money, research capacity, computer capacity): "*Today's computers cannot handle the complexity of some of the critical processes being modelled. For example, simulating atmospheric feedback processes alone requires a large program with thousands of input variables. Linking such a program to a model of ocean processes produces a model that is even bigger and more complex. Using a coupled ocean atmosphere model to perform a single several-century simulation might take a year on a super computer. Exploring the effect of a single uncertain input parameter would require 50 or 100 runs - a task so expensive and time-consuming that it precludes any formal uncertainty or sensitivity analysis in such models.*" (E-lab January-March 1995, MIT).

For the IMAGE-2 model, Monte-Carlo-based uncertainty analysis has so far been carried out on a few sub models only (Hordijk, 1993; Krol and Van der Woerd, 1994). Further, in order to obtain insight into error propagation in the entire integrated model one needs to do more than simply apply Monte Carlo modeling solely to isolated sub models. The IMAGE example shows that even in the area of uncertainty management where adequate tools are currently available (namely the upper left corner of Table 5), the mapping of uncertainties is only partly done because of lack of resources and competing priorities (Van der Sluijs, 1995).

6.3.2. Dealing with uncertainty regarding model structure and functional relations

Quantitative assessment of spread in key model outcomes caused by uncertainty regarding the relations in the model that represent the modeled processes is an uncultivated area in IAM modeling. This is also where the issue of model validation comes in.

We use a taxonomy for errors in model structure, based partly upon (ERL, 1984). The taxonomy distinguishes (1) functional error, (2) errors introduced by the technique of modeling, including process error, technical fix error, resolution error and aggregation error, and (3) bugs, including numerical error, programming error, and hardware error. We will discuss each category below.

Functional error

Functional error arises from uncertainty about the nature of the process represented by the model. Uncertainty about model structure frequently reflects disagreement between experts about the underlying scientific or technical mechanism. In practice, functional error is almost only assessed indirectly through inter-IAM comparison. The rare exception is the approach chosen by the TARGETS modeling group at RIVM. TARGETS uses an innovative approach in uncertainty analysis. The subjective component in uncertainty is operationalized through well-defined cultural perspectives that guide the choice of values for input data and parameters but also for different model relations. The three cultural perspectives are called Hierarchist, Egalitarian, and Individualist, which is a simplification of the original group-grid classification by Douglas and Wildavsky (1982) and Schwartz and Thompson (1990), who distinguish 2 more categories (the Fatalist and the Hermit). The axiom is that each perspective consists of a different myth of nature (fragile, robust, or robust within limits) and a different management style for managing the risk (prevention, adaptation or control). Cultural theory has been criticized for its oversimplification of reality, for being too static (in reality one can be a hierarchist at work, a fatalist in leisure time, and an egalitarian at home), for its undue universal claims (whereas, in reality, one can for instance act as a hierarchist when confronted with problem A, and as an egalitarian when confronted with problem B) and for not taking account of complex systems of myths of nature (for a critique on cultural theory, see Trisoglio, 1995). In the TARGETS model, cultural theory is used as a rationale for non-random sampling from the almost infinite number of combinations of tenable (discrete) choices for model assumptions and parameter values and distributions, in order to provide consistent routes through the tree of (discrete) choices constituted by the interpretive space in present-day knowledge. It is, however, questionable whether cultural-theory sampling really differs from, or provides better insights than, for instance the more traditional 'best-case/mean-case/worst-case' rationale for non-random sampling. Further, non-random sampling structurally ignores a large number of possible routes through the tree of choices. Consequently, cultural-theory sampling structurally ignores a part of the existent uncertainty. The only way to reveal the total distribution of the outcome associated with the total set of model assumptions and their specified uncertainties is Monte-Carlo analysis with random sampling or Latin Hyper cube Sampling[1]. This does not mean that cultural-theory sampling is useless, but its value lies somewhere else. Cultural-theory sampling allows sensitivity studies of the type: "suppose I choose a hierarchist management style for managing the risk of climate change, based on the sampling from the uncertainty ranges and alternative model structures that best correspond to the

[1] On the other hand, completely-random sampling runs the risk of overestimating the total uncertainty in a model result, because it can include realizations based upon combinations of parameter values that are unlikely or impossible in reality.

hierarchist myth of nature, what then will happen if the hierarchist myth is wrong and nature acts instead according to the egalitarian myth?" The method is in an experimental phase and has been tested on a few of the sub-models of TARGETS (Van Asselt *et al.*, 1995; Van Asselt and Rotmans, 1996; Van Asselt, Beusen and Hilderink, to be published).

Process error

Process error arises from the fact that a model is by definition a simplification of the real system represented by the model. Examples of such simplifications are the assumption of constants for entities that are functions in reality, or focusing on key processes that affect the modeled variables but omitting processes that play a minor role and are considered not significant. What processes are considered relevant is often related to the time-scale of the model. For instance, for very long term-carbon cycle modeling, many mechanisms that are dominant on shorter time-scales become insignificant, whereas mechanisms that are dominant on the very long time-scales, such as the carbonate silicate geochemical cycle, are insignificant for models of the shorter time scales (see e.g. Van der Sluijs, De Bruyn and Westbroek, 1996). Another example is that if GCMs simulate climates that are far removed from the current climate, they become inadequate e.g. with an increase of 5.5°C you might completely wipe out antarctic sea-ice, with massive changes in physical processes affecting climate. This mechanism is not yet taken into account by the models (see Van der Sluijs *et al.*, 1997).

Process errors can also occur from the way in which sub-models in IAMs are linked. In most of the current IAMs feedbacks between variables and relations located in different sub-models are not evaluated after each time step of integration of the model. In other words, the models are not really integrated.

The magic key word in handling this type of uncertainty is model validation. It has however been argued that model-validation is in principle impossible (Beck *et al.*, 1996a; Beck *et al.*, 1996b). For that reason Konikow and Bredehoeft (1992, cited in Beck *et al.* 1996) prefer terms such as *"model testing, model evaluation, model calibration, sensitivity testing, benchmarking, history matching, and parameter estimation."* Beck *et al.* (1996) discussed the problems of model validation and observed that "*The difficulty of model validation tends to increase as the degree of extrapolation from observed conditions in the past increases. And not surprisingly, the greater the degree of extrapolation so the greater is the necessity of relying on a model for the conduct of an assessment.*" (See also Beck, 1994).

Toth (1995, p226) has proposed three routes for model verification:

1. check against historical records;
2. adoption of models and codes from other modeling groups for conceptual verification; and

3. model inter-comparisons.

For most sub-models of IMAGE, the data from 1970-1990 were used for model verification. In 1993, the sub-models for agricultural demand, terrestrial vegetation, land use change and atmospheric composition were tested against historical data (Hordijk 1993). In 1994, the IMAGE team started to construct a hundred-year historical scenario and database for model verification (Batjes and Goldewijk, 1994). A major IAM inter-comparison that included IMAGE was carried out by the Energy Modeling Forum (Energy Modeling Forum, 1994a, 1994b, 1995a, 1995b; Tol, 1994, 1995).

Resolution error

Resolution error arises from the spatial and temporal resolution of the model. The possible error introduced by the chosen spatial and temporal resolutions can be assessed by means of sensitivity analysis. However, this is not as straightforward as it looks, since the change of spatial and temporal scales might require significant changes in the model structure. For instance, going from annual time steps to monthly time steps requires the inclusion of the seasonal cycle of insolation.

Aggregation error

The scaling up or scaling down of variables to meet the required aggregation level of the sub-modules they feed into, is another possible source of error. In cases of non-additive variables, the scaling-up or scaling-down relations are always to a certain degree arbitrary (Hellström, 1996). For instance, so that it can take biospheric feedbacks between climate and vegetation into account, IMAGE uses a down-scaling procedure to produce local climate in one grid cell. In each time step, the interaction between the down-scaled climate in the grid cell and the vegetation type in the grid cell is evaluated, followed by an upscaling step from the grid level to regional figures by means of aggregation. The credibility and soundness of this high resolution two-dimensional modeling practice to produce climate on a grid scale is subject to scientific controversy because of the difficulties of regional climate predictions with our present state of knowledge. In the evaluation of the first phase of the Dutch NRP it was concluded that *"A crucial area of weakness in IMAGE is the regionalization of climate change. Previous reviews recommended abandoning the 2D model, but this does not seem to have been done."* The reason that it was not abandoned is that the IMAGE modelers believe that the methodology is scientifically sound, and that it is the only available way to dynamically include the feedbacks between vegetation and climate. Alcamo (1994b) maintains that unbiased errors on a geographic grid sometimes cancel each other out when gridded data are aggregated up to regional averages. He believes that as long as you don't see the grid specific calculations as predictions, and only use the regionally aggregated figures, the uncertainty becomes acceptable.

Technical fix error

Technical fix error arises from the introduction of non-existent phenomena to bridge the mismatch between model behavior and observation and or expectation. An example is the flux adjustment in coupled Atmosphere Ocean General Circulation Models used for climate projections. Flux adjustment is a technical fix in coupled atmospheric and ocean GCMs to get round the problem that the heat fluxes between the ocean and the atmosphere calculated by the two models are different. Without such adjustments the reference climate undergoes drifts. For a critical review of the issue of flux adjustment we refer to Shackley *et al.* (1996).

Many forms of calibration bridge the gap between model and observation using non-physically based technical fixes. The effect of such fixes on the reliability of the model outcome will be bigger if the simulated state of the system is further removed from the (range of) state(s) to which the model was calibrated.

Numerical error

Numerical error arises from approximations in numerical solution, rounding of numbers and numerical precision (number of digits) of the represented numbers. When we asked Leen Hordijk how these uncertainties are being assessed in the IAM practice, he said that *"in the models I have seen there is not much attention paid to this. But there are not very many numerical approximations. The models are usually simple and linear. Behind these linear equations there is sometimes non-linearity. I agree that the IMAGE model should pay attention to numerical errors. But there is always the tension between further development of the model, its use in policy making, and other issues in the modeling team. It is also a matter of money.*"

Programming error

Programming error arises from bugs in software, design errors in algorithms, type-errors in model source code, etc.. Here we encounter the problem of code verification which is defined as (ASTM E 978-84, cited in Beck *et al.*, 1996): examination of the numerical technique in the computer code to ascertain that it truly represents the conceptual model and that there are no inherent numerical problems in obtaining a solution.

