Defining projects and scenarios for integrated assessment modelling using ontology

Janssen, S.J.C.[1,2], J.J.F. Wien[3], Hongtao Li[4], I. N. Athanasiadis[4], F. Ewert[1], M.J.R. Knapen [3], D. Huber[5], O. Thérond[6], A.E. Rizzoli[4], H. Belhouchette[7], M. Svensson[8] and M.K. van Ittersum[1]

[1] Wageningen University, Plant Production Systems Group, Wageningen
[2] Wageningen University, Business Economics Group, Wageningen
[3] Alterra, Environmental Sciences Group, Wageningen University and Research Centre, Wageningen
[4]Dalle Molle Institute for Artificial Intelligence (IDSIA), Lugano
[5] AntOptima SA, Lugano
[6]INRA- UMR-AGIR, Toulouse
[7] INRA-Agro M. UMR SYSTEM, Montpellier
[8] Lund University, Centre for Sustainability Studies, Lund
Email: sander.janssen@wur.nl

Keywords: *Scenarios, integrated assessment, modelling, ontology*

EXTENDED ABSTRACT

Integrated Assessment Modelling provides a systematic inter-disciplinary approach to support coherent ex-ante decision-making by a flexible integration of (reusable) models and datasets across scales. Within integrated assessment modelling a coherent and robust description of projects and scenarios is required to facilitate data preparation, model integration and the graphical user-interface development.

This paper explains our experiences with a challenging and time-consuming task, e.g. arriving at a shared understanding on the definition of projects, experiments and scenarios among researchers coming from different disciplines, who have been exposed to dissimilar education and research experience. We demonstrate the use of ontologies in building this shared set of definitions and the relationship between the ontology and the human computer interaction through a case study. With a common ontology that represents the joint conceptualization of the projects, experiments and scenarios each researcher can refer at any later stage to the semantics of the concepts used. A collaborative approach was used to build such a common ontology in the SEAMLESS-Integrated Project, funded through the EU sixth Framework Programme, which aims at developing an integrated modelling framework (SEAMLESS-IF) to assess, ex-ante, agricultural and environmental policy options, allowing cross-scale analysis of a broad range of sustainability issues.

Through several iterations a common ontology for projects, experiments and scenarios was built. In our common ontology a project has one and only one problem definition, and it can handle at least one or more *Experiments.* Experiments represent the assessment of one or a combination of policy options in a given context and outlook on the future. The *indicator(s)* should be the same between experiments which are part of the same project, allowing the comparison of different experiments. Each of the concepts *Policy Option, Context and Outlook* capture one part of the input parameters required for running each of the models.

As a first validation of the project ontology, a set of four fictitious sample projects were made. One of these sample projects is an integrated assessment for one region Midi-Pyrénées in the South of France concerning the impacts of the CAP2003 reform, which is described in this paper

The common project ontology highlighted the imprecise meaning of the word *scenario* and it links projects to problems, outlooks on the future, indicators, context of the problem, policies and ultimately to model runs in experiments. Also, by this common ontology the assumptions in building the assessment are clarified, moving the focus away from the tools to the assumptions underlying models and scenarios. In any integrated assessment project, it is recommended to clarify with its participants the meaning of scenario and associated concepts. We achieved this by the use of a common ontology, which forces participants to be clear, precise and coherent in their description of concepts and relationships between concepts, while the common ontology can be directly used for development of databases, models and graphical user interfaces.

1. INTRODUCTION

Integrated Assessment Modelling (IAM) is more and more frequently used methodology to assess the impacts of policies, technologies or societal trends on the environmental, economic and social future sustainability (Parker *et al.*, 2002), for example in mitigation to climate change (Weyant *et al.*, 1996; Cohen, 1997) or water quality in catchment areas (Turner *et al.*, 2001). Integrated assessments are defined by Rotmans and Van Asselt (1996) as an interdisciplinary, participatory and future-oriented process of combining, interpreting and communicating knowledge from diverse scientific disciplines to allow a better understanding of complex phenomena.

