Managing the Hydra in integration: developing an integrated assessment tool for agricultural systems

Sander J.C. Janssen

Thesis committee

Thesis supervisors

Prof. dr. ir A.G.J.M. Oude Lansink

Professor of Business Economics

Wageningen University

Dr. ir. M.K. van Ittersum

Associate Professor of Plant Production Systems

Wageningen University

Other members

Prof. ir. A.J.M. Beulens, Wageningen University

Prof. dr. T. Heckelei, University of Bonn, Duitsland

Prof. dr. H.B.J. Leemans, Wageningen University

Prof. dr. H.B. Meinke, Wageningen University

This research was conducted under the auspices of the C.T. de Wit Graduate School Production Ecology and Resource Conservation

Managing the Hydra in integration: developing an integrated assessment tool for agricultural systems

Sander J.C. Janssen

Thesis
submitted in partial fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Doctorate Board
to be defended in public
on Tuesday 8 September 2009
at 11 AM in the Aula.

Sander J.C. Janssen
Managing the Hydra in integration: developing an integrated assessment tool for agricultural systems
272 pages

Thesis Wageningen University, Wageningen, NL (2009).
With references, with summaries in Dutch and English

ISBN 978-90-8585-436-4

Abstract

Ex-ante assessment through science-based methods can provide insight into the impacts of potential policy measures or innovations to manage complex problems (e.g. environmental pollution, climate change, or farmers' welfare). Integrated Assessment and Modelling (IAM) is a method that supports ex-ante assessment through modelling and modelling tools. One type of IAM links models focusing on particular processes on a specific scale into model chains covering multiple scales and disciplines. To achieve an operational model chain for IAM, methodological, semantic and technical integration is required of models, data sources, indicators and scenarios. In this thesis, methodological, semantic and technical integration focuses on two case studies. The first case study is on integration within bio-economic farm models covering two hierarchical systems levels involving a small team of scientists. The second case refers to modelling European agricultural systems. In this case, the integration covers five hierarchical systems levels and different types of models were linked by a large team of about hundred scientists. In the context of these two case studies, many different integration topics and challenges have been addressed: a review of the state-of-the-art in bio-economic farm models, a generic method to define alternative agricultural activities, development of a generic bio-economic farm model, development of an integrated database for agricultural systems, linking different agricultural models and a shared definition of scenarios across disciplines, models and scales. Ultimately, elaborating the methodological, semantic and technical integration greatly contributed to the development of an integrated assessment tool for European agricultural systems. This integrated assessment tool can be used across disciplines and for multi-scale analysis, and allows the assessment of many different policy and technology changes.

Keywords: modelling, bio-economic, farm, simulation, ontology, knowledge management, Europe, agricultural management, database, scenario

Preface

Alea iacta est (Julius Caesar, 49BC)

'Alea iacta est' is the best way to describe my feeling in finishing this PhD-thesis. I have reached the point of no return and such an end-point is always sad and happy at the same time. The PhD thesis that now lies in front of you is the outcome of four years of work on integration in science and bio-economic farm models in the context of the SEAMLESS project. This project provided an excellent opportunity for me to do a PhD. I have learned a tremendous amount from a lot of different people and I developed not only as a scientist, but also as an European by working with many different cultures, as a team worker in trying to achieve a joint goal and a leader organising a task force on semantic integration.

What remains now is to thank a (large) number of people, but one person in particular. Martin, thanks, thanks, thanks! I have learned a lot from our cooperation and I really appreciate your supervision during my PhD time. I have great respect for the way you managed the SEAMLESS project, while still having time to read and in detail consider my emails, notes, discussion pieces and draft articles. Thanks also for your patience with my often hastily compiled draft articles and the good and constructive discussion we had about them. I am grateful for the confidence and room for manoeuvre I got to develop activities within the SEAMLESS project, even if neither you nor me were specialized in it. I am looking forward to future collaboration.

I would also like to thank Alfons, my promotor for his supervision and support. Your reflections on my articles helped to really shape and improve them as scientific publications. I can imagine it was not always easy to have a PhD so emerged in a complex project like SEAMLESS.

During these four years, I had many colleagues at many different locations across Wageningen (Plant Production Systems Group, Business Economics Group, Alterra Team Systems) and Europe (SEAMLESS partners in many countries, with even one in the USA). Thanks all for the nice work atmosphere and collaboration we had in

those four years. I always felt at ease in each different environment. Special thanks to my roommates Argyris Kanellopoulos, Lenny van Bussel, Lusine Aramyan, Grigorios Emvalomantis, William Bellotti, Felix Bianchi and Marrit van den Berg. Anne Houwers and Karin Sijlmans, thanks for taking care of many things, as arranging all my trips and a lot of things that I probably do not realize. Also, thanks to my paranymphs Ioannis Athanasiadis and Wanne Kromdijk. I would like to thank Ken Giller for helping to convince me to do a PhD at his group.

Elizabeth, thank you for your support, love, laughs and healthy skepticism of my multi-disciplinary research. Norah, you are my sunshine. My parents John and Christien and sister Maaike, you are always close and around, even if you are far or very far away.

Sander Janssen, Wageningen, July 14 2009

Table of Contents

Index of Tables

Index of Figures

To Elizabeth and Norah

'Yet he, Herakles, son of Zeus, of the line of Amphitryon, by design of Athene the spoiler and with help from warlike Iolaos, killed this beast with the pitiless bronze sword.' Hesiod, Theogony (Greek epic 7th B.C.)

Chapter 1. Introduction to integration in Integrated Assessment and Modelling and bio-economic farm models

1.1. Introduction

1.1.1. Ex-ante assessment

For decision makers in the domains of agriculture and environment, for instance in government agencies, farmers, environmental NGOs and farmers' unions, it is beneficial to evaluate ex-post or to asses ex-ante the impacts of their choices. An ex-post evaluation occurs after such a choice has been made, while an ex-ante assessment tries to simulate the potential impacts of choices before these are made. In ex-post evaluation, data is likely to be available or can be collected on relevant variables in the period after the choice took effect. In contrast, an ex-ante assessment tries to shed some light onto the future and data is not available. Modelling and modelling tools can be helpful by providing a simplified representation of reality and simulating potential contrasting pathways into the future.