If one realizes that some of the models have hundreds of thousands of lines of source code, errors in it cannot easily be excluded and code verification is difficult to carry out in a systematic manner. In our interview with Hordijk we obtained some examples of programming errors in the RAINS model: *"In the RAINS model I once encountered a serious error. You could view the RAINS model not only for the whole of Europe, but also for a geographical rectangle. If you take for instance West Germany, and you put a rectangle over it, then you have not only West Germany, but also a part of the Netherlands, a part of East Germany, a part of Switzerland, a part of the Czech republic. The old soil model showed a bar-diagram for Germany indicating the pH distribution of the soil. In addition it gave a bar-diagram of the entire rectangle. That one was different, of course. For England the diagrams should be equal, because the rectangle did not include surrounding countries. To my surprise this turned out not to be the case when I tried it once. Max Posch, who did the computer programming at that time, searched for a very long time. Finally it turned out that for a big part of Europe the file with soil-types - we distinguish 88 soil types - was read the wrong way round. Consequently, the diagram of the rectangle over England used the soil data of a different recangle, which containined a part of France. So all the diagrams for any rectangle in Europe were wrong."*

And another example from RAINS mentioned by Hordijk: *"It hasn't had any effect on the results, but it was discovered only recently, although the model has existed already for 10 years or so. If one chooses Bulgaria from one of the menus from RAINS, then Rumania appears on the screen. And if you chose Rumania, then you get Bulgaria on the screen. So, somewhere in the early period of RAINS someone put a wrong pointer in the software. And once such a map has been made, it is passed from one version to the next version of the model. The error goes with it and can remain undiscovered for years. Nobody can guarantee that nothing else is is not yet discovered, which could have influence on the results. But you can minimize that risk, by what I did as project leader. When my team had finished a new version of the model, I sat for many days behind my PC, running the model with many different scenarios, the one just a bit different from the other, and then comparing the results; a sort of a on-line sensitivity analysis."*

The latter method is what Hordijk calls the 'rack method', which he learned from Professor Somermeijer: *"That means, you take the model and you enter very extreme values and see what happens. Repeat this for a range of extreme values. If the model does not exhibit strange behavior, there is a fair chance that the model is stable. If it does show strange behavior, then you've got to search. If it exhibits strange behavior for values that deviate significantly from the default settings but that are not really extreme, then you have a serious problem. That is a rough first recipe, but it does work."*

To secure against undiscovered bugs in the compiler-software, one can test the sensitivity of critical model outputs to the choice of the compiler-software used to compile the source code. Just use another compiler and see if the result is reproducible. In practice, one often encounters portability problems when trying to compile the source code with another compiler, due to a diversity of standards in programming languages, which means extra work (adjustment of the source code to match the standard of the compiler).

Hardware error

Hardware error arises from bugs in hardware. An obvious example is the bug in the early version of the Pentium processor for personal computers, which gave rise to numerical error in a broad range of floating point calculations performed on that processor. The processor had already been widely used worldwide for quite some time, when the bug was discovered. It cannot be ruled out that hardware used for IAMs contains undiscovered bugs that might affect the outcomes, although it is unlikely that they will have a significant influence on the models' performance. To secure against hardware error, one can test critical model output for reproducability on a computer with a different processor before the critical output enters the policy debate.

6.3.3. Dealing with uncertainties regarding model completeness

The IAM can have omissions at three levels: causes, processes, and impacts. We might overlook or underestimate anthropogenic causes of climate change, e.g. water emissions by aircraft in the upper atmosphere; indirect climate effects of perturbation of geochemical cycles other than carbon (e.g. phosphate) via its effect on the biota; climate effects of ocean pollution via its effects on the biota etc. We might overlook greenhouse gases. For instance, the greenhouse effect of SF_6 entered the assessments only recently (Cook, 1995), whereas the greenhouse gas NH_3 is not included in any current climate risk assessments. NH_3 was mentioned as an anthropogenic greenhouse gas in earlier assessments (Wang *et al.*, 1976; Hekstra, 1979; Schuurmans *et al.*, 1980; RMNO, 1984, see also Van der Sluijs and Van Eijndhoven, forthcoming) and it plays a crucial role in geological models of the earths' early atmosphere (e.g. Margulis and Lovelock, 1974). NH_3 has a pre-industrial ambient concentration of 6 ppbv and an atmospheric life-time of about one week in the present day atmosphere (figures from Margulis and Lovelock, 1974 and Wang *et al.*, 1976). Its concentration is therefore not well-mixed over the globe, but in the immediate environment of permanent emission sources such as intensive cattle-breeding it might significantly affect the local radiation balance and hence the local climate.

Further, we overlook some processes: unknown feedback loops and supposed feedback loops that are not yet mathematically representable due to lack of knowledge are not included. It is possible that important modules are missing because scientists and modelers are not (yet) aware of their importance. We know that not all relevant substances and processes in complexly coupled atmospheric chemistry pathways of greenhouse gases have been identified. The uncertainties in the carbon budgets still allow for the existence of missing sinks. Recently it was discovered that fungi play an important role in the CH_4 cycle in peats because they effectively fix CH_4 emitted from lower peat layers (personal communication Mark Kilian, May 1996). This might affect our understanding of what could happen if the permafrost start thawing and the CH_4 fixed in clathrates will be emitted to the atmosphere. We might also overlook impacts of climate change on aspects that are valued by actors who are not yet participating in the assessment projects. Currently, the macro-economic-oriented IAMs are dominated by economists, whereas the process-oriented IAMs are dominated by natural scientists. According to Kasemir *et al.* (1996), *"scientific modeling without sufficient input from public discussion risks focusing on irrelevant issues while ignoring questions of interest to the public."* Examples of omissions at the impact side of IAMs can be the impact on human migration patterns and environmental refugees or the extinction of species.

Systematic methodologies for assessing model completeness directly are not available. It can only be addressed indirectly by quality control processes such as advisory boards, peer review, model inter-comparison, competition between modeling groups, etc. We will discuss these in the next section.

6.4. Quality control in IAM practice

The lack of quality control and good scientific practice in policy-oriented modeling was already stressed in 1984 by Keepin and Wynne (1984). Their twelve-year-old findings on energy modeling are still very topical and highly relevant to current IAM practice. Keepin and Wynne (1984) analyzed the energy models of IIASA and found that despite the appearance of analytical rigor, IIASA's widely acclaimed global energy projections were highly unstable and based on informal guesswork. According to Keepin and Wynne, this results from inadequate peer review and quality control, which raised questions about political bias in scientific analysis. They concluded amongst other things: "*First, many crucial components of the scenarios are generated informally and supplied as inputs to the formal computer models, which then reproduce these projections with only minor alterations. Thus, the models have not analytically "discovered" feasible energy futures. Indeed, despite the appearance of analytic sophistication and rigor, the models serve primarily as a static accounting framework of the analysts.*" and "*One important lesson is that the most exquisite formal analytic modeling still embodies informal* *assumptions (often about sociopolitical values and institutional behavior) that affect what technical outcomes are conceivable. Peer review of formal models can expose these assumptions for external debate and evaluation. Indeed, rather than attempting to identify objective policy truths, perhaps a more realistic role for policy modeling is to explore origins and consequences of different social and institutional assumptions. Such an approach would embrace (rather than deny) the interpenetration of science and politics in policy analysis.*". More than ten years later Toth (1995) still concludes: *"My review of current integrated assessments indicates the emerging need for a systematic and critical appraisal."*

J. Ravetz (E-mail message to J. van der Sluijs, 8 July 1996) suggests that the lack of quality control and good scientific practice might have three causes:

"1. Some people have effective "good practice" but their standards are not diffused among all practitioners. This is quite common in science, where a "leading" lab can get genuine results that few others can emulate.

2. The discussion of "good practice" is of the sort I call "lamp-posting", from the old story about a man who was seen by a neighbour in the early hours of the morning, crawling on the ground near the lamp-post. Asked what he was doing, he replied that he was looking for his keys. "Did you drop them there by the lamp-post?", the neighbour asked. "No, near my front door". "Then why are you looking near the lamp-post?" "Because at least it is light here, so if they were here I would find them." (I saw this recently as a joke about a drunken man; but it is also quoted as an original Sufi story!)

Translated into practical terms, this means that the researchers concentrate on the soluble problems, even if the insoluble ones are more important. I got this impression from the discussion with Alcamo.

3. Finally, there is the possibility that discussions of quality are only a game. This might be played for political advantage within the field (who can demolish the other's research more effectively?), or to comply with external requirements."

To Ravetz's Number 1, we can add the cost-problem. We saw that despite the availability of adequate tools for uncertainty assessment such as Monte Carlo Simulation, these are hardly being used in IAMs yet, due to their resource-consuming character. For the issue of cost-benefit analysis of the costs of generating more information on uncertainty and the benefits of this information, we refer to Hirshleifer and Riley (1992).

We found support for Ravetz's Number 2 in the evaluation report of the Dutch NRP, where it was observed that the program focused on the strengthening of existing areas of excellence in Dutch climate research, rather than on the key questions that should be answered (Science and Policy Associates, 1995).

Ravetz's Number 3 is supported by the 'backlash phenomenon' which is most prominent in the US, where industrial organizations such as the Western Fuels Association fund research to actively undermine the scientific credibility of the IPCC, while on the other hand the IPCC does its best to enhance its own credibility by maximizing representativeness and maximizing the process legitimacy of the consensus-building process (see e.g. Lunde, 1991). Recently the greenhouse 'rebels'[1] (although some of the established IPCC scientists would call them 'cranks') organized themselves in the European Science and Environment Forum that *"will seek to provide a platform for scientists whose views are not being heard, but who have a contribution to make"* (Emsley, 1996). They strongly criticize the 'science by consensus' approach of the IPCC and have issued a book in which almost every link in the IPCC chain of arguments is challenged (Emsley, 1996).

We should keep in mind that **all** actors with a stake in global warming have agendas of their own and are not always averse to manipulating uncertainty for various reasons. Uncertainties are often magnified and distorted to prevent public insight into the policy-making process and to obstruct the policy process (see e.g. Helström, 1996). The uncertainty question can be (and is) actively used as a strategy to end up in the lower right-hand corner of Mermet and Hordijks' diagram (Table 3).

Clark and Majone (1985) have designed a taxonomy of criteria for quality control of policy-oriented science (Table 6). The taxonomy acknowledges that each actor that has a stake in quality control has a different critical role. Further, they distinguish three critical modes: input, output and process. As Ravetz (1986) stressed ten years ago, mastery of Clark and Majone's Table would make an excellent introduction to the methodological problems of policy-related science.

Shackley and Wynne (1995) showed that the criteria for good scientific practice with respect to climate research are not solely determined from within science itself. Most of them emerge by a process of mutual construction with government policy institutions. According to them, this may now risk inadvertently foreclosing the consideration of potentially significant alternative scientific research and policy approaches. Boehmer-Christiansen (1994a, 1994b, 1994c, 1995) stresses an even more dominant role for institutions in this mutual construction of criteria for good scientific practice. Funtowicz and Ravetz (1993) have made a case for opening up the issue of good scientific practice through what they call 'extended peer communities'.

---

[1] The terms 'rebels' and 'cranks' stem for Funtowicz and Ravetz (1990), see also Table 8b. 'Rebels' have some standing among their colleagues, whereas 'cranks' have none. Who is a 'crank' and who a 'rebel' may be time bound.

Since the IMAGE 1 model was developed with virtually no external quality control, the model received only low peer acceptance. IMAGE 2 did better, partly because it was incorporated in the Dutch NRP, partly because it had learned from the criticism leveled at IMAGE 1 and partly because Joe Alcamo brought with him his experience of the RAINS model where peer review was common practice.