Scenario analysis is identified as an important tool in integrated assessment (Rotmans, 1998), where scenarios are used in the interaction between scientists and stakeholders to describe what the future could be. Many different definitions of scenario exist in scenario literature. For example, Rotmans (1998) defines scenarios as ‘archetypal descriptions of alternative images of the future, created from mental maps or models that reflect different perspectives on past, present and future developments’, while Parry and Carter (1998) define scenarios as ‘a coherent, internally consistent and plausible description of a possible future state of the world.’ Peterson *et al.* (2003) provides a definition of scenario which is closer to modelling, ‘as variation in the assumptions used to create models.’

Given that a wide range of definitions is available for scenarios, there is a risk for confusion and misunderstanding in any integrated assessment modelling project. For example, does scenario refer to the outcomes of model runs? Or does it refer to the set of input parameters to a model? Is it only related to policies as suggested by the term policy scenario or is it broader? This reinforces the need for a clear set of rules and protocols for integrated assessment, in particular with respect to the understanding of scenarios, as concluded by Rotmans and Van Asselt (1996), to avoid the dangers of unclear, inconsistent, narrowly-defined scenarios and ad-hoc setting of parameters (Rotmans, 1998; van Asselt and Rotmans, 2002).

SEAMLESS is an integrated assessment modelling project (Van Ittersum *et al.*, 2007), which aims to provide an computerized framework to assess the sustainability of agricultural systems in the European Union at multiple scales. As in SEAMLESS over one hundred of scientist participate from different disciplines and dissimilar research background, many different views exist on the meaning of scenario and its implications for the computerized SEAMLESS-IF framework.

As SEAMLESS builds a computerized framework, one and only unified view on the meaning of scenarios and the definition of assessment projects is required within the group of scientist to enable data preparation, model integration and graphical user-interface development. There is no established procedure to develop such a unified view on the meaning of scenarios and the definition of assessment projects. This paper explains our experiences with this challenging and time-consuming task, e.g. arriving at a shared understanding on the definition of projects and scenarios among researchers from different disciplines with dissimilar research experience. We demonstrate the use of common ontologies in building this shared conceptual model through a case study.

In the next section, some background will be provided on common ontologies and the process of ontology engineering. Also, our case study set-up will be introduced for the SEAMLESS-project. In Section 3, the shared conceptual model on scenario and project definition is presented, supported by a fictitious example of the use of the common concept in a regional integrated assessment project, while this common concept will be discussed in Section 4. In the final section we highlight the most important lessons we learned in our case study.

2. MATERIALS AND METHODS

2.1. Common Ontologies

In the context of integrated modelling, ontologies are useful to define the shared conceptualization of a problem, as ontologies are written in a language, e.g. Web Ontology Language, that is understandable by computers and as ontologies consists of a finite list of concepts and the relationships between these concepts (McGuinness and van Harmelen, 2004). The term ontology originates from philosophy, originally coined by classical philosophers Plato and Aristotle (Aristotle, 336-332 BC) in the study of types of being and their relationships (metaphysics). An ontology in computer science is considered as a specification of a conceptualization (Gruber, 1993), where a conceptualization is ‘an abstract, simplified view of the world “e.g. systems under study (addition by author)” that we wish to represent for some purpose’ (Gruber, 1993). In integrated modelling research, scientists from various disciplines can define a common conceptual schema that their domains share. A common project ontology, i.e. ontology which is shared by all domains to-be-integrated , serves as a knowledge-level specification of the joint conceptualization of the project and scenario definition. Each scientist can refer to and should adhere to the semantics of the concepts in the common project ontology, including restrictions on the concepts and relationships between the concepts.