Ex-ante assessments through models and modelling tools could provide valuable insights on potential choices affecting complex societal and environmental problems (e.g. climate change, achievement of the Millennium Development Goals (UN, 2005), securing ecosystem services (Carpenter et al., 2009)). A prominent example of the use of models and modelling tools is the assessment of the likely impacts of climate change on the biophysical environment and society (IPCC, 2007) by the Intergovernmental Panel on Climate Change. An example on a lower spatial scale is the FARMSCAPE project (Carberry et al., 2002), in which farmers, advisory services and researchers jointly applied a simulation tool to assess the potential for

alternative management strategies of cropping systems on Australian farms. Such ex-ante assessments need to involve multiple disciplines and cover multiple scales.
There is an increasing interest in multi-disciplinary and multi-scale research, which is reflected in institutional regulations and calls for funding. For example, the European Commission now requires an impact assessment (EC, 2005) of its policies before these are implemented. Calls for research funding more often require collaboration between scientists and research groups from different disciplines (Metzger and Zare, 1999; Bruce et al., 2004; EC, 2009). Although projects market themselves as interdisciplinary, the interdisciplinarity of the research achieved is often quite limited and remains a claim, as Bruce et al. (2004) observed.

1.1.2. Interdisciplinary research

The current division of science in disciplines in our universities has started some four hundred years ago. In 1637 Rene Descartes (1596-1650) wrote his treatise 'Discours de la méthode pour bien conduire sa raison, et chercher la verité dans les sciences' (English title: 'Discours on the method'), which advocates the necessity to break down a scientific problem into smaller parts, solve the smaller problems first and solve the whole problem by solving step by step the smaller problems. Descartes used this reductionist approach to arrive at his famous 'cogito ergo sum,' which represents the core of his metaphysic theories. Over the centuries, the breaking up into smaller parts led to the separation of science in disciplines to study problems in isolation. An advantage is that researchers are shielded from an overwhelming complexity (Bruce et al., 2004). A mono-disciplinary scientist cannot, by definition, combine his analysis of one smaller part or problem into one holistic science. By contrast, Descartes achieved a holistic metaphysics in his treatise using solutions to smaller problems to prove the existence of God.
The role of interdisciplinary research is to connect the answers to the smaller problems to answer the over-arching societal problem and to identify emerging properties (De Ridder, 1997) that do not appear in mono-disciplinary research (e.g. the whole of the parts is different from the sum of the parts). In this thesis, we define

interdisciplinary research as research in which the researchers purposefully cross the disciplinary borders and jointly use methodologies that cover a range of disciplines to create new knowledge and achieve a common research goal (Tress et al., 2007). Transdisciplinary research is defined as interdisciplinary research, that includes stakeholder involvement through participatory approaches (Tress et al., 2007). This thesis limits itself to interdisciplinary research, although it was part of a larger project which had a participatory approach to stakeholder involvement (Van Ittersum et al., 2008).

Many institutional barriers exist to successful interdisciplinary science in funding agencies (Metzger and Zare, 1999), editorial teams of journals (Tress et al., 2006), universities (McEvoy, 1972) and research project organization (Tress et al., 2007). Also, the researcher himself may act as a barrier to interdisciplinary research. Interdisciplinary research requires other types of researchers and/or organization of research projects than mono-disciplinary research as researchers need to work and invest in teams (Bruce et al., 2004; Tress et al., 2007), be able to transcend the comfort zone of their own discipline (Norgaard, 1992), and have adequate knowledge across disciplines (Harris, 2002). Personality traits important to interdisciplinary researchers are creativity, curiosity, open-mindedness to other disciplines, good team worker, flexibility, good listening skills and a fast learner (Bruce et al., 2004).

1.1.3. Integrated Assessment and Modelling

Integrated Assessment and Modelling (IAM; Harris, 2002; Parker et al., 2002) is a method that supports ex-ante assessment through modelling and modelling tools. IAM is a type of integrated assessment, which 'is an interdisciplinary and participatory process of combining, interpreting and communicating knowledge from diverse scientific disciplines to allow a better understanding of complex phenomena' (Rotmans and Asselt, 1996). Usually IAM combines several quantitative models representing different systems and scales into a framework for Integrated Assessment (Parker et al., 2002). In this thesis, a model is defined as a deliberate simplification of reality that represents part of reality as a quantitative system. IAM is a quantitative, future-oriented, interdisciplinary and participatory methodology, that aims to supply tools to support the process of integrated assessment. Different types of IAM exist, e.g. meta-modelling, Bayesian networks, agent-based systems and linking of comprehensive models into model chains. This thesis focuses on the latter IAM approach for assessing changes in agriculture and agricultural land use (Verburg and Lesschen, 2006) due to policy changes, technological innovations and biophysical or societal trends.

Integrated assessment and IAM usually involve multiple disciplines, multiple scales and multiple dimensions and integration is an essential and challenging task (Rotmans and Asselt, 1996; Brandmeyer and Karimi, 2000; Harris, 2002; Parker et al., 2002; Hinkel, 2008). In this thesis, integration is defined as a communication process of combining parts (e.g. scales, disciplines, scientific methods) into a whole (e.g. a procedure or model chain). As Parker et al. (2002), Harris (2002), Tress et al. (2007) and Hinkel (2008) noted, different types of integration are relevant for integrated assessment, IAM and land-use modelling:

1. methodological integration of models, methods and process descriptions across scales;
2. semantic integration of knowledge, data and meaning, e.g. speaking the same language;
3. technical integration of programming paradigms, data and models into modelling frameworks and graphical user interfaces;
4. social integration within a research team, with stakeholders and across cultures;
5. institutional integration in the existing disciplinary university systems, funding schemes and merit systems.