Table 6 A taxonomy of criteria for quality control of policy-oriented science (Clark and Majone, 1985)

	Critical mode		
Critical role	Input	Output	Process
Scientist	Resource and time constraints; available theory; institutional support; assumptions; quality of available data; state of the art.	Validation; sensitivity analyses; technical sophistication; degree of acceptance of conclusions; impact on policy debate; imitation; professional recognition.	Choice of methodology (e.g., estimation procedures); communication; implementation; promotion; degree of formalization of analytic activities within the organization.
Peer group	Quality of data; model and/or theory used; adequacy of tools; problem formulation; input variables well chosen? Measure of success specified in a advance?	Purpose of the study; conclusions supported by evidence? Does model offend common sense? Robustness of conclusions; adequate coverage of issues?	Standards of scientific and professional practice; documentation; review of validation techniques; style; interdisciplinarity
Program manager or sponser	Cost; institutional support within user organisation; quality of analytic team; type of financing (e.g. grant versus contract)	Rate of use; type of use (general education, program evaluation, decision making etc.); contribution to methodology and state of the art; prestige; can results be generalized, applied elsewhere?	Dissemination; collaboration with users; has study been reviewed?
Policy maker	Quality of analysts; cost of study; technical tools used (hardware and software); does problem formulation make sense?	Is output familliar and intelligible? Did study generate new ideas? Are policy indications conclusive? Are they constant with accepted ethical standards?	Ease of use; documentation; are analysts helping with implementation? Did they interact with agency personnel? With interest groups?
Public interest groups	Competence and intellectual integrity of analysts; are value systems compatible? Problem formulation acceptable? Normative implications of technological choices (e.g. choices of data)	Nature of conclusions; equity; analysis sued as rationalization or to postpone decisions? All viewpoints taken into consideration? Value issues	Participation; communication of data and other information; adherence to strict rules of procedure

A problem in quality control is that due to the resolution and the aggregation level, IAMs contain many parameters (and other constituents) that are constructs resulting from the simplification process and are hence not well-specified, that is, they fail to pass Howard's clarity test. Howard's clarity test reads (cited in: Morgan and Henrion, 1990): *"Imagine a clairvoyante who could know all facts about the universe, past, present and future. Could she say unambiguously whether the event will occur or had occurred, or could she give the exact numerical value of the quantity? If so, it is well-specified."*

An example from IMAGE 1 of a parameter that is not well-specified is the "thickness of the warmer ocean mixed layer". An example from IMAGE 2 is the "effective depth of the surface" in the surface heat balance equation in the zonal atmosphere climate model. These constructs are, in fact, technical fixes to simplify the model and allow modeling at higher aggregation levels. The validity of the use of such constructs depends on the validity of the assumptions made in each step of the simplification process. This chain of successive assumptions is usually poorly documented and relies highly upon tacit knowledge and wisdom of the modelers. From our attempts to understand the IMAGE 2 model from its documentation and its peer reviewed publications, we feel that there is an urgent need for a comprehensive documentation of the chains of assumptions followed in the construction process of IAMs in order to allow for quality assessment and quality control. Without such documentation we can only resort to debriefing through depth-interviews with the modelers to get hold of the full set of assumptions on which the model is based.

A good example from the current practice of quality control of IMAGE is the multidisciplinary and international IMAGE Advisory Board, which was set up in 1992 with the explicit task of addressing the scientific quality of the model and its usability for policy development support (Hordijk, 1993; Solomon, 1994). The IMAGE advisory board consists of nine scientists and one policy-maker, and its composition has changed over time. The IMAGE modeling team also attended the meetings. According to Alcamo (1994b), the advantage of a mixed advisory board is that the policy makers hear from the scientists what the limitations are of the science, and the scientists can hear from the decision makers what the needs of the model should be. Alcamo calls the advisory board also a *"device to obtain the support of not only policy makers, but also scientists. Because a model like IMAGE should be steered through the best science available, and the most policy relevant fashion as possible."* This implies that the board also is important as a means of obtaining legitimation and negotiating credibility for the model.

The advisory board played a significant role in the construction process of the model. At the first meeting (December 1992) the first design of the IMAGE 2 model was revealed to the board to test whether it complied with both scientific and policy requirements. Using their comments, the first fully operational version, 2.0 was completed by the middle of 1993. Changes in the model induced by policy makers (both directly from DGM and via the board) were, *inter alia*, more focus on the impacts on the food system, and on the adaptation capacity of ecosystems (both due to Article 2 of the FCCC), and a change in the regional break-down in that the US and Canada were treated as separate regions, to allow comparison and exchange with national research programs in that region (Alcamo, 1994b).

We further asked Alcamo to describe the selection procedure used in the IMAGE project to make sure that the best scientific understanding is included, and how they find out whether all relevant processes are included in the model. He answered *"For the selection it is based on having a good interdisciplinary team. We propose something in our best judgement, and then test it against a wide variety of scientists. The procedure I use is as follows. (1) Design a preliminary version of the sub-model as simple as possible, yet with an "adequate" representation of key processes. (2) Review this preliminary design with an expert in the subject covered by the model, and add detail (begrudgingly) until the expert is more or less satisfied that key processes are represented. For example, we began modeling global land cover in IMAGE 2 by assuming that future agricultural land will be located according to very simple "land use rules". Later, experts in land cover studies advised us about more detailed rules and factors that we should include in our calculations, for example including the role of rivers in steering inland agricultural development. These and other factors we included in later versions of our land cover sub-model."*

Alcamo pointed out that another important mechanism of quality control of the IMAGE model is exposing the model to peer review by submitting model results for publication in scientific journals. At an early stage of IMAGE 2, an overview of the entire model was published as a special issue of the journal Water Air and Soil Pollution, and later published as a book. As in the case of the advisory board, quality control was not the only motivation for publication. Another important aspect was credibility and acceptance by policy-makers and scientists. As a further means to strengthen the embedding of IMAGE in the scientific community, the IMAGE team also contributes to the International Geosphere Biosphere Program (IGPB).

Not all actors who have a stake in climate change and climate policy (of which Table 6 shows a selection) are currently involved in the quality control procedures. A recent innovation in this field of quality control of IAMs is the use of focus groups; these are currently being used in a European research project called ULYSSES (Urban LifestYles, SuStainability and Environmental aSsessment). Experiments are conducted on the interaction of monitored discourse groups with IAMs. *"The discourse group which we will call the IEA Jury [IEA=Integrated Environmental Assessment], will be a* *microcosm of social learning, and will include the concerns of the intended users of IEA, emphasizing the role of ordinary citizens"* (Kasemir *et al.*, 1996).

6.5. Areas for improvement in uncertainty management

In Table 7 we summarize the various tools available or currently being developed for managing uncertainties in integrated models. The results can be summarized in terms of the vertical dimension of Table 4: Inexactness can be addressed by stochastic modeling and subjective probability distributions and by 'cultural-theory sampling' from the ensemble of tenable expert interpretations. Unreliability requires quality assessment (see also section 8) and quality control. Ignorance is both unassessable and irreducible, so the only thing we can do is explore the border with ignorance. The paradox is that we try to reduce ignorance by doing more research whereas more research increases the border with ignorance and ignorance increases with increased commitments based on given knowledge (e.g. Wynne, 1992). Pascal once said: *"Science is like a ball in a universe of ignorance. The more we expand knowledge, the greater the ignorance encountered by the ball's expanding surface."* (Cited in: Giarini and Stahel, 1993). Giarini and Stahel (1993, p219/220) have put forward the philosophical notion that *"Our ignorance and our imperfect information are an instance of disequilibrium, a condition of life and of evolution. Our growing ignorance, determined by the growth of our knowledge which increases the number of unanswered questions, is the best evidence that we are part of the flow of life. Experience tells us that whenever we have the feeling of having completely mastered and understood a problem, it is often because the object or the situation of reference no longer exists: we are just about to discover that our confidence in our capacity "totally" to understand is at least partly misplaced."*.

Table 7 Tools available to address different sorts of uncertainties in IAMs

source / type	input data	model structure: parameters	model structure: relations	model completeness
inexactness	-sensitivity analysis -Monte Carlo Simulation -Subjective Probability Distributions	-sensitivity analysis -Monte Carlo Simulation -Subjective Probability Distributions	-inter-model comparison (EMF) -rack-method (Somermeijer) -cultural theory (Van Asselt) -advisory boards (IMAGE) -peer review	-advisory boards -inter-model comparison -peer review -competition among IAM groups -focus groups (ULYSSES)
unreliability	-expert judgement -quality assessment (NUSAP)	-expert judgement -quality assessment (NUSAP)	-quality assessment (NUSAP) -testing against historic data -inter-model comparison (EMF) -competition among IAM groups -rack-method (Somermeijer) -advisory boards (IMAGE) -peer review	-advisory boards -inter-model comparison -peer review -competition among IAM groups -focus groups (ULYSSES)
ignorance	-research	-research	-research	-research

From our analysis it follows that techniques currently available for uncertainty analysis and uncertainty treatment in IAMs have three major shortcomings:

1. They do not fully address all relevant aspects within the whole spectrum of types and sources of uncertainty;

2. They fail to provide unambiguous comprehensive insight to both the modeler and the user into:

(a) the quality and the limitations of the IAM;

(b) the quality and the limitations of the IAM-answers to the policy questions addressed;

(c) the overall uncertainties;

3. They fail to systematically address the subjective component in the appraisal of uncertainties (with the exception of the cultural-theory-sampling method in the TARGETS model).

In terms of Tables 5 and 7, the main areas for improvement are the second row, namely the assessment of error due to the unreliability of the knowledge about input data and model structure. and the third column, namely the assessment of error due to uncertainty about model structure. Regarding the latter, we mentioned the promising developments within the TARGETS model, namely addressing this problem with the help of cultural-theory-based consistent model routes. The important innovative elements of the TARGETS/cultural-theory sampling method are that it is for the first time in the IAM history that expert disagreement is incorporated in the model and that variable functional relations are introduced between the variables in the model. Although this method is of great value for managing uncertainty that stems from disagreement amongst experts, we think that the cultural theory approach is only a partial solution to the problems we identified. Cultural theory cannot be the panacea because it does not address the per se quality of the assumptions concerning model structure. It even runs the risk of marginalizing the quality issue by taking as an axiom that all thinkable perspectives are equally legitimate. When you apply that in its extreme form to model structure, then any conceivable model structure is legitimate and any plausibility-ranking of possible model structures based on scientific soundness would be just one of many legitimate views rather than conclusive intersubjective guidance for preferring one conceivable model structure to another. This would laed to total relativism, making science useless in the policy process. We take the position that science can provide conclusive arguments and that only scientifically tenable model structures should be used in science for policy, but we acknowledge the co-existence of multiple scientifically tenable interpretations of reality. In our view, systematic evaluation of the quality of each interpretation is desirable to further limit the interpretative space and (at least to attempt) to provide intersubjective rationales for discriminating between - or at least for ranking - different possible model structures.

Van der Sluijs (1995) proposed taking Funtowicz and Ravetz's (1990) NUSAP methodology as a starting point for better uncertainty management in IAMs. In the next section we will discuss the NUSAP methodology and explore how it can be used to improve uncertainty management in IAMs.