2.2. Ontology Engineering

In developing a common ontology, the scientific challenge of adopting tight, well-reasoned and shared conceptualizations among a group of scientists or one individual scientist should be overcome. The development of a common ontology by a group of researchers is a complex, challenging and time-consuming task (Musen, 1992; Gruber, 1993; Farquhar *et al.*, 1995; Holsapple and Joshi, 2002). Tools are available that help in ontology development (Farquhar et al., 1995) and to store the ontology once it was developed (Knublauch, 2005). To achieve ontological commitment, i.e. the agreement by multiple parties to adhere to a common ontology, when these parties do not have the same experiences and theories (Holsapple and Joshi, 2002), a collaborative approach is suggested to be used. A collaborative approach has the advantages that researchers from different disciplines are diverse in their contributions, which avoids blindspots and which has more chances of getting a wide acceptance (Holsapple and Joshi, 2002) and that it can incorporate the other approaches, e.g. synthetic approach, as required for development of parts of the ontology.

2.3. Case Study: SEAMLESS project

The SEAMLESS integrated project (System for Environmental and Agricultural Modelling; Linking European Science and Society), EU sixth Framework project, develops a computerized and integrated framework (SEAMLESS-IF) to assess the impacts on environmental and economic sustainability of a wide range of policies and technological improvements across a number scales. This aim should be achieved by overcoming the gap between micro-marco level analysis, overcoming the bias in integrated assessments towards either economic or environmental issues, facilitating the re-use of models and providing methods to technically link different models together (Van Ittersum *et al.*, 2007).

Within SEAMLESS, both modelling and stakeholder involvement are seen as important elements of the assessment procedure proposed by SEAMLESS-IF. With respect to modelling, macro-level economic partial or general equilibrium models are linked to micro-level farm optimization models and field crop growth models, while in between macro and micro-level steps of aggregation and des-aggregation occur by other models. A participatory approach is foreseen for the use of SEAMLESS-IF with stakeholders at the end of the project. The SEAMLESS-IF should be designed to facilitate such a participatory approach. Prime Users as the Directorates General of the European Commission are involved in this process through a User Forum. In the project 30 partners and more than 100 researchers participate.

Thus, the common ontology for project and scenario definition acts on these interfaces between modellers and other scientists and between scientist and stakeholders after the development of the SEAMLESS-IF (Fig. 1). The common ontology is used to construct the database schema to store data on projects and scenarios and Java^{tm}-beans (Athanasiadis *et al.*, 2007) for development of the Graphical User Interface (GUI) and for structuring input for the models.

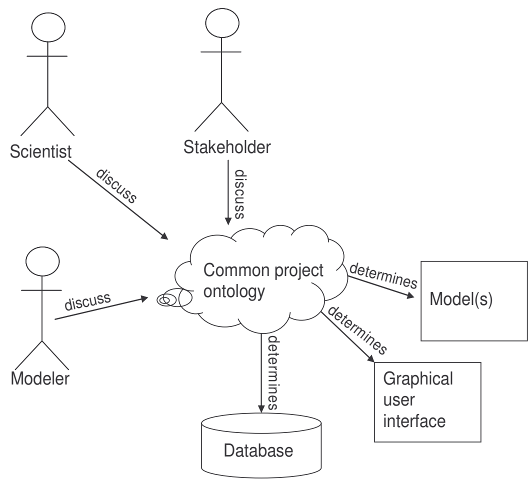

Figure 1 Role of a common project ontology in a integrated assessment modelling project

The collaborative approach for ontology engineering in our case study was based on developing one shared document between a group of seventeen researchers working in different parts of the SEAMLESS project. Ten iterations of the document were used and after each iteration a small ontology constructed in Protégé OWL (Knublauch, 2005) was synchronized with the iteration. With each iteration, more scientist were involved starting from four for this first iteration up to seventeen for the tenth iteration. At the tenth iteration both the document and the ontology were 'closed' after the approval of the SEAMLESS management group and a set of actions was formulated to elaborate specific parts of the project and scenario definition.