Integration appears as a multi-headed Hydra snake for IAM-projects (Fig. 1.1), due to the many different types of integration that have to be achieved in parallel and due to the important role of communication. If a research project can manage all but one head of the Hydra, it is bound to fail in its interdisciplinarity. If it manages all the heads of the Hydra, it stands a better change of succeeding. This thesis focuses on methodological, semantic and technical integration.

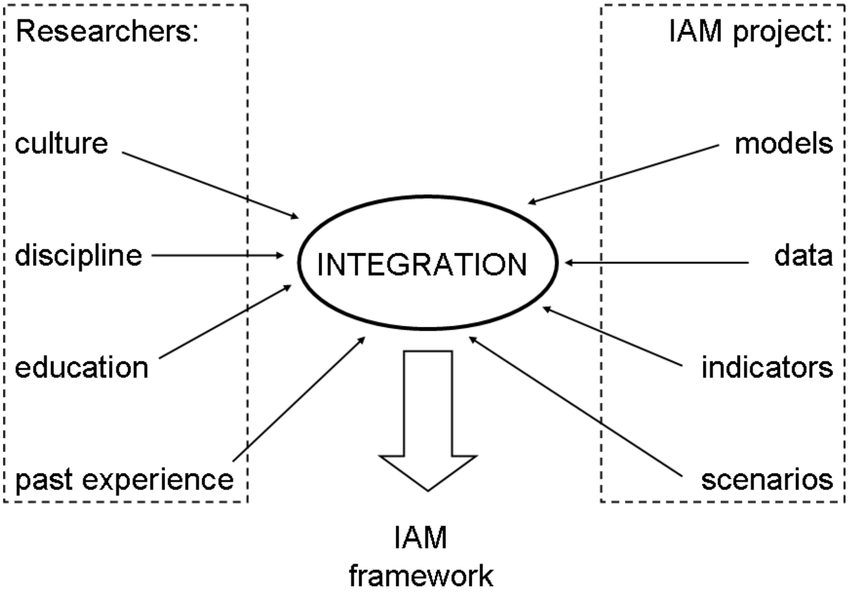


Figure 1.1. The integration challenge in an IAM project.

1.1.4. Bio-economic farm models

The thesis focuses on integration required for one type of model, bio-economic farm models (BEFMs). BEFMs typically combine methods and data from biophysical and economic disciplines. A BEFM links resource management decisions of farms to current and alternative production possibilities describing input-output relationships and associated externalities. Bio-economic farm models have been proposed for ex-ante assessments (Flichman and Jacquet, 2003) and many recent applications (Donaldson et al., 1995; Flichman, 1996; Judez et al., 2001; Berentsen, 2003; Veysset et al., 2005; Riesgo and Gomez-Limon, 2006; Onate et al., 2007; Semaan et al., 2007) have been published to assess the impacts of policy changes on economic, environmental and social indicators of agricultural systems.

If a bio-economic farm model is to be used for ex-ante assessments of agricultural and environmental policies and technology changes, it has to fulfill several requirements. First, results (particularly product supply) must enable upscaling to higher hierarchical levels (e.g. country or market). Second, data with respect to farm types, their locations and their agricultural activities must be available throughout regions. Third, the model must be applicable to different farm types including mixed farm types. Fourth, the application and calibration to a farm type or region requires only few specific steps or ad hoc constraints. Finally, many different policy instruments and potential technology changes are implemented in a generic way.

1.2. Problem definition

1.2.1. State-of-the-art in integration

One important question for an interdisciplinary research project is 'how' or 'what-to-do' to achieve a methodological, semantic and technical integration. Harris (2002) most aptly formulated this as:

'So just how do we integrate across disciplines and synthesise knowledge so as to produce useful outcomes? How do we do this in an environment where data sources have different types and degrees of error, where some data types from disparate disciplines are even incompatible? How do we keep the community on side and committed to change — and at the same time convince our political and economic masters to keep funding the whole enterprise? ….This is not rocket science, it is much more difficult!'

Already ten years before Harris (2002), Norgaard (1992) identified the same problem in relation to sustainability science for agriculture: 'Discipline boundaries have impeded true implementation of interdisciplinary methodologies and the development of generalized models because the assumptions, cultures, and paradigms within the disciplines have not been overcome.' Semantic integration, e.g. the synthesis of knowledge and the achievement of a common language between researchers, is often lacking (Jakobsen and McLaughlin, 2004; Bracken and Oughton, 2006; Scholten et al., 2007; Tress et al., 2007; Hinkel, 2008). Literature on semantic integration presents 'largely anecdotal or non-empirical discussions (Jakobsen and McLaughlin, 2004).' Jakobsen and McLaughlin (2004) investigated communication between multi-disciplinary teams (e.g. landscape ecologist, economists, social scientists), that jointly had to manage a large ecosystem. They found that team members did not have sufficient understanding of the central domains and the problem at hand and stressed the importance of communication to achieve those. Hinkel (2008) developed a method for transdisciplinary knowledge integration, which consists of building a shared language through meta-concepts as a first step and integration of methodologies and theories by representing them in the same mathematical language, and using a common mathematical integration method as a second step. Hinkel (2008) studied these two steps in several case studies, but does not explicitly discuss the role of communication as do Jakobsen and McLaughlin (2004). He only mentions that mathematical formalisms have to be carefully explained to non-mathematical scientists. Striving for communication-intensive integration might not always fit in the research strategies of scientists, who may feel most comfortable as 'lone boffin doing as (s)he pleased' (Harris, 2002) or in their ivory tower.