7. Disentangling the uncertainty problem

If uncertainties in IAMs are to be managed better, a first step is to distinguish between *quality*, which can be viewed as the inverse of "potential for improvement", and *limitations*, which refers to our limited capacity to know and understand and the inherent uncertainty in the system that remains if the "potential for improvement" has declined to zero. It should be kept in mind that what is an inherent limitation and what is, in principle, reducible uncertainty will change over time because of ongoing research and innovations that enlarge the toolbox (invention of better computers, invention of new mathematics for complex systems, invention of new modeling techniques, paradigm changes). This implies that "limitations" and "potential for improvement" can never be absolute and should always be treated as tentative wisdom rather than stable truth.

We illustrate the usefulness of the distinction between uncertainty due to limitations and uncertainty due to lack of quality by presenting the assessment diagram for the evaluation of model output as designed by Funtowicz and Ravetz (1990) (Figure 2). In this diagram one can map model constituents (parameters, model input, model assumptions etc.), to find out how they contribute to the overall quality of the calculated model output.

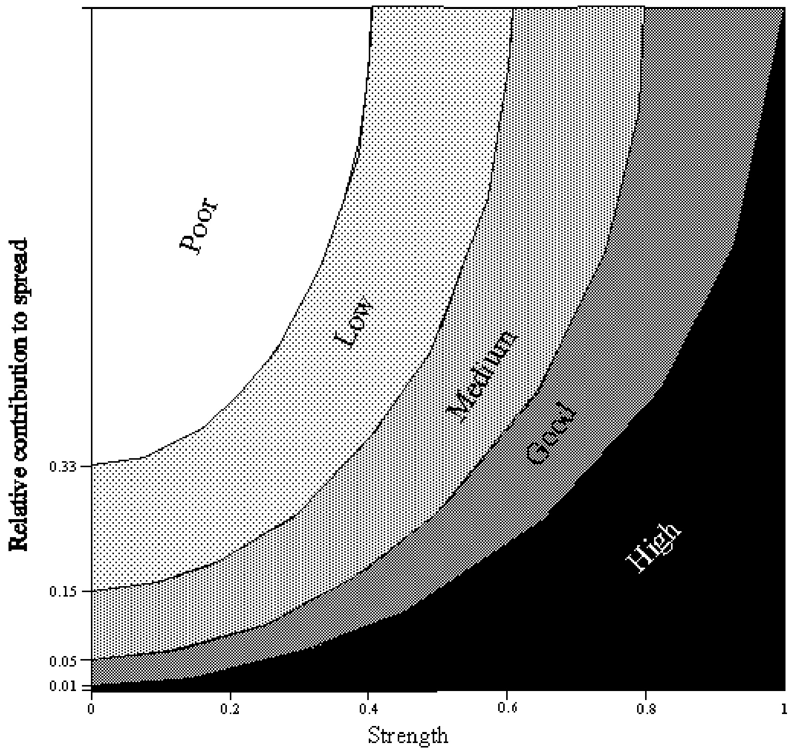

Figure 2 Assessment diagram, showing zones for evaluation of model output (Funtowicz and Ravetz, 1990)

The vertical axis of this diagram shows the relative contribution of uncertainty regarding a model constituent (e.g. a parameter value) to the total spread in model output. This relative contribution can be estimated using sensitivity analysis and Monte-Carlo-based uncertainty analysis. It should be noticed that this indicator is imperfect, because it is calculated with fixed values for all other parameters. It might well be that the sensitivity of model output to uncertainty in one parameter will change if other values are taken (within the bounds of their uncertainty ranges) for the fixed other parameters. This implies that from a perfectionist point of view, an analysis of the sensitivity of the outcome to the uncertainty in one variable, provided fixed values for the other parameters, should be followed by a sensitivity analysis of the sensitivity to the uncertainties in the fixed values of the other parameters. Deterministic and stochastic dependencies between parameters, in particular co-variances and conditional distributions (histograms), should also be assessed. The more parameters a model has, the more computing time it will take to do this meta-sensitivity analysis systematically. If after such a meta sensitivity analysis it should turn out that the 'sensitivity of the sensitivity' is high, it will be more and more difficult to disentangle the uncertainty problem in such a way that you really can attribute a certain percentage of the error in model output to the error in each single parameter. It might instead be more adequate to attribute it to for instance a joint distribution of a cluster of parameters. This is an important methodological problem that is not discussed in Funtowicz and Ravetz' (1990) book but certainly deserves attention. A soultion could be to search for another measure of performance for the vertical axis, one that is more unequivocally attributable to the uncertainties in individual model components. However, it is doubtful whether such a measure will exist in all cases, especially in complex coupled non-linear models and in cases of joint distributions of parameters. As a first approximation and despite its imperfection, the relative-contribution-to-spread provided-all-others-fixed, provides at least an objective measure for disentangling the uncertainty problem.

The horizontal axis maps the strength or quality of the model constituents. For well-established physical constants such as the Stefan-Boltzmann constant, the strength is very close to one. Parameters whose value is just an educated guess, such as the carbon dioxide fertilization parameter in the first generations of carbon cycle models, the strength will be low, e.g. 0.1. Mapping all model constituents in the assessment diagram reveals the weakest links in the model. It also helps greatly in the setting of priorities for model improvement. But the most important advantage of this diagram is that it makes it possible to distinguish between what part of the uncertainties in model outcome is solvable (by increasing the strength of the model constituents through further research) and what part is intrinsic to the modeled system. This distinction is important when designing strategies to cope with the uncertainties. For the intrinsic uncertainties, the challenge is to design response strategies that are robust against these uncertainties (for the same reason - and this is an example of a robust strategy - Henderson-Sellers (1996b) makes a case for adaptive strategies, diminishing the vulnerability of agricultural and ecosystems to future changes in soil moisture, because of fundamental limits to the predictability of future regional changes in soil moisture). On the other hand, the uncertainties due to lack of quality of the model constituents will not be dispelled without adequate research programs. An estimate of the time frame needed to jack up the quality to the required level is also needed. Note that the required level of quality for each model constituent is a function of the relative contribution to spread in model output caused by the uncertainty in that constituent, as can be seen from Figure 2. If the estimated time-frame to attain the required quality for a model constituent is too long, the challenge is to design response strategies that are robust against these uncertainties.

The NUSAP methodology designed by Funtowicz and Ravetz (1990) provides a good starting point to identify the strength or quality of model constituents. NUSAP is a notational scheme for scientific information. It 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:

- *Spread* conveys an indication of the inexactness;
- *Assessment* expresses a judgement on the reliability and indicates the *strength* of the data;
- *Pedigree* conveys an evaluative account of the production process of the information, and indicates the *scientific status* of the knowledge.

Pedigree is expressed in a set of evaluation criteria, a so-called pedigree matrix. Evaluation criteria used in a pedigree matrix are in fact yard-sticks for quality. These yard-sticks can be cognitive (e.g. theoretical structure) or social (e.g. peer acceptance). What criteria should be included in a pedigree matrix depends on the portion of information whose pedigree has to be determined. Examples of pedigree matrixes are given in Table 8. Note that the columns are independent of each other.

Table 8a. 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	Theory-based model	Historic/Field data	High	All but rebels
2	Computational model	Calculated data	Medium	Competing schools
1	Statistical processing	Educated guess	Low	Embryonic field
0	Definitions	Uneducated guess	None	No opinion

Table 8b. The pedigree matrix for environmental models as designed by Funtowicz and Ravetz (1990).

Code	Model structure	Data input	Testing
4	Comprehensive	Review	Corroboration
3	Finite-element approximation	Historic/field	Comparison
2	Transfer function	Experimental	Uncertainty analysis

1	Statistical processing	Calculated	Sensitivity analysis
0	Definitions	Expert guess	None

Table 8c. The pedigree matrix for radiological data entries (Funtowicz and Ravetz, 1990)

Code	Type	Source	Set-Up
4	Constants	Reviewed	Universal
3	Deduced	Refereed	Natural
2	Estimated	Internal	Simulated
1	Synthesized	Conference	Laboratory
0	Hypothetical	Isolated	Other

The choice of the quality criteria in a pedigree matrix, the choice of the quality modes in each column and the ranking of quality modes within one column, all are open to discussion. This makes it desirable to search for intersubjective procedures to establish pedigree matrixes. It is quite a complicated and not self-evident matter to select adequate quality modes for each column
and to rank them from low to high quality. We will illustrate that with the following example.

For "Theoretical structure" Funtowicz and Ravetz propose the quality modes (we put them in reverse order:)

0. Definition
1. Statistical processing
2. Computational model
3. Theory-based model
4. Established Theory

We tend to interpret this ranking also as a linear path through which research usually progresses. From that point of view we have difficulties with the mode 'definition' and we do not know the precise difference between 'statistical processing' and 'computational model'. A typical evolution over time of a 'theoretical structure' of a model constituent, resulting from research, might look something like this:
[*mode 0*] The first stage is, for instance, a (theory driven or empirically driven) notion that there is a correlation between quantities (e.g. global CO_2 concentration and global plant growth), but without sufficient knowledge about the precise nature of the correlation in both a qualitative (that is: theoretical) and a quantitative sense (either empirical or theoretical & empirical).

[*mode 1*] The notion of mode 0 is extended with empirical data. Statistical analysis of the data supports the notion.

[*mode 2*] A quantitative input-output 'black box' model is constructed that fits with the empirical data. Theory that explains from high process-detail the nature of the mathematical relations in the input-output model is lacking or is rather incomplete.

[*mode 3*] The mechanisms that constitute the correlation between the quantities are understood well enough to construct a theory-based model with high process-detail: the black box from mode 2 is now transparent. However, more empirical data are needed to quantify certain parameters in the model with high process-detail. Thorough verification and testing are still required.

[*mode 4*] A theoretically and empirically sound high process detail model, free of technical fixes and sufficiently validated, has been achieved. Often this model is simplified to a straightforward meta-model (which the users-community can use as a black-box model).

The resulting set of modes of Theoretical structure is:

0. Notion
1. Statistically indicated
2. Black-box model
3. Theory based model
4. Established theory

[5. simple meta-model - backed by established theory]

This ranking of these modes gives rise to a few problems that we would like to discuss. First, the evolution of a portion of scientific information over time might not proceed linearly from mode 0 to mode 4. A theoretically based model can precede statistical indication and black-box models; e.g. the paths 0-3-4 or 0-3-1-4 are also conceivable, or - with some imagination - even 3-0-4 is possible. If we assume that quality of the knowledge about a model constituent improves with increasing research on that constituent, then we have a ranking problem for the quality modes of the quality-yard stick in this example.

Second, The ranking of the modes depends strongly on the use (and the user) of the scientific information concerned (See Van der Sluijs, 1995). Someone who builds integrated assessment models would prefer a working black-box model to a scientifically sound theoretically based model with insufficient empirical data to quantify the key parameters. However, a scientist from a deterministic school would consider any "black box" relation inferior to even poorly developed scientific theory on that relation; even if the black-box performs better in the light of available empirical data.[1] The nature of science is to open black-boxes.