3. RESULTS

3.1. The project ontology

A project ontology was developed for integrated assessments as carried out in the SEAMLESS-projects, and here, a description in words is provided, together with an overview in diagrams. In SEAMLESS, we foresee in an application of the SEAMLESS-IF an integrative modeller (scientist) working together with a policy expert or other stakeholders. In general, a project in SEAMLESS refers to the assessment of the effects on agricultural sustainability and profitability of changes in policies. When the integrative modeller wants to work with the SEAMLESS-IF, he/she always has to build a project that will handle the specific problem the integrative modeller wants to tackle, after discussing it with a policy expert. A project has one and only one problem definition, and it can handle at least one or more *Experiments* (Fig. 2). This implies that different perspectives on a problem can be investigated through different Experiments, representing the assessment of one or a combination of policy options in a given context and outlook on the future.

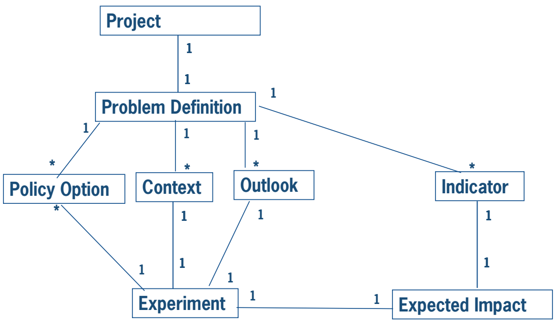

Figure 2. A data model describing the project ontology.

Thus, an experiment is ONE run of the models within SEAMLESS-IF that assesses ONE or A COMBINATION of *policy option(s)* within a context and an outlook on the future. The *indicator(s)* should therefore be the same between experiments which are part of the same project, allowing the comparison of different experiments. Each quantitative indicator selected for a project gets one value for each Experiment. *Impacts* are the changes in indicator-value due to changes in policy options, context and outlook on the future as compared to a reference situation (Fig. 3). Expected impacts are then changes in indicator values defined by the Policy Expert before running the experiment in the SEAMLESS-IF, while calculated impacts are the changes in indicator values as calculated by the SEAMLESS-IF.

Each of the concepts *Policy Option*, *Context* and *Outlook* capture one part of the input parameters required for running each of the models. These parameters are unchangeable by the models, meaning that the model run does not affect the value of the parameter, and so these are exogenous. Next to these exogenous parameters, the models have endogenous parameters, which are parameters that can be changed by other models in the model chain. A policy option refers to one or more policy measures as part of it and each policy option is described by a set of parameters that are exogenous to the models. An example of a policy option is a reform of the Common Agricultural Policy of the European Union, which has as measures a decoupling of the subsidy payments from the area of the farm to income support and the lowering of the support prices for products. Relevant policy parameters are for this example the percentage of decoupling for a region, the reference yield for a region and the cut in premiums.

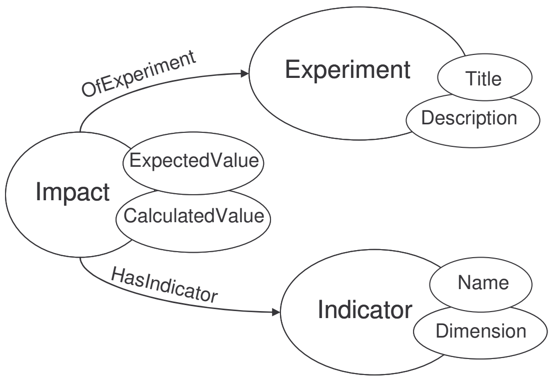

Figure 3. An ontology snapshot showing the concepts 'Impact', 'Indicator' and 'Experiment' (large circles), their relationships (arrows) and their data-properties (small circles).