1.2.2. Integration in SEAMLESS

Integration in an IAM project requires communication (Jakobsen and McLaughlin, 2004) and a rigorous formal approach to achieve a shared understanding (Hinkel, 2008). This thesis describes the methodological, semantic and technical integration within the European Sixth Framework research project 'System for Environmental and Agricultural Modelling; Linking European Science and Society' (SEAMLESS) (Van Ittersum et al., 2008). SEAMLESS aimed to overcome fragmentation in efforts of modelling agricultural systems and to achieve model integration. It developed a computerized and integrated framework (SEAMLESS-IF) to assess the impacts on environmental, social and economic sustainability of a wide range of policies and technological improvements across a number of scales. With respect to the models (Fig. 1.2), macro-level economic partial or general equilibrium models (Heckelei and Britz, 2001) are linked to a micro-level bio-economic farm model (Louhichi et al., 2009) and a cropping system model (Donatelli et al., 2009), using micro-macro upscaling methods (Pérez Domínguez et al., 2009). Next to models, data and indicators are derived from different dimensions (e.g. economic, biophysical, climatic, policy), different scales (e.g. field, farm, regional, national, continental) and provided by different institutions. SEAMLESS-IF and its components (e.g. models, databases and indicators) are designed to be generic and reusable and applicable to a range of policy questions, technology changes and trend changes.

Challenges in methodological integration in SEAMLESS were to ensure a meaningful exchange of data between the models, consistency and integrity of data and alignment of modelling methodologies across models, scales and disciplines (Fig. 1.2). The semantic integration aimed to develop a shared understanding and language concerning models, indicators, data and scenarios (e.g. policy questions, technology changes or biophysical and societal trends) for the circa 100 participating scientists from 30 different institutions. Finally, technical integration entailed the development of a computer program that reflects the joint knowledge, an intuitive and easily understandable design to the Graphical User Interface that is not overly complex with disciplinary jargon, and a modelling framework that enables the execution of models in model chains.

In SEAMLESS, BEFMs were chosen as one model in the model chain representing the farm level and farm responses. A BEFM was developed for different regions, different farm types and different applications with two main purposes. These purposes are to provide the possibility to upscale supply responses from farm to market scale and to enable detailed regional integrated assessment, throughout regions and farm types of the European Union for a wide range of agricultural and environmental policies.

The SEAMLESS project consists of work packages and tasks within the work packages. Different organizational structures and communication strategies were used in SEAMLESS for methodological, semantic and technical integration. Methodological integration was achieved by organizing a group of seniors scientists from different disciplines and institutes in one work package. For semantic integration, a cross-work package task force was established, that consisted of two knowledge engineers and a large group of domain experts. From an organizational point of view, the technical integration was jointly achieved by the work packages of the modellers and computer scientists.

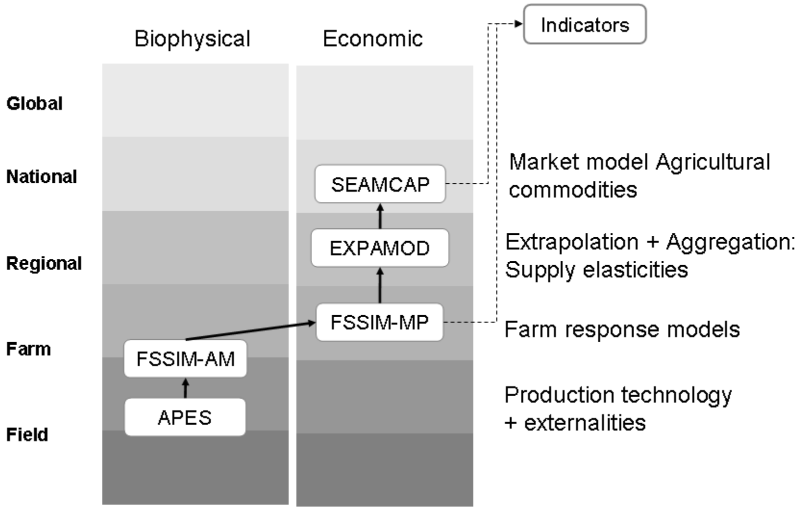


Figure 1.2. The model chain in SEAMLESS

1.3. Objectives

The objective of this thesis is to develop integration approaches for interdisciplinary model-based research:

1. by developing methodological, semantic and technical modes of integration;
2. by developing communication processes required to achieve one joint understanding of the research in a group of researchers from different disciplines;
3. by explicitly generating meaningful and coherent knowledge-level specifications across models, disciplines and scales through the use of the modes of integration and the communication processes.

We consider methodological, semantic and technical integration in two case studies:

1. integration within one type of model (i.e. a bio-economic farm models) used for assessment of impact indicators covering two hierarchical systems levels (i.e.

field to farm) involving a relatively small team of ca. ten agronomic and farm economic scientists, referred to in this thesis as “bio-economic integration”;

2. integration covering five hierarchical systems levels (i.e. field, farm, region, country and continental) linking different types of models (Fig. 1.2) for the calculation of impacts on European agricultural systems through indicators. This case study involves a large team of about hundred scientists from agronomy, economics, landscape ecology, information technology and environmental sciences, referred to in this thesis as “multi-issue integration.”

1.4. Research questions

The research questions are defined on the basis of the two case studies, bio-economic integration and multi-issue integration. For bio-economic integration, first the strengths and weaknesses of current Bio-Economic Farm Models (BEFMs) will be investigated with the research question: “What is state-of-the-art in BEFMs?” Second, an important part of any BEFM is the definition of agricultural activities, which requires linking many different data sources and linkage to cropping system models to evaluate these activities. The research question: “What is a suitable generic method to define alternative agricultural activities for the future, usable in different scientific methods?” captures the integration of methods to define alternative activities in Chapter 3. Third, the development of a generic and widely applicable BEFM is targeted with the research question: “What is a suitable methodological and technical design of a BEFM, that is applicable to many farm, soil and climate types and to different purposes at different levels of detail with links to other models?” The second and third research question both target the development of generic methods. Generic methods reinforce the importance of deliberate and transparent integration efforts, because these methods must be widely re-usable, widely accepted by the research and stakeholder community and easily adaptable to new circumstances.

In multi-issue integration, the methodological, semantic and technical integration of a set of models will be investigated with respect to their data sources, model linking and scenario definition. First, the integration of data sources is considered with the research question: "What is an appropriate integration and data structure to manage the multiple data-sources required by the models linked in a model chain?" Second, the model linking is further investigated with the research question: "What is a suitable integration of large complex models from different disciplines based on different modelling techniques operating on different time and spatial scales?" The last research question targets the efforts required to arrive at a shared understanding of scenarios across models, data sources and indicators from different disciplines. The research question is formulated as: "What is a conceptualization of a scenario and assessment project for use in multiple models, on multiple scales, in multiple dimensions and in the software implementation of the graphical user interface and database?"