The above discussion shows that the measurement of the strength or quality of model constituents is not trivial and can be a function of the perspective taken by the actor that uses the model, even with regard to single dimensions of the quality hyper space constituted by all thinkable quality criteria.

A second problem in the operational use of the assessment diagram is the inherent impossibility of objectively aggregating the scores regarding a set of quality evaluation criteria to a single number between zero and one which represents strength as represented by the horizontal axis of the assessment diagram. We will show that these problems do not detract in any way from the value of the assessment diagram as a heuristic tool in model evaluation and priority setting for research. In their book Funtowicz and Ravetz (1990, chapter 12) elegantly show how the diagram can be applied to a radiological model of milk contamination with Cesium 137. This model however is extremely simple when compared to for instance the IMAGE 2 model. The data-file with input-data of IMAGE 2 is about 20 megabytes (personal communication Rik Leemans, October 1996). It is an impossible task to map all the constituents of a complex IAM such as IMAGE 2 in the same way in the assessment diagram. And it could be a thousand-year research project to map all constituents of a coupled ocean-atmosphere general circulation model in the assessment diagram. However, when the application of the assessment diagram is combined with expert judgement in selecting what should be mapped, it can be a valuable tool, helping in quality control, model improvement, communication of model uncertainties, and, most important, helping to distinguish between limitations and lack of quality. By providing a sound basis for this distinction, the method can generate key input for the design of response strategies, as we argued earlier in this section.

As an example, we will sketch a possible procedure by which the assessment diagram from Figure 2 can be used to evaluate the quality of an outcome of an IAM. A good example of such an outcome are the 'safe emission corridors' (see section 2) that are calculated with the IMAGE 2 model (Alcamo and Kreileman, 1996; Swart *et al.*, 1996). In our example-procedure, objective information obtained from sensitivity and uncertainty analysis will be combined with systematically obtained expert judgement to determine the X-axis and Y-axis position in the assessment diagram of the model constituents that are to be mapped. The example is inspired by the methodology developed by Morgan and Keith (1995) for ranking key uncertainties in climate modeling.

---

[1] Compare this to a medicine that is effective in curing a certain disease, regardless scientists do not know why it works. The physician will be satisfied with it and use it in her or his practice, whereas the medical researcher will not be satisfied.

Step 0 Select an expert panel. For the IMAGE 2.0 model this panel can e.g. exist of the modelers and the IMAGE Advisory Board. But, it might be a good idea to include modelers from rival IAM-groups as well.

Step 1. Let the expert panel select model constituents to be mapped in the assessment diagram. Instead of mapping every single parameter and model assumption of an IAM, one can start at a higher aggregation level and map, for instance, entire sub-models of the IAM.

Step 2. Carry out sensitivity analysis and Monte-Carlo-based uncertainty analysis to find out how much each constituent contributes to the total spread in model outcome. It is likely that adequate information and resources are insufficient for a complete Monte-Carlo-based uncertainty analysis for each constituent. For this reason we added step 9.

Step 3. Choose (or design if necessary) adequate pedigree matrixes to apply to each model constituent that is to be mapped.

Step 4. Let the panel determine the pedigree scores for each constituent.

Step 5. Make one card for each model constituent with its pedigree scores from step 4 and all available information from sensitivity and uncertainty analysis from step 2.

Step 6. Let each panel member individually sort the cards according to strength. After sorting, let each member position the cards along a linear scale from 0 to 1.

Step 7. Determine the average scores and their standard deviations for all constituents and discuss the results at a plenary meeting of the panel. Pay special attention to scores with a high standard deviation.

Step 8 Allow re-ranking by individual members if they changed their mind after the plenary discussion. Now the X-axis scores are available for each constituent to be mapped in the assessment diagram.

Step 9. If the panel thinks the information of step 2 too incomplete to determine the Y-axis scores of each constituent straightforwardly, repeat steps 6 to 8 but let the members sort the cards according to its relative contribution to spread in the model outcome.

Step 10. Map the resulting average scores and their standard deviations in both dimensions, the latter as error bars, in the assessment diagram. The size of the error bars represents the disagreement among the members on the position of each constituent in the diagram.

Steps 4 and 9 can be further improved by adding debriefing sessions with key experts in the field for each constituent. The debriefing sessions can reveal tacit knowledge regarding (the quality of) that constituent, making the ranking-exercise better informed. To measure the social dimensions of quality such as 'peer acceptance', scientometric methods could be of help in determining the pedigree scores.

The procedure can be used iteratively. After step 10, one has insight into the weakest constituents of the model - in view of the policy question addressed - and one can further disentangle the uncertainty problem by going back to step 1 and selecting model constituents at a lower aggregation level.

The NUSAP methodology can also be of help in the drafting of subjective probability distributions for Delphic Monte Carlo Analysis. In the current studies (e.g. Titus and Narayanan, 1996), the experts consulted are only asked to provide the first three qualifiers: a numeral, a unit and a distribution function. Adding the qualifiers Pedigree and Strength in Delphic Monte Carlo Analysis would make it possible to apply the assessment diagram to the outcomes of Delphic Monte Carlo Analysis as well. The relative contribution to spread would then be determined by analyzing the sensitivity to the spread in the distribution functions, for instance by comparing model-outcome-distribution assuming half the spread and double the spread for each parameter distribution function. Adding such a quality analysis would remove some of the criticism concerning the use of subjective distribution functions.

8. Conclusions

In this paper we have explored the problems of uncertainty management in Integrated Assessment Models (IAMs) of Climate Change in relation to their mission to model the entire causal chain and to guide and inform the climate policy debate and the negotiations on the climate convention. We have identified areas for improvement in uncertainty management in IAMs and we propose a methodology, based on the work by Funtowicz and Ravetz, for disentangling the uncertainty problem in IAMs. This methodology will enable us to assess the quality of the model results and to identify the weakest links in the models.

We conclude that:

i. Man's knowledge and understanding of the modeled causal chain of climate change (see Table 2) is incomplete and characterized by large uncertainties and limits to predictability. In each stage of the causal chain there are both potentially reducible and probably irreducible uncertainties affecting the estimates of future states of key variables and the future behavior of system constituents. The potentially reducible parts stem from incomplete information, incomplete understanding, lack of quality in data and model assumptions and disagreement between experts. The probably irreducible parts stem from ignorance, epistemological limits of science, undeterministic system elements, practical unpredictability of chaotic system components, limits to our ability to know and understand, limits to our ability to handle complexity, unmodelability of surprise, non-smooth phenomena, intransitive system components and multiple equilibria.

ii. None of the currently available IAMs really integrates the entire causal chain, nor do they take dynamically into account all feedbacks and linkages between the different stages of the causal chain.

iii. The state of science backing the (mono-disciplinary) sub-models of IAMs differs across sub-models. In other words, the current climate IAMs consist of a mixture of constituents covering the whole spectrum from educated guesses to well-established knowledge. Further, we know that the models are incomplete, but it is uncertain to what extent.

iv There is a controversy about the usefulness of IAMs for the assessment of climate change. The positions in the debate vary from "We are not ready to do integrated modeling, we must wait until all science used in the model has the status of well established knowledge", to "We have the responsibility to use our best scientific understanding to develop reasonable policies. Integrated modeling is the best way of combining our knowledge in such a way that we can evaluate the consequences of different policy scenarios, do cost-benefit framing or optimize cost effectiveness to reach a target."

There is however agreement that IAMs are not truth-machines and cannot reliably predict the future, but rather they are heuristic tools. IAMs are capable of testing sensitivity, of answering 'what if' questions (although each answer has to be followed by ", given the total set of assumptions of this model"), of ranking policy options, of assessing the relative importance of uncertainties, of identifying research priorities and of providing insights that cannot easily be derived from the individual natural or social science component models that have been developed in the past.

v. Initially, IAMs were promoted as '**the** tools' for integrated environmental assessment, but now they are becoming 'one of several tools' in a broader integrated environmental assessment process. In this new position, IAMs are a part of the assessment, not the whole.

vi Stochastic modeling by Monte Carlo Simulation is needed to fully assess error propagation through the model and to identify the distribution function of the outcome of an IAM that results from the distribution functions of the input data, the model parameters and the model relations. Due to its resource-consuming character and due to lack of information on the distribution functions of all individual model constituents, Monte Carlo Simulation is not common practice in IAMs. For the IMAGE-2 model, Monte-Carlo simulation has been applied to a few individual sub-models only and these instances rely on (imperfect) subjective probability functions. This implies that error-propagation throughout the IAM has not yet been addressed.

vii When IAMs are made stochastic, it should be possible to use them for goal-searching in order to find solutions that are robust against the specified uncertainty in the model.

viii We identified a mismatch between the types and sources of uncertainty that should be addressed on the one hand (Table 5) and the current practice of uncertainty management in IAMs and the tools available, on the other hand. From our analysis it follows that techniques currently available for uncertainty analysis and uncertainty treatment in IAMs have three major shortcomings:

1. They do not fully address all relevant aspects within the whole spectrum of types and sources of uncertainty;

2. They fail to provide both the modeler and the user with unambiguous comprehensive insight into:

(a) the quality and the limitations of the IAM;

(b) the quality and the limitations of the IAM-answers to the policy questions addressed;

(c) the overall uncertainties;

3. They fail to systematically address the subjective component in the appraisal of uncertainties (with the exception of the cultural-theory-sampling method in the TARGETS model).

The main areas for improvement are the assessment of error in model output due to unreliability (that is the lack of quality; not to be confused with the assumed spread or the assumed distribution function) of the knowledge about input data, parameters, and model structure, and the quantitative assessment of error in model output due to uncertainty about model structure.

ix Building further on Funtowicz and Ravetz' (1990) NUSAP methodology and their Assessment Diagram, we propose a 10-step iterative methodology to disentangle the uncertainty problem in IAMs. The methodology discriminates between the potentially solvable and the currently unsolvable uncertainties. This information is crucial for the development of adequate response strategies. The response strategy has to be robust against the currently insoluble uncertainties, whereas adequate research programs should be designed to reduce the potentially solvable uncertainties. The methodology further has the potential to assess the quality of model output and to identify the parts of the model whose individual lack of quality contributes the most to the overall lack of quality. The latter information is very useful for setting research priorities.

x The methodology can also improve the drafting procedures for subjective probability for Delphic Monte Carlo analysis, because it adds information on the quality of each subjective distribution function. It can then be used to identify which subjective distribution functions should first be improved to improve the quality of the model output. The proposed method has not been tested so far, and standardization of the sets of quality yard-sticks to apply (the so called pedigree matrixes), and of procedures to draft them for individual model constituents, is desirable.

9. Acknowledgments

We thank Jill Jäger, Joanne Linnerooth-Bayer, Yuri Ermoliev, Jerry Ravetz, Rob Swart, Joe Alcamo, Josee van Eijndhoven, Anne Stijkel and Stephen Healy for making useful comments on earlier versions of this paper. We are grateful to Lauren Hale and Sheila McNab for making linguistic and stylistic improvements.