Outlook on the future describes trends and trend deviations foreseen to occur in society that might affect the implementation of policy options within a given context. These trends or trend deviations are not modelled endogenously in SEAMLESS. One reference outlook is always required that describes the prolongation of the current situation into the future, sometimes called *business-as-usual* outlook. Outlooks are usually contrasting, for example a positive versus a negative outlook, a globalization versus a regionalization outlook. Each outlook has several exogenous parameters that capture the different trends occurring in society. Examples of these parameters are atmospheric CO_2-concentration, GDP-growth and unemployment rate.

Finally, the context of a problem defines the object of interest, which is delimited by the boundaries to the biophysical and agro-management system. These boundaries determine what is inside and what is outside the system. Each experiment within a problem will be based on one agro-management and biophysical context that can be different from those of other experiments.

The experiments thus define the changes or driving forces as compared to the reference situation, by capturing the changes in policy options, context and outlook, either as changes in isolation (only one policy option/outlook/context-change) or simultaneously (more than one policy option/ outlook/ context-change). This implies that the maximum number of experiments to be defined is the factorial combination of policy options, outlooks and context, but not all of these experiments make sense.

3.2. An example project

As a first validation of the project ontology, a set of four fictitious sample projects were made. One of these sample projects is an integrated assessment for one region Midi-Pyrénées in the South of France concerning the impacts of the CAP2003 reform as requested by Mrs X and Mr. Y of a regional government agency. They also want to evaluate if CAP2003 reform will favour conservation agriculture in the Midi-Pyrénées region and they are curious to know if a subsidy on conservation agriculture would increase the uptake of conservation agriculture. They discuss their assessment with three scientists, working at a research institute in Toulouse, Midi-Pyrénées.

The three scientists define the following two policy options, based on their discussions with the policy experts. The policy option *CAP2003 reform* comprises a set of European Union policies, related to income support for farmers, while the policy option *subsidies for conservation agriculture* also comprise a set of policies, but targeted at the sustainability of agriculture in terms of soil conservation.

After discussing with the policy experts, the scientists think they need three outlooks. One *business-as-usual*-outlook in which there are no trend deviations in society, so the current situation is prolonged. Next to this a *globalization* and *regionalization* outlook are defined. In the globalization outlook, atmospheric CO2-concentration is expected to rise steeply by 5%, and unemployment in Midi-Pyrénées is expected to decrease by 3%. In the regionalization Outlook atmospheric CO2-concentration is expected to rise mediocre by 2%, and unemployment in Midi-Pyrénées is expected to increase by 3%. (Table 1)

Table 1. Outlooks for the fictitious example concerning Midi-Pyrénées

Outlook	Atmospheric CO_2 concentration	Unemployment
Business-as-usual	No change	No change
Globalization	+5%	-3%
Regionalization	+2%	+3%

Two contexts should be defined. These contexts describe the technological innovation as driving forces. One context is the situation without the possibility for the farmers to choose conservation agriculture, and in the other context farmers have the option to choose conservation agriculture on the farm.

The scientists define four experiments, which are a combination of one context, one outlook and one or a combination of policy options (Table 2). The first experiment is the *business-as-usual* experiment where no CAP 2003 reform takes place, the farmers do not have options for conservation agriculture and the outlook for the future is that no particular important trends occur. This experiment acts as a reference point for comparison of the other experiments (Table 2).

Table 2. The experiments for the fictitious sample project in Midi-Pyrénées

Experiments	Policy option	Outlook	Context
1. Business as Usual	Only current policies	Business as Usual	No conservation agriculture
2. CAP 2003 reform	CAP 2003 Reform	Globa-lization	No conservation agriculture
3. No support	CAP 2003 Reform	Regiona-lization	Conservation agriculture
4. Conservation oriented in regional world	CAP 2003 Reform and subsidies for conservation agriculture	Regiona-lization	Conservation agriculture

Finally, the scientist discuss with the policy experts the relevant indicators, which are the regional cropping pattern, the farmer income, the amounts of subsidies, the % of no-plowing tillage, the area for the intercrops mustard and clover and the level of erosion.