1.5. Methods

1.5.1. Methodological integration

The methodological integration focuses on aligning different scientific methodologies, aligning spatial and temporal levels and identifying required model improvements and extensions. It is a vital first step to facilitate communication between modellers, non-modelling researchers and stakeholders (Liu et al., 2008). Good practice guidelines (Refsgaard and Henriksen, 2004; Jakeman et al., 2006; Scholten, 2008) exist for methodological development of a model in all steps of model building for a mono-disciplinary model. Methodological challenges for land use models are to model appropriate '(1) level of analysis; (2) cross-scale dynamics; (3) driving forces; (4) spatial interaction and neighbourhood effects; (5) temporal dynamics; and (6) level of integration (Verburg et al., 2004).'

The models referred to in this thesis are a cropping system model, a bio-economic farm model, an econometric estimation model, and a partial equilibrium optimization model. Each of these have a different spatial and temporal scale and are based on different modelling techniques (Fig 1.2.). The cropping system models (Van Ittersum and Donatelli, 2003) operates at the field systems level, and represents one hectare (or a point). It is a dynamic simulation and it usually simulates a period of 10 to 25 years with a daily time step. Cropping system models typically have deterministic mechanistic process descriptions, which means that processes are implemented according to established scientific theories. The bio-economic farm model and partial equilibrium optimization model are both optimization models based on mathematical programming. These models are built based on assumptions with respect to the functioning of economic agents, e.g. farms or market parties. These models are comparative static, e.g. they have no interdependence of outcomes across years, and model results represent the equilibrium situation for a year. Bio-economic farm models operate at a farm systems level, either a real farm or a representative or average farm for a group of farms. The partial equilibrium optimization model in our model linking operates at the continental systems level, covering several member states. Finally, the econometric estimation model is based on statistical procedures and the model functional form depends on the data quality and availability. The model uses a sample of farm types across regions as a basis for estimation. The estimation model uses annual data combined with outcomes from the bio-economic farm model.

Methodological integration in a model chain requires that the data produced by one model are a valid input to another model. Temporal and spatial scaling of model outputs is crucial to move between systems levels and temporal scales, e.g. from field to farm or from daily to yearly time steps. Different modelling techniques (e.g. optimization, estimation, simulation) affect the interpretation of outputs, as different modelling techniques have different levels of uncertainty and model different outputs. For example, if the cropping system model is used in an ex-ante assessment to simulate yield of crops, it is usually run for a simulation period of 25 years. The BEFM expects the cropping system model to provide an estimate for the yield of wheat for the year 2003 in a region. Two options exist for the cropping system model to provide this output. Either, it can provide only its estimate of 6 tonnes wheat yield based on weather data of the year 2003, or it provides the average wheat yield of 6.4 tonnes over the whole simulation period of 25 years. Another example is the over-estimation of farm income in bio-economic farm models, because fixed costs are difficult to incorporate and transaction costs to adopt different activities are ignored.

1.5.2. Semantic integration

Semantic integration entails the development of a common language to achieve a shared understanding between modellers and their models. This is a crucial challenge for any integrated modelling project (Jakobsen and McLaughlin, 2004; Bracken and Oughton, 2006; Tress et al., 2007; Hinkel, 2008; Scholten, 2008), as it is provides the building blocks, consistency and transparency in definitions required for the methodological and technical integration. Models from different disciplines have a different representation of data, space and time, and linking them implies that the inputs of one model have to be matched to the outputs of another model, while the modellers and their models should have a common understanding of the space and time in which they operate.

Very few practical applications of possible methods for semantic integration of models could be found in literature, with the exceptions of Hinkel (2008) and Scholten (2008). Possible methods are variable mapping, mathematical formalism, concept maps and ontologies. Variable mapping is an ad hoc process of investigating which variables could be exchanged between models and then mapping the variables to each other. As variable mapping is not formalized and ad hoc, it remains a black box approach. Hinkel (2008) uses mathematical formalism as a methodology to align firstly terminology between models and secondly the model equations across models. Hinkel (2008) uses this for a semantic integration to link models in a number of modelling projects. One disadvantage of mathematical formalism is, as Hinkel (2008) mentioned, that non-modellers need explanation and training in order to be involved. Concept maps (Novak and Cañas, 2006) are graphs representing knowledge, in which concepts are in circles and relationships are shown by lines connecting two concepts. Finally, ontologies consist, like concept maps, of concepts and relationships between concepts (Antoniou and Van Harmelen, 2004). Ontologies have the advantage that these are expressed in man readable and machine understandable format. This thesis focuses on the use of ontologies for semantic integration and model linking, since i. ontologies are in machine understandble format, i.e. as the Web Ontology Language (McGuinness and Van Harmelen, 2004); ii. ontologies are based on first order logic upon which a computer can reason; iii. the developed ontologies are a separate product, that are independent of the models on which they are originally based and that can be used in developing new models and iv. both modellers and non-modellers can contribute to the ontology development.

1.5.3. Technical integration

Technical integration means designing an adequate computer-based framework to integrate and exchange data across different models, which often are based on diverse programming paradigms. Technical integration of linking models in model chains is classified by Brandmeyer and Karimi (2000) in five hierarchical levels: i. one way data transfer, in which output files of one model are used as input files to the next model, ii. loose coupling, in which two models automatically send each other data, iii. shared coupling, in which two models are executed through a common graphical user interface and common data storage, iv. joined coupling, in which a shared graphical user interface and data storage is used and in which one model embeds other model(s) and v. tool coupling, in which models are linked together in a modelling framework with a common graphical user interface and data storage.