10. References

H. Abarbanel, S. Koonin, H. Levine, G. MacDonald, and O. Rothaus, *Issues in Predictability*, JSR-90-320, JASON, The MITRE Corporation, McLean, VA, 1991.

J. Adams, *Risk*, UCL Press, London, 1995a.

J. Adams, *Cost Benefit Analysis: Part of the Problem, Not the Solution*, Green College Centre for Environmental Policy & Understanding, University College London, 1995b.

J. Alcamo, L. Hordijk, J. Kämäri, P. Kauppi, M. Pösch, and E. Runca, Integrated Analysis of Acidification in Europe, in: *Journal of Environmental Management*, **21**, 1986.

J. Alcamo, R. Shaw, and L. Hordijk (eds.), *The RAINS Model of Acidification: Science and Strategies in Europe*, Kluwer Academic Press, Dordrecht, Boston, 1990, 402 pp.

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.

J. Alcamo, M. Krol, and R. Leemans, *Stabilization of Atmospheric* CO_2 *and Global Changes in Land Cover and Crop Productivity*, Paper presented to First Special IPCC workshop on Article 2 of the Climate Convention, Forteleza, Brazil, 17-21 October, 1994.

J. Alcamo, M. Krol, and R. Leemans, *Stabilizing Greenhouse Gases: Global and Regional Consequences*, *Results from the IMAGE 2.0 Model*, Prepared as background report for the Dutch delegation to the first session of the Conference of Parties to the UN Framework Convention on Climate Change Berlin, Germany 28 March - 4 April, 1995, RIVM, Bilthoven, 1995.

J. Alcamo and E. Kreileman, *The Global Climate System: Near Term Action for Long Term Protection*, RIVM, Bilthoven, 1996.

W.B. Arthur, Competing Technologies: an overview in: G. Dosi, Ch. Freeman, R. Nelson, G. Silverberg, and L. Soete, *Technical Change and Economic Theory*, Pinter Publishers, London, 1988 p.590-607.

J. Austin, N. Butchart and K.P. Shine, Possibility of an Arctic Ozone Hole in a Doubled-CO_2 climate, *Nature*, **360**, 1991 p. 221.

P. Bailey, C. Gough, M. Chadwick, and G. McGranahan, *Methods for Integrated Environmental Assessment: Research Directions for the European Union*, Stockholm Environmental Institute, 1996.

J.J. Batjes and C.G.M. Goldewijk, *The IMAGE 2 Hundred Year (1890-1990) Database of the Global Environment (HYDE)*, RIVM report no, 481507008, Bilthoven, 1994.

M.B. Beck, Understanding Uncertain Environmental Systems, in: *Predictability and Nonlinear Modeling in Natural Sciences and Economics*, J. Grasman and G. van Straten (Eds.); Kluwer Academic Publishers: Dordrecht / Boston / London, 1994, p. 294-311.

M.B. Beck, J.R. Ravetz, L.A. Mulkey, and T.O. Barnwell, *On The Problem of Model Validation for Predictive Exposure Assessments*, 1996a.

M.B. Beck, L.A. Mulkey, T.O. Barnwell, and J.R. Ravetz, *Model Validation for Predictive Exposure Assessments*, EPA, 1996b.

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

S. Boehmer-Christiansen, Global Climate Protection Policy: the Limits of Scientific Advice, Part 1, *Global Environmental Change*. **4** (2), 1994a, p. 140-159.

S. Boehmer-Christiansen, Global Climate Protection Policy: the Limits of Scientific Advice, Part 2, *Global Environmental Change*. **4** (3), 1994b, p. 185-200.

S.A. Boehmer-Christiansen, A Scientific Agenda for Climate Policy?, *Nature*, **372**, 1994c, p. 400-402.

S.A. Boehmer-Christiansen, Comments on S. Shackley and B. Wynne Paper "Global Climate Change: the Mutual Construction of an Emergent Science-Policy Domain", in: *Science and Public Policy*, **22** (6), 1995, p.411-416.

H. Brooks, The Typology of Surprises in Technology, Institutions and Development, in: W.C. Clark (ed.), *Sustainable Development of the Biosphere*, Cambridge University Press, New York, 1986, p.325-348.

J. Casti, What if..., in: *New Scientist*, 13 July 1996.

J. Casti, *Would-Be Worlds*, John Wiley & Sons, 1996.

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.

E. Cook, *Lifetime Commitments: Why Climate Policy-makers can't Afford to Overlook Fully Fluorinated Compounds*, World Resources Institute, 1995, 16 pp.

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

K. Denman, E. Hofmann and H, Marchant, Marine Biotic Responses to Environmental Change and Feedbacks to Climate, in: J.T. Houghton, L.G. Meira Filho, B.A. Callander, N. Harris, A. Kattenberg, and K. Maskell (eds.), *Climate Change 1995, The Science of Climate Change*, Contribution of WGI to the Second Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 1996.

B. de Vries and R. van den Wijngaart, *The Targets/IMage Energy (TIME) 1.0 Model*, RIVM report no. 461502016, Bilthoven, 1995.

M. Douglas and A. Wildavsky, *Risk and Culture: An Essay on the Selection of Technological and Environmental Dangers*, University of California Press, Los Angeles, 1982, 221 pp.

H. Dowlatabadi, Integrated Assessment Models of Climate Change, An Incomplete Overview, in: *Energy Policy*, **23** (4/5), pp 289-296, 1995

C. Ebert and A. Karmali, Uncertainties in Estimating Greenhouse Gas Emissions, Environment Working Paper No. 52, World Bank, April 1992, 44 pp.

J. Emsley (ed.), *The Global Warming Debate, The Report of the European Science and Environment Forum*, Bourne Press, Dorset, 1996, 288 pp.

Energy Modeling Forum, *Final First Round Study Design for EMF 14: Integrated Assessment of Climate Change*, Energy Modeling Forum, Stanford University, Stanford, 1995a.

Energy Modeling Forum, *Model Results*, Stanford University, Stanford, 1995b.

Energy Modeling Forum, *Second Round Study Design for EMF 14: Integrated Assessment of Climate Change*, Energy Modeling Forum, Stanford University, Stanford, 1996.

Energy Modeling Forum, *Second Round Scenario Results*, Energy Modeling Forum, Stanford University, Stanford, 1996b.

Environmental Resources Limited, *Handling Uncertainty in Environmental Impact Assessment*, Environmental Resources Limited, London, 1985.

Yu. Ermoliev and J.B. Wets, *Numerical Techniques for Stochastical Optimisation*. Springer Verlag, Berlin, 1988, 571 pp.

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

S.O. Funtowicz and J.R. Ravetz, Science for the Post Normal Age, *Futures*, September 1993, p. 739-755.

S.O. Funtowicz and J.R. Ravetz, Uncertainty Complexity and Post Normal Science, *Environmental Toxicology and Chemistry*, **13** (2), 1994, p. 1881-1885.

W.L. Gates, A. Henderson-Sellers, G.J. de Boer, C.K. Folland, A. Kitoh, B.J. McAvaney, F. Semazzi, N. Smith, A.J. Weaver, and Q.-C. Zeng, Climate Models - Evaluation, in: J.T. Houghton, L.G. Meira Filho, B.A. Callander, N. Harris, A. Kattenberg, and K. Maskell (eds.), *Climate Change 1995, The Science of Climate Change*, Contribution of WGI to the Second Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 1996.

O. Giarini and W.R. Stahel, *The Limits to Certainty, Facing Risks in the New Service Economy*, 2nd revised edition, Kluwer Academic Publishers, Dordrecht, 1993.

J. Goudriaan, *Model, Schaal en Aggregatieniveau*, (Model, Scale and Aggregation Level) inaugural oration, 30 September 1993.

M.C.G. Hall, Estimating the Reliability of Climate Model Projects, Steps towards a Solution. In: *The Potential Climate Effects of Carbon Dioxide*, M.C. MacCracken and F.M. Luther (eds), US-DOE/ER-0237, Washington D.C., 1985, p. 337-364.

K. Hasselman, Stochastic Climate Models, Part 1: Theory, *Tellus*, **28**, 1976, p. 473-485.

G.P. Hekstra, The Case for a Global CO_2-Emission Ceiling. Paper presented at *UNEP/ECE Regional Seminar on Alternative Patterns of Development and Lifestyles*, Ljubljana, 3-8 December 1979, 10 pp.

T. Hellström, The Science-Policy dialogue in transformation: model-uncertainty and environmental policy, in: *Science and Public Policy*, **23** (2), 1996, p.91-97.

J.C. Helton, Treatment of Uncertainty in Performance Assessments for Complex Systems, *Risk Analysis*. **14** (4), 1994, p. 483-???

A. Henderson-Sellers, Bridging the Climate Gaps, in: T.W. Giambelluca and A. Henderson-Sellers, *Climate Change: Developing Southern Hemisphere Perspectives*, John Wiley & Sons, New York, 1996a.

A. Henderson-Sellers, Climate Modeling, Uncertainty and Responses to Predictions of Change, in: *Mitigation and Adaptation Strategies for Global Change*, **1** (1), 1996b.

A. Henderson-Sellers and R. Robinson, *Contemporary Climatology*, Longman, Harlow, 1986.

J. Hirshleifer and J.G. Riley, *The Analytics of Uncertainty and Information*, Cambridge University Press, Cambridge, 1992, 465 pp.

F.O. Hoffman and J.S. Hammonds, Propagation of Uncertainty in Risk Assessments: The Need to Distinguish Between Uncertainty Due to Lack of Knowledge and Uncertainty Due to Variability, *Risk Analysis*, **14** (5), 1994, p. 707-712.

C.S. Holling, The Resilience of Terrestrial Ecosystems: Local Surprise and Global Change, in: W.C. Clark and R.E. Munn (eds.) *Sustainable Development of the Biosphere*, Cambridge University Press, 1986, p. 292-317.

L. Hordijk, *An Integrated Assessment Model for Acidification in Europe*, Ph.D. thesis, Free University of Amsterdam, 1991a, 270 pp.

L. Hordijk, "Use of the RAINS Model in Acid Rain Negotiations in Europe." *Environmental Science and Technology* **25** (4), 1991b, p. 596-602.

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.

J.T. Houghton, L.G. Meira Filho, B.A. Callander, N. Harris, A. Kattenberg and K. Maskell (eds.), *Climate Change 1995, The Science of Climate Change*, Contribution of WGI to the Second Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 1996.

J.T. Houghton, G.J. Jenkins, and J.J. Ephraums (eds.), *Climate Change, The IPCC Scientific Assessment*, Cambridge University Press, 1990, 365 pp.

J.T. Houghton, B.A. Callander and S.K. Varney (eds.), *Climate Change 1992 The Supplementary Report to the IPCC Scientific Assessment*, Cambridge University Press, 1992, 200 pp.

J.T. Houghton, L.G. Meira Filho, J. Bruce, Hoesung Lee, B.A. Callander, E. Haites, N. Harris, and K. Maskell (eds.), *Climate Change 1994 Radiative Forcing of Climate Change and An Evaluation of the IPCC IS92 Emission Scenarios*, Cambridge University Press, 1994, 339 pp.