4. DISCUSSION

4.1. Scenario and its meaning

The fictitious example for the integrated assessment in the Midi-Pyrénées region demonstrated that all relevant aspects of the integrated assessment could be captured by the project ontology. In our project ontology as presented in the previous Section, we have no concept *Scenario* as part of it. In the iterative process of building the common project ontology, it was apparent that scenario had different meanings for scientists. Some scientists thought of scenarios as experiments, so a perspective of what could change in terms of policies, outlooks and context, and thereby determining the input parameters for the models. Other scientists thought of scenarios as a set of impacts, so values for indicators that were the 'end-result' of changes in policies, outlooks and contexts. Economic modellers limited their definition of scenario to policy options, while biophysical modellers were more inclined to think of scenario as outlook. The concept *scenario* was thus not included in the project-ontology for risk of confusion and other concepts were chosen that could be un-ambiguously defined and agreed upon. This proves that the project ontology is able to cover all the different meanings scenario can have, and offers an opportunity to comprehensively describe an integrated assessment problem. One definition of scenario could be centrally decided on by scientists participating in an integrated assessment.

4.2. Project ontology and models

Models are not explicitly mentioned as a separate concept in the project ontology and the fictitious sample projects as presented in Section 3, although a link exists between the properties of the context, outlook and policy option and input parameters for the models. The required models to analyse an assessment problem can be deduced from the selections by the integrative modeller with respect to the properties of context, policy option and outlook. This allows to focus on the assumptions made while defining values for the different model input parameters and defining the experiments instead of focusing solely on the technical capabilities of the models (Rotmans, 1998; Greeuw *et al.*, 2000). Many different types of models could be linked to the project ontology, for example optimization models and deterministic simulation models.

4.3. Use of ontologies and ontology engineering

By using ontology engineering as our methodology, scientists participating in this collaborative process had to be precise in their meaning of concepts they proposed for the common ontology, and they had to ensure consistency and coherence between the concept they proposed and the other concepts in the project ontology. With ten iterations and seventeen participating scientist, the collaborative approach required a clear objective and set of actions for each iteration, which lead it to be a time-consuming task. The collaborative approach was an appropriate solution in our case, as knowledge from scientist from different domains could be disclosed and as the project ontology has become a common reference point for scientists in the project, reflecting the shared understanding.

4.4. Future developments

More evaluation is required by peer review and stakeholders outside the science community of the common project ontology. Also, the concept of 'scale' needs to be included, which is recognised as an important concept in integrated assessments (Parker *et al.*, 2002).

5. CONCLUSIONS AND RECOMMENDATIONS

As concluded by Cohen (1997), 'if stakeholders and their knowledge are to be drawn into integrated assessments, integrated assessments must become less of a "black box" and more human.' Through the development of a common ontology on projects and scenarios, we opened the black box surrounding scenarios in terms of meaning and content. Our common ontology improves the consistency and transparency of scenarios, as (i) a set of concepts is provided to describe different types of model input parameters, as (ii) the focus is more on assumptions made in defining these input parameters instead of the tools themselves and as (iii) experiments should be explicitly constructed capturing the different perspectives on the future.

In any integrated assessment project, it is recommended to clarify with its participants the meaning of scenario and associated concepts. We achieved this by the use of a common ontology, which forces participants to be clear, precise and coherent in their description of concepts and relationships between concepts, while the common ontology can be directly used for development of databases, models and graphical user interfaces.

6. ACKNOWLEDGMENTS

We thank all scientists in the SEAMLESS project who contributed to development of the common ontology on projects and scenarios. This work has been carried out as part of the SEAMLESS Integrated Project, EU sixth Framework Programme, Contract No. 010036-2.

7. REFERENCES

Aristotle (336-332 BC), Metaphysics,

Athanasiadis, I. N., F. Villa and A. E. Rizzoli (2007), Enabling knowledge-based software engineering through semantic-object-relational mappings, Paper presented at 3rd International Workshop on Semantic Web Enabled Software Engineering, 4th European Semantic Web Conference, 15p., Innsbruck, Austria.