SEAMLESS opted for the last option, i.e. a tool coupling. A tool coupling requires first a central repository for data storage for all models, scenarios and data sources. Second, tool coupling requires one graphical user interface from which all models can be parameterized and executed. Third, tool coupling requires a modelling framework, that supports the execution of models in a model chain (Hillyer et al., 2003; Rahman et al., 2003; Moore and Tindall, 2005). To achieve the three requirements of a central data repository, a model-independent graphical user interface and a modelling framework, an advanced software architecture is needed that bridges different programming paradigms used in the models, data sources and graphical user interface. The software architecture for the SEAMLESS project makes extensive use of the ontology achieved in the semantic integration to provide a shared conceptual schema across different programming paradigms (Wien et al., 2009) and to generate part of the source code based on the ontology (Athanasiadis et al., 2007b). As a modelling framework, the Open Modelling Inteface (OpenMI - Moore and Tindall, 2005) standard was adopted and extended. This thesis will focus on aspects of the software architecture of SEAMLESS-IF that are relevant to integration problems, for example preparing models for execution in model chains with OpenMI and storing data from different data sources in one relational database.

1.6. Reading guide

The thesis is organized along the two case studies bio-economic integration and multi-issue integration. Chapters 2, 3 and 4 answer the research questions for Bio-economic integration. Chapter 2 entitled "Assessing farm innovations and responses to policies: A review of bio-economic farm models" answers the research question: "What is state-of-the-art in BEFMs?" It concludes with a research agenda for BEFMs and some items of the research agenda are elaborated in Chapters 3 and 4. Chapter 3 develops a generic and comprehensive conceptualization of alternative agricultural activities for bio economic farm models and focuses on the research question: "What is a suitable generic method to define alternative agricultural activities for the future, usable in different scientific methods?" A generic method is also targeted in Chapter 4, where a generic BEFM is discussed and evaluated on criteria for a generic model. This chapter tackles the research question "What is a suitable methodological and technical design of a BEFM, that is applicable to many farm, soil and climate types and to different purposes at different levels of detail with links to other models?”

Chapters 5, 6 and 7 describe different aspects of multi-issue integration. Chapter 5 focuses on a database for integrated assessment of European agricultural systems. This Chapter attempts to answer the research question “What is an appropriate integration and data structure to manage the multiple data-sources required by the models linked in a model chain?” A crucial challenge in Integrated Assessment and Modelling (IAM) is elaborated in Chapter 6. This challenge concerns the integration of models in a model chain and relates to the research question: “What is a suitable integration of large complex models from different disciplines and modelling techniques operating on different time and spatial scales?” The final research question “What is a conceptualization of scenario and assessment project for use in multiple models, on multiple scales, in multiple dimensions and in the software implementation of the graphical user interface and database?” is answered in Chapter 7. Chapter 7 describes both the process to achieve a shared conceptualization and the shared conceptualization achieved of scenario and assessment projects.

Chapter 8 discusses methodological, semantic and technical integration in depth. Both case studies bio-economic integration and multi-issue integration will be more closely examined for lessons learned in relationship to integrated modelling, interdisciplinary research, IAM and BEFMs.

Chapter 2. Assessing farm innovations and responses to policies: a review of bio-economic farm models

Abstract

Bio-economic farm models (BEFMs) are developed to enable assessment of policy changes and technological innovations, for specific categories of farming systems. A rapidly growing number of research projects is using these models and there is increasing interest for application. The chapter critically reviews past publications and applications of BEFMs on their strengths and weaknesses in assessing technological innovation and policy changes for farmers and policy makers and highlights key issues that require more attention in the use and methodology of BEFMs. A Bio-Economic Farm Model (BEFM) is defined as a model that links formulations describing farmers' resource management decisions, to formulations that represent current and alternative production possibilities in terms of required inputs to achieve certain outputs, both yield and environmental effects. Mechanistic BEFMs are based on available theory and knowledge of farm processes and these were the focus of our study. Forty-eight applications of mechanistic BEFMs were reviewed as to their incorporation of farmer decision making and agricultural activities, comprehensiveness, model evaluation, and transferability. A clear description of end-use of the BEFM, agricultural activities, model equations and model evaluation are identified as good practices and a research agenda is proposed including the following issues: 1. development of a thorough and consistent procedure for model evaluation; 2. better understanding and modelling of farmer decision making and possible effects of the social milieu; 3. inclusion of several economic and environmental aspects of farming including multifunctionality and 4. development of a generic, modular and easily transferable BEFM.

Sander Janssen and Martin K. van Ittersum, 2007. Assessing farm innovations and responses to policies: A review of bio-economic farm models. Agricultural Systems 94 (3), 622-636.

2.1. Introduction

Policy makers and farmers have an interest in making ex-ante assessments of the outcomes of their choices in terms of policy and farm plan, (cf. Rossing et al. (1997); Zander and Kächele (1999); Leeuwis (1999); EC (2005)). This interest mainly concerns the assessment of socio-economic and environmental performance of farms as a result of innovations, and the assessment of socio-economic and environmental effects of policies on the major categories of farms. Mathematical models based on systems analysis are suited to explore and assess uncertain future states of systems. As expressed by Edwards Jones and Mc Gregor (1994) "the utility of a series of whole farm models for the European situation would be substantial, particularly in the ex-ante policy assessment and marketing of on-farm technology." Certainly, not only the European situation would benefit from assessments of agricultural innovations or agricultural and environmental policies.