J.T. Houghton, L.G. Meira Filho, B.A. Callander, N. Harris, A. Kattenberg, and K. Maskell (eds.), *Climate Change 1995, The Science of Climate Change*, Contribution of WGI to the Second Assessment Report of the Inter Governmental Panel on Climate Change, Cambridge University Press, 1996.

IGBP, *Reducing Uncertainties*, International Geosphere Biosphere Programme, Stockholm, 1992.

S.L.S. Jacoby and J.S. Kowalik, *Mathematical Modeling with Computers*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1980

C.C. Jaeger and B. Kasemir, Climatic Risks and Rational Actors, in: *Global Environmental Change*, 1996 (forthcoming).

J. Jäger (ed.), *Responding to climate change: Tools for Policy Development*, Stockholm Environmental Institute, Stockholm, 1990.

J. Jäger, J. Cavender Bares, N. Dickson, A. Fenech, P. Haas, F. Hampson, P. Hughes, A. Liberatore, E.A. Parson, V. Sokolov, F. Toth, C. Waterton, J.P. van der Sluijs, J.C.M. van Eijndhoven, and B. Wynne 1997. Risk Assessment, in: The Social Learning Group, *Learning to Manage Global Environmental Risks: A Comparative History of Social Responses to Climate Change, Ozone Depletion and Acid Rain*. MIT Press, forthcoming (1997?).

P.H.M. Janssen, W. Slob, and J. Rotmans, *Sensitivity Analysis and Uncertainty Analysis: A Statement of Affairs of Ideas, Methods and Techniques*, National Institute of Public Health and the Environment (RIVM) report no. 958805001, Bilthoven, The Netherlands, 1990 (in Dutch).

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.A. Janssen and W.J.M. Martens, *Sex, Drugs and Climate Change: Modeling Malaria as a Complex Adaptive System*, Report no. GWP95-2, RIVM, Bilthoven, 1995.

M. Janssen and J. Rotmans, *The Evolution of Integrated Assessment Models for Global Climate Change: Towards a new Generation*, paper prepared for the Energy Modeling Forum Working Group Meeting at Stanford University, California, 24-26 May 1995.

R.W. Kates and W.C. Clark, Expecting the unexpected, in: *Environment*, **38** (2), 1996, p. 6-11 and 28-34.

B. Kasemir, C.C. Jaeger, and G. Dürrenberger, Embedding Integrated Assessment Models in Social Discourse, in: *Science and Public Policy*, **23** (2), 1996, p. 124-125.

A. Kattenberg, F. Giorgi, H. Grassl, G.A. Meehl, J.F.B. Mitchell, R.J. Stouffer, T. Tokioka, A.J. Weaver, and T.M.L. Wigley, Climate Models - Projections of Future Climate, in: J.T. Houghton, L.G. Meira Filho, B.A. Callander, N. Harris, A. Kattenberg, and K. Maskell (eds.), *Climate Change 1995, The Science of Climate Change*, Contribution of WGI to the Second Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 1996.

B. Keepin and B. Wynne, Technical Analysis of IIASA energy scenarios, in: *Nature*, **312**, p. 691-695.

N. Keilman, Ex-Post Errors in Official Population Forecasts in Industrialized Countries, *Journal of Official Statistics*, (accepted for publication; forthcoming).

D.W. Keith, When is it Appropriate to Combine Expert Judgements? An Editorial Essay, in: *Climatic Change*, **33** (2), 1996, p.139-143.

L.F. Konikow and J.D. Bredehoeft, Ground-water Models cannot be Validated, in: Advances in Water Resources, **15** (2), 1992, p.75-83.

F. Krause, *From warming fate to warming limit: benchmarks for a global climate convention*, El Cerrito, California, 1989.

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 Modeling 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.

J. Leggett, J. (eds.) *Global Warming; the Greenpeace Report*, Oxford University Press, Oxford, 1990.

C. Lorius and H. Oescher, Palaeo-perspectives: Reducing Uncertainties in Global Change?, *Ambio*, **23** (1), 1994, p. 30-36.

J. Lovelock, The Ages of Gaia, A Biography of Our Living Earth, Oxford University Press, Oxford, 1988, 252 pp.

L. Lunde, *Science or Politics in the Greenhouse? A Study Towards Scientific Consensus on Climate Change*, Report no. 1991/8, Fridtjof Nansen Institute, Lysaker, 1991.

W. Lutz, W. Sanderson, and S. Scherbov, *Probabilistic World Population Projections Based on Expert Opinion*, IIASA working paper WP-96-17, Laxenburg, February 1996, 27 pp.

A.S. Manne and G.R. Richels, Buying greenhouse insurance, in: *Energy Policy*, **19** (1), 1991, pp. 543-552.

L. Margulis and J.E. Lovelock, Biological Modulation of the Earth's Atmosphere, in: *Icarus*, **21**, 1974, p. 471-489.

W.J.M. Martens, *Modeling the Effect of Global Warming on the Prevalence of Schistomomiasis*, Report no. 461502010, RIVM, Bilthoven, 1995.

W.J.M. Martens, J. Rotmans, and L.W. Niessen, *Climate Change and Malaria Risk*, Report no. 461502003, RIVM, Bilthoven, 1994.

M.P. McCormick, L.W. Thomason, and C.R. Trepte, Atmospheric Effects of the Mt Pinatubo Eruption, in: *Nature*, **373**, 1995, p. 399-404.

M.D. McKay, R.J. Beckman, and W.J. Conover, A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code, in: *Technometrics*, **21** (2), 1979, 239-245.

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.

I.M. Mintzer, *A Matter of Degrees: The Potential for Controlling the Greenhouse Effect*, World Resources Institute, Washington D.C., 1987

M.G. Morgan and M. Henrion, *Uncertainty, A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, Cambridge University Press, 1990.

M.G. Morgan and D.W. Keith, Subjective Judgements by Climate Experts', in: *Environmental Science and Technology*, **29**, 1995, p.468A-476A.

NOP-MLK, *Programming Memorandum - Second Phase 1995-2000*, Internal report II-4, Programmabureau NOP, RIVM, Bilthoven, 1994

V. Norberg-Bohm, W.C. Clark, M. Koehler, and J. Marrs, Comparing Environmental Hazards: The Development and Evaluation of a Method based on a Causal Taxonomy of Environmental Hazards. In: W.C. Clark (Ed.), *Usable Knowlegdge for Managing Global Climatic Change*, Stockholm Environmental Institute, 1990.

W.D. Nordhaus, *The Economics of the Greenhouse Effect*, paper presented at the June 1989 meeting of the International Energy Workshop, Laxenburg, 1989

W.D. Nordhaus, To Slow or not to Slow: The Economics of the Greenhouse Effect, in: *The Economic Journal*, **101**, 1990, pp 920-937.

NSF (U.S. National Science Foundation) *NSF Awards $17.8 Million to Study*
Human Dimensions of Global Change, press release NSF PR 96-40, August 6, 1996.

J. Oerlemans and H.M. van den Dool, Energy Balance Climate Models: Stability Experiments with a Refined Albedo and Updated Coefficients for Infrared Emission, in: *Journal of the atmospheric sciences*, **35**, 1978, p. 371-381

E.A. Parson, *Searching for Integrated Assessment: A Preliminary Investigation of Methods and Projects in the Integrated Assessment of Global Climatic Change*, Paper presented at the 3rd meeting of the CIESIN-Harvard Commission on Global Environmental Change Information Policy, NASA Headquarters, Washington DC, February 17-18, 1994.

E.A. Parson, Integrated Assessment and Environmental Policy Making, In Pursuit of Usefulness, in: *Energy Policy*, **23** (4/5), 1995, p. 463-476.

E.A. Parson, *How Should we Study Global Environmental Problems? A Plea for Unconventional Methods of Assessment and Synthesis*, 1996.

Property Claim Services, *Hurricane Record 1949-1995, Property Claim Services Report Prepared for International Institute for Applied Systems Analysis*, Property Claim Services (A division of the American Insurance Services Group, Inc.), 1996.

S. Rahmstorf, Bifurcations of the Atlantic Thermohaline Circulation in Response to Changes in the Hydrological Cycle, *Nature*, **378**, 1995, p. 145-149.

J. Ravetz, Usable Knowledge, Usable Ignorance: Incomplete Science with Policy Implications, in: W.C. Clark (ed.), *Sustainable Development of the Biosphere*, Cambridge University Press, New York, 1986, p.415-432.

F.R. Rijsberman and R.J. Swart, *Targets and Indicators of Climatic Change*, Stockholm Environmental Institute, Stockholm, 1990.

RMNO, *Onderzoek in Nederland naar de gevolgen van toenemende CO_2 en andere sporegassen in de atmosfeer door menselijke activiteiten.* RMNO, Rijswijk, 1984, 48 pp.

J. Rotmans, *IMAGE An Integrated Model to Assess the Greenhouse Effect*, (Thesis), Rijksuniversiteit Limburg, 1990, 295 pp.

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

J. Rotmans and H. Dowlatabadi, Integrated Assessment of Climate Change: Evaluation of Methods and Strategies, in: S. Rayner and J. Emonds (eds.) State of the Art report. (referred to in Van Asselt and Rotmans, 1995)

J. Rotmans and H.J.M. de Vries (ed.), *Insights, no Answers* (forthcoming)

W.D. Rowe, Understanding Uncertainty, *Risk Analysis*, **14** (5), 1994, p. 743-750.

D. Schimmel, D. Alves, I. Enting, M. Heimann, F. Joos, D. Raynaud, T. Wigley, M. Prather, R. Derwent, D. Ehhalt, P. Fraser, E. Sanhueza, X. Zhou, P. Jonas, R. Charlson, H. Rodhe, S. Sadasivan, K.P. Shine, Y. Fouquart, V. Ramaswamy, S. Solomon, J. Srinivasan, D. Albritton, R. Derwent, I. Isaksen, M.Lal, and D. Wuebbles, Radiative Forcing of Climate Change, in: J.T. Houghton, L.G. Meira Filho, B.A. Callander, N. Harris, A. Kattenberg, and K. Maskell (eds.), *Climate Change 1995, The Science of Climate Change*, Contribution of WGI to the Second Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 1996.

D. Schimmelpfennig, 1996, Uncertainty in Economic Models of Climate-Change Impacts, in: *Climatic Change*, **33** (2), 1996, p. 213-234.

M. Schlesinger and J. Mitchell, 'Climate Model Simulations of the Equilibrium Climate Response to Increased Carbon Dioxide', *Review of Geophysics*, **25**, 1987, 760-98.

S. Schneider, The Greenhouse Effect: Science and Policy, *Science*, **243**, 1989, p. 771-781.

U. Schotterer and H. Oeschger, The Stability of the Climate System in the light of Recent Ice Core Measurements, in: N. Naki_enovi_, W.D. Nordhaus, R. Richels, and F.L. Toth (eds.) *Integrative Assessment of Mitigation, Impacts, and Adaptation to Climate Change*, CP-94-9, International Institute for Applied Systems Analysis, Laxenburg, 1994.