Cohen, S. J. (1997), Scientist–stakeholder collaboration in integrated assessment of climate change: lessons from a case study of Northwest Canada, *Environmental Modelling and Assessment* 2(4), 281.

Farquhar, A., R. Fikes, W. Pratt and J. Rice (1995), Collaborative Ontology Construction for Information Integration, Secondary Collaborative Ontology Construction for Information Integration, 32 p. Knowledge Systems Laboratory, Department of Computer Science, Stanford University,

Greeuw, S. C. H., M. B. A. Van Asselt, J. Grosskurth, C. A. M. H. Storms, N. Rijkens-Klomp, D. S. Rothman and J. Rotmans (2000), Cloudy crystal balls- An assessment of recent European and global scenario studies and models, Secondary Cloudy crystal balls- An assessment of recent European and global scenario studies and models, 112 p. European Environment Agency, Copenhagen

Gruber, T. R. (1993), A Translation Approach to Portable Ontology Specifications, *Knowledge Acquisition* 5, 199-220.

Holsapple, C. W. and K. D. Joshi (2002), A collaborative approach to ontology design, *Communicationss of the ACM* 45(2), 42-47.

Knublauch, H. (2005), Protege OWL, p., Stanford: Stanford Medical Informatics

Mcguinness, D. and F. Van Harmelen (2004), OWL Web Ontology Language Overview, Secondary OWL Web Ontology Language Overview, p. WWW Consortium,

Musen, M. A. (1992), Dimensions of knowledge sharing and reuse, *Computers and Biomedical Research* 25(5), 435-467.

Parker, P., R. Letcher, A. Jakeman, M. B. Beck, G. Harris, R. M. Argent, M. Hare, C. Pahl-Wostl, A. Voinov and M. Janssen (2002), Progress in integrated assessment and modelling, *Environmental Modelling & Software* 17(3), 209-217.

Parry, M. and T. Carter (1998), Climate Impact and Adaptation Assessment, Earthscan Publications Ltd., London, UK

Peterson, G. D., G. S. Cumming and S. R. Carpenter (2003), Scenario Planning: a Tool for Conservation in an Uncertain World, *Conservation Biology* 17(2), 358-366.

Rotmans, J. (1998), Methods for IA: The challenges and opportunities ahead, *Environmental Modelling and Assessment* 3(3), 155.

Rotmans, J. and M. Asselt (1996), Integrated assessment: A growing child on its way to maturity, *Climatic Change* 34(3), 327.

Turner, R. K., L. Ledoux and R. Cave (2001), The Use of Scenarios in Integrated Environmental Assessment of Coastal-Catchment Zones, Secondary The Use of Scenarios in Integrated Environmental Assessment of Coastal-Catchment Zones, 25 p. School of Environmental Sciences, University of East Anglia, Norwich

Van Asselt, M. B. A. and J. Rotmans (2002), Uncertainty in Integrated Assessment Modelling, *Climatic Change* 54(1), 75.

Van Ittersum, M. K., F. Ewert, T. Heckelei, J. Wery, J. Alkan Olsson, E. Andersen, I. Bezlepkina, F. Brouwer, M. Donatelli, G. Flichman, L. Olsson, A. Rizzoli, T. Van Der Wal, J.-E. Wien and J. Wolf (2007), Integrated assessment of agricultural systems- a component based framework for the European Union (SEAMLESS), *Agricultural Systems* In Press.

Weyant, J., H. Davidson, H. Dowlatabadi, J. Edmonds, M. Grubb, R. Richels, J. Rotmans, P. Shukla, R. S. J. Tol, W. Cline and S. Fankhauser (1996), Integrated Assessment of climate change: An overview and comparison of approaches and results, In Climate Change 1995 - Economic and Social Dimensions Eds J. P. Bruce, H. Lee & E. F. Haites), Cambridge University Press, Cambridge