For such assessments research has proposed the use of methods such as Bio-Economic Farm Models (BEFMs), Multi-Agent Systems, Environmental Risk Mapping, Life Cycle Analysis, Environmental Impact Assessment and Agri-Environmental Indicators, which are each briefly reviewed in Payraudeau and Van der Werf (2005). A Bio-Economic Farm Model (BEFM) is defined as a model that links formulations describing farmers' resource management decisions to formulations that describe current and alternative production possibilities in terms of required inputs to achieve certain outputs and associated externalities. The focus of this article is on BEFMs as they have some clear advantages with respect to the other methods reviewed by Payraudeau and Van der Werf (2005): (i) they are based on an constrained optimization procedure and thereby seem to match the reality of small farmers, striving, with limited resources, to improve their lot (Anderson et al., 1985); (ii) many activities, restrictions and new production techniques with sound technical specifications can be considered simultaneously (Wossink et al., 1992; Ten Berge et al., 2000; Weersink et al., 2004), including linkages between crop and livestock production (Antle and Capalbo, 2001); (iii) the effects of changing parameters, for example prices, can easily be assessed through sensitivity analysis (Wossink et al., 1992), and (iv) they can be used both for short term predictions and long term explorations (Van Ittersum et al., 1998). A BEFM permits the (ex-ante) assessment of technological innovations and policies over a range of different geographic and climatic circumstances. A rapidly growing number of research projects is using these models and there is increasing interest for application (Deybe and Flichman (1991); Donaldson et al. (1995); Rossing et al. (1997); Louhichi et al. (1999); Vatn et al., (2003); Gibbons et al. (2005); and Torkamani (2005)).

The presently available publications and applications of BEFMs can be subdivided in three broad classes based on their purpose: i. exploring the suitability of alternative farm configurations and technological innovations, i.e., assessing whether a technology will be viable financially and will have positive environmental effects, for example Abadi Ghadim (2000), usually focused at (groups of) farmers and extensionist; ii. predicting or forecasting the effects of changing policies on agriculture, focusing at policymakers or facilitating discussion between multiple groups of stakeholders, for example, Berentsen and Giesen (1994) and Bartolini et al. (2006), and iii. efforts to highlight methodological aspects of BEFMs and their improvement; for example Apland (1993), usually targeted at researchers.

Currently many descriptions and applications of BEFMs are being published (cf. Bartolini et al. (2007), Acs et al. , Onate et al. (2007) and Semaan et al. (2007)). A critical analysis of the methodological strengths and shortcomings of these BEFMs and their applications, as related to ex-ante assessment of farm innovation and policies for farmers, policy makers and other stakeholders is lacking. From such analysis, an overarching research agenda can be derived to help and guide efforts on the third class of purposes mentioned above, i.e., methodological improvement of BEFMs.

The objectives of this article are to critically review past publications and applications of BEFMs as to their strengths and weaknesses in assessing technological innovation and policy changes for farmers and policy makers and to highlight key issues that require more attention in the use of BEFMs. As a result, this article tries to draw up a research agenda and to identify good practices in the use of BEFMs. An in depth analysis of 48 model studies was carried out (see Table 2.1), which was supplemented with information from text books and methodological articles. These 48 model studies used 42 different models, as sometimes a model was used in subsequent studies. The review and examples focus on agriculture in industrialized countries, though many aspects will be equally valid for agriculture in developing countries.

In the next Section a classification of BEFMs and their use will be presented. In the subsequent Sections 2.3, 2.4 and 2.5 ‘Farmer decision making’, ‘Agricultural activities’ and ‘Comprehensiveness’ of BEFMs are discussed. We then analyse the quality of BEFMs. Finally, conclusions in the form of good practices and a research agenda are presented.

2.2. Methodology and use of BEFMs

2.2.1. A classification

For this article the term bio-economic farm model (BEFM) is proposed, but literature uses a wide range of terms for the same type of models. Publications use terms such as ‘bio-economic’, ‘ecological-economic’ or ‘combining the environmental and economic,’ referring to the integration of economic and biophysical processes and models.

The distinction between on the one hand empirical and mechanistic BEFMs and on the other hand normative and positive approaches is proposed here to classify BEFMs. These distinctions between empirical versus mechanistic and normative versus positive are sometimes mentioned in publications (cf. Flichman and Jacquet (2003), Calker et al. (2004) and Thornton and Herrero (2001)), but poorly defined.

Table 2.1 Model studies included in this research

Reference	Farm type(s)	Country	End use[2]	model type: name
Abadi Ghadim (2000)	arable	Australia	1	LP: MIDAS
Acs et al. (2007)	arable	Netherlands	4	Dynamic LP
Annetts and Audsley (2002)	arable	United Kingdom	2	MCDM
Apland (1993)	arable	USA	1	LP (DDP and DSP[3])
Barbier and Bergeron (1999)	mixed	Honduras	2	dynamic recursive LP
Bartolini et al. (2007)	arable	Italy	2	MCDM: MAUT[5]
Benoit and Veysset (2003)	livestock	France	1	Static LP: Opt'INRA
Berentsen (2003)[1]	dairy	Netherlands	4	Static LP
Berentsen and Giesen (1994)[1]	dairy	Netherlands	2	Static LP
Berentsen et al. (1998)[1]	dairy	Netherlands	1	Static LP
Berger (2001)	several	Chile	4	LP coupled to MAS
Berntsen et al. (2003)	mixed	Denmark	2	Static LP :FASSET
Beukes et al. (2002)	livestock	South Africa	1	Dynamic LP
Bos (2000)	livestock	Netherlands	1	MGLP
Cain et al. (2007)	Mixed and dairy	Pakistan	1	Normative LP
Calker et al. (2004)	dairy	Netherlands	2	Static LP
de Buck et al. (1999)	arable	Netherlands	3	Normative LP
Deybe and Flichman (1991)	arable	Argentina	2	LP
Dogliotti et al. (2005)[1]	vegetable	Uruguay	1	MILP[4]: Farm Images
Dogliotti et al. (2003)[1]	arable	Netherlands	3	LP
Donaldson et al. (1995)	arable	England and France	2	dynamic recursive LP
Dorward (1999)	subsistence	Malawi	3	LGP with DSP and SSP[3]