C.J.E. Schuurmans, H.M. van den Dool, and J.A. Wisse, Luchtverontreiniging en wereldklimaat (1) *Zenit*, **7** (1), 1980, p. 8-13.

M. Schwartz and M. Thompson, *Divided we Stand: Redefining Politics, Technology and Social Choice*, Harvester Wheatsheaf, London, 1990, 176 pp.

Science and Policy Associates, inc. and Holland Consulting Group, *Evaluation of the Technical Emphasis, Policy Relevance and Management Performance of the Dutch National Research Program on Global Air Pollution and Climate Change,* *Executive Summary*, Science and Policy Associates, inc., Washington DC and Holland Consulting Group, Amsterdam, October 1992.

Science and Policy Associates, inc., *Dutch National Research Programme On Global Air Pollution and Climate Change, NRP Phase I Final Review Report*, Science and Policy Associates, inc., Washington D.C., March 1995.

S. Shackley and B. Wynne, Global Climate Change: the Mutual Construction of an Emergent Science-Policy Domain, in *Science and Public Policy*, **22** (4), 1995, p.218-230.

S. Shackley, J. Risbey, P. Stone and B. Wynne, Adjusting to Policy Expectations in Climate Change Modeling: An Interdisciplinary Study to Flux Adjustments in Coupled Atmosphere-Ocean General Circulation Models, (June 1996 draft of forthcoming paper, CSE, University of Lancaster).

A.M. Solomon, *Report of International Review Meeting IMAGE 2.0*, 20-22 June 1994 Amsterdam, The Netherlands, Dutch National Research Program on Global Air Pollution and Climate Change (NRP) Report 00-09, 1994.

H. Sonnenschein, Market Excess Demand Functions, in: *Econometrica*, 40 p. 549-562.

N. Stehr, and H. von Storch, The Social Construct of Climate and Climate Change, *Climate Research*. **5**, 1995a, p. 99-105.

N. Stehr, H. von Storch, and M. Flügel, *The 19th Century Discussion of Climate Variability and Climate Change: Analogies for the Present Debate?*, Report no. ISSN 0937-1060, Max-Planck-Institut für Meteorologie, Hamburg, 1995b.

P.C. Stern, O.R. Young and D. Druckman (eds.), *Global Environmental Change: Understanding the Human Dimensions, National Research Council*, National Academy Press, Washington, 1992 (cited in Swart, 1994a).

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

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

R. Swart, M. Berk, E. Kreileman, M. Janssen, J. Bollen, R. Leemans, and B. de Vries, *The Safe Landing Approach: Risks and Trade-offs in Climate Change*, RIVM, Bilthoven, 1996.

K.E. Taylor and J.E. Penner, Response of the Climate System to Atmospheric Aerosols and Greenhouse Gases, *Nature*, **369**, 1994, p. 734-737.

H. Tennekes, Karl Popper and the Accountability of Numerical Weather Forecasting, *Weather*, **47** (9), 1992, p. 343-346.

H. Tennekes, The Limits of Science, The Conversion of a Meteorologist to Ecology, in: Zweers and Boersma (eds.), Ecology, Technology and Culture, White Horse Press, UK, 1994.

J.G. Titus and V. Narayanan, The Risk of Sea Level Rise, A Delphic Monte Carlo Analysis in which Twenty Researchers Specify Subjective Probability Distributions for Model Coefficients within their Respective Areas of Expertise, in: *Climatic Change*, **33** (2), 1996, p.151-212.

R.S.J. Tol, *The Climate FUND, Modeling Costs and Benefits*, Institute for Environmental Studies, December 1993.

R.S.J. Tol, *The Climate FUND, Optimal Greenhouse Gas Emission Abatement*, Institute for Environmental Studies, 7 May 1994.

R.S.J. Tol (ed.), *Integrated Assessment of Climate Change, Report EMF14 Meeting #2*, IvM, Free University of Amsterdam, 1994.

R.S.J. Tol (ed.) *Integrated Assessment of Climate Change, Report EMF14 Meeting #3*, IvM, Free University of Amsterdam, 1995

M. Toman, Climate Change, Cataclysmic Risk and Sustainable Development: An Economic Perspective, paper presented at IIASA/LOS meeting on *Climate Change, Cataclysmic Risk and Fairness*, IIASA, Laxenburg, 20-22 July 1996.

F.L. Toth, Practice and Progress in Integrated Assessments of Climate Change, A Workshop Overview, in: *Energy Policy*, **23** (4/5), 1995, p. 253-267.

K.E. Trenberth (ed.), *Climate System Modeling*, Cambridge University Press: Cambridge, 1992, 788 pp.

W.C. Turkenburg and A.J.M. Van Wijk, *Onze Kennis over Klimaatverandering, Verslag van een Workshop,* Report Lucht 99, DOP (Depot Overheids Publicaties), Leiden, 1991.

A. Trisoglio, *Plurality and Political Culture, A Response to Thompson and Drake*, paper Presented at workshop "Risk, Policy and Complexity", IIASA, Laxenburg, 7-9 August 1995, 18 pp.

M.B.A. van Asselt and J. Rotmans, Uncertainty in Perspective, in: *Global Environmental Change, Human and Policy Dimensions*, (forthcoming, scheduled for the September 1996 issue).

M.B.A. van Asselt, A.H.W. Beusen, and H.B.M. Hilderink, Uncertainty in Integrated Assessment, A Social Scientific Perspective, in: *Environmental Modeling and Assessment*, **1** (1), forthcoming.

J.P. van der Sluijs, Uncertainty Management in Integrated Modeling, the IMAGE case, in: S. Zwerver, R.S.A.R. van Rompaey, M.T.J. Kok, and M.M. Berk (eds.), *Climate Change Research; Evaluation and Policy Implications*, Studies in Environmental Science 65 B, Elsevier Science B.V., 1995.

J.P. van der Sluijs, C.A. Hendriks, K. Blok, "Feasibility of polymer membranes for carbon dioxide recovery from flue gases", In: *Energy Conversion and Management*, **33**, 429-436 (1992) (Proc. First Int. Conf. on Carbon Dioxide Removal).

J.P. van der Sluijs, G.J. de Bruyn, and P. Westbroek, Biogenic Feedbacks in the Carbonate-Silicate Geochemical Cycle and the Global Climate, *American Journal of Science*, (forthcoming, scheduled for the October 1996 issue).

J.P. van der Sluijs, J.C.M. van Eijndhoven, B. Wynne, and S. Shackley, Anchoring Devices in Science For Policy: The Case of Consensus Around Climate Sensitivity, *Social Studies of Science* (forthcoming, scheduled for the spring 1997 issue).

J.P. van der Sluijs and J.C.M. van Eijndhoven, Closure of Disputes in the Assessments of Climate Change in the Netherlands Arena, submitted to *Environmental Management*.

J.C.M. van Eijndhoven and P. Groenewegen, The Construction of Expert Advice on Health Risks, *Social Studies of Science*, **21**, 1991, p. 257-278.

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

T.S. Wallsten, Measuring Vague Uncertainties and Understanding their use in Decision Making, in: G.M. von Furstenberg (ed.) *Acting under Uncertainty, Multidisciplinary Conception*, Kluwer Academic Publishers, Dordrecht, 1990.

W.G. Wang, Y.L. Yung, A.A. Lacis, T. Mo, and J.E. Hansen, Greenhouse effects due to man-made perturbations of trace gases, *Science*, **194**, 1976, p.685-690.

WCPR, *World Climate Research Program*, 1979.

WEC and IIASA, *Global Energy Perspectives to 2050 and Beyond*, World Energy Council, London, 1995.

P. Westbroek, *Life as a Geological Force, Dynamics of the Earth*, W.W. Norton & Compagny, New York, 1991, 240 p.

P. Westbroek and G. Muyzer, Modeling the Biosphere, *Terra Nova*, **4**, 1992, p. 539-541.

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.

E. Worrell, Potentials for Improved Use of Industrial Energy and Materials, Ph.D. Thesis, Utrecht University, 1994, 183 pp.

B. Wynne, Uncertainty and Environmental Learning, in: *Global Environmental Change*, **2**, 1992, p.111-127.

P. Young and S. Parkinson, Simplicity out of Complexity in Forecasting Environmental Change, in: B. Beck (ed.) *Forecasting Environmental Change* (forthcoming monograph).

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

APPENDIX

List of Integrated Assessment Models for the climate issue

Model	Full name	Source
AIM	Asian-Pacific Integrated Model	I;D;W;E
AS/ExM	Adaptive Strategies/Exploratory Model	I
ASF (EPA)	Atmospheric Stabilization Framework	
CETA	Carbon Emission Trajectory Assessment	I;D;W;E
Connecticut	(model by G. Yohe)	I;E
CRAPS	Climate Research And Policy Synthesis model	I;E
CSERGE	Centre for Social and Economic Research into the Global Environment	I;W;E
DGEM		D
DIALOOG	(Model developped by KEMA for a dialogue with the Netherlands electricity sector)	R
DIAM	Dynamics od Inertia and Adaptability Model	I
DICE	Dynamic Integrated Climate Economy model	I;D;W;E
Edmonds-Reilly-Barns		D
FUND	The Climate Framework for Uncertainty, Negotiation and Distribution	I;E
GCAM	Global Change Assessment Model	W
GEMINI		D
GLOBAL 2100		D
ICAM-2	Integrated Climate Assessment Model	I;D;W;E
IIASA	International Institute for Applied Systems Analysis	I;E
IMAGE 1	Integrated Model to Assess the Greenhouse Effect	D
IMAGE 2	Integrated Model to Assess the Greenhouse Effect	I;D;W;E
ISM	Integrated Science Model for assessment of climate change	W

MAGICC	Model for the Assessment of Greenhouse gas Induced Climate Change	W
MARIA	Multiregional Approach for Resource and Industry Allocation	I;E
MARKAL	Market Allocation model	D
MBIS	Mackenzie Basin Impact Study	D
MCW (WRI)	Model of Global Warming Commitment	
MERGE 2	Model for Evaluating Regional and Global Effects of GHG Reduction Policies	I;W;D;E
MIT	Massachusetts Institute of Technology	I;W;E
MiniCAM *	Mini Climate Assessment Model	I;W
MORI	(Also known as MARA of DICE+e)	E
New EARTH 21		W
OECD-GREEN		D
PAGE	Policy Analysis of the Greenhouse Effect	I;D;W
PEF	Policy Evaluation Framework	I;D;W
PoleStar	(Stockholm Environmental Institute)	
ProCAM	Process Oriented Global Change Assessment Model	I;D
RAND		E
RICE	Regional DICE	I
SLICE	Stochastic Learning Integrated Climate Economy Model	I
TARGETS	Tool to Assess Regional and Global Environmental and Health Targets for Sustainability	I;D;W

Explanation of the codes in the third column: D = Mentioned in the overview by Dowlatabadi (1995); W = Mentioned in the overview by Weyant (1994); E = Part of the comparison by the Energy Modeling Forum (1995b, 1996b); I = Mentioned in the overview by IPCC WG3 (Weyant *et al.*, 1996); R = Personal communication, Walter Ruigrok, KEMA (17 Sept. 1996).