Table 2.1 (Cont.) Model studies included in this research

Reference	Farm type(s)	Country	End use[2]	model type: name
Falconer and Hodge (2000)[1]	arable	United Kingdom	2	Normative LP
Falconer and Hodge (2001)[1]	arable	United Kingdom	2	Normative LP
Gibbons et al. (2005)	arable	United Kingdom	4	LP: farm-adapt
Gutierrez-Aleman et al. (1986a) and Gutierrez-Aleman et al. (1986b)	mixed	Brazil	1	LP
Jannot and Cairol (1994)	arable	France	1	Normative LP
Louhichi et al. (1999)	mixed	Tunisia	2	dynamic non linear model
Meyer-Aurich (2005)	mixed	Germany	2	MCDM LP: MODAM
Morrison et al. (1986) [1]	Mixed	Australia	1	Normative LP: MIDAS
Oglethorpe (1995)[1]	livestock	England	2	Static LP (MOTAD)
Oglethorpe and Sanderson (1999) [1]	livestock	Scotland	2	Static LP (MOTAD)
Oñate et al. (2007)	arable	Spain	2	PMP
Pacini (2003)	mixed	Italy	2	LP
Pfister et al (2005)	subsistence	Nicaragua	1	Dynamic mathematical programming model
Ramsden et al. (1999)	dairy	United Kingdom	2	Static LP
Riesgo and Gómez-Limón (2006)	arable	Spain	2	MCDM: MAUT[5]
Schilizzi and Boulier (1997)	mixed	Mexico	3	MCDM
Semaan et al. (2007)	arable	Italy	2	LP with risk
Ten Berge et al. (2000)	several	Netherlands	2	MGLP[5]
Thompson (1982)	mixed	New Zealand	2	LP
Kruseman and Bade (1998)	mixed	Mali	2	MGLP[5]

Table 2.1 (Cont.) Model studies included in this research

Reference	Farm type(s)	Country	End use[2]	model type: name
Vatn et al. (1997) and Vatn et al. (2003)	arable	Norway	3	Dynamic LP: ECECMOD
Wallace and Moss (2002)	livestock	Ireland	3	Dynamic Recursive LP and WGP
White et al. (2005)	arable	Peru	3	LP
Wossink et al. (1992)[1]	arable	Netherlands	4	Static LP
Wossink et al. (2001)[1]	arable	Netherlands	4	Static LP: MIMOSA.
Zander and Kächele (1999)	several	Germany	4	MCDM LP: MODAM

[1] = If authors wrote more than one article based on the same model, both articles were included as the research question was often different, which lead to different model structures and analysis.; [2] = different end uses are assisting farmer decision making (=1), policy assessment (=2), developing methodologies (=3), both assisting farmer decision making and policy assessment (=4); [3] = DDP and DSP refer to Discrete Deterministic Programming and Discrete Stochastic Programming, where DSP has several time periods in one year and DDP has not. SSP is semi sequential programming and a form of DSP; [4] = MILP is Mixed Integer Linear Programming.; 5 = see Table 2.2

Hereby we propose a set of definitions. Mechanistic BEFMs are built on a certain image the researcher has of the processes on farms occurring in reality (Pandey and Hardaker, 1995); in other words a mechanistic model is built on existing theory and knowledge (Austin et al., 1998). Mechanistic models are suitable both for extrapolations and long-term predictions, as these models can simulate system "behaviour outside the range of observed data in ways consistent with established scientific understanding" (Antle and Capalbo, 2001). Empirical models are constructed from the data that are incorporated in them, and try to find relationships in the observed data that are not known ex-ante (Austin et al., 1998). In empirical models prediction of future changes is mostly based on an extrapolation of historical time-series of observed past behaviour and a description of past agricultural technologies. Therefore, they cannot easily deal with specific alternative technological options or new constraints and polices (Ruben et al., 1998; Falconer and Hodge, 2000; Flichman and Jacquet, 2003). This chapter will further focus on mechanistic BEFMs.

BEFMs can be used according to a positive and normative approach. Positive approaches try to model the actual behaviour of the farmer by describing farm responses and trying to understand them, while normative approaches try to find the optimal solutions and alternatives to the problem of resource management and allocation (Flichman and Jacquet, 2003). BEFMs based on a normative approach are setting a 'norm.' The 'norm' describes what farmers ought to do in order to achieve a certain objective, for example optimise profits (Berntsen et al., 2003). Farmers often do not succeed or do not desire to manage the farm according to model outcomes (the norm) due to various reasons, such as imperfect information, risk aversion, management quality and skills (Wossink and Renkema, 1994; Falconer and Hodge, 2000; Calker et al., 2004).

Mechanistic farm models generally use mathematical programming or optimization models, which are often based on Linear Programming (LP), see Table 2.1. Linear programming represents the farm as a linear combination of so called 'activities'. An activity is a coherent set of operations with corresponding inputs and outputs, resulting in e.g. the delivery of a marketable product, the restoration of soil fertility, or the production of feedstuffs for on-farm use (Ten Berge et al., 2000). An activity is characterised by a set of coefficients (technical coefficients (TCs) or input-output coefficients) that express the activity's contribution to the realisation of defined goals or objectives in modelling terms (Ten Berge et al., 2000). As inputs are limited resources, constraints to the activities are defined, which represent the minimum or maximum amount of a certain input or resource that can be used. This system of activities and constraints is then optimised for some objective function, reflecting a user-specified goal, for example profit. Standard mathematical formulations of different types of LP models can be found in Hazell and Norton (1986).

2.2.2. Major types of application

Often mechanistic BEFMs are used in a normative approach, for example Wossink et al., (1992); Ten Berge et al. (2000); Berntsen et al., (2003); Berentsen (2003) and Pacini (2003). Normative mechanistic approaches may be used in assessments of alternative farm configurations and technological innovations targeted at farmers and explorations of the long term effects of policies and technological innovations targeted at policy makers or groups of stakeholders. However, the predictive power of such models is restricted and hence their usefulness in policy assessment.

To assess technological innovations to their economic viability and environmental effects static BEFMs focusing on one or more technologies with exogenous input/output prices are often constructed, for example Abadi Ghadim (2000) and Benoit and Veysset (2003). A problem with mechanistic BEFMs is that when a technological innovation becomes available to the model, it is instantaneously used as it is a better option than existing technologies (Wossink and Renkema, 1994). This problem of instantaneous adjustment goes well with the aims of a normative approach, like demonstrating farmers promising alternative set-ups, but less so with a positive approach aiming at predicting actual responses. To solve this problem of instantaneous adjustment, the process of diffusion of an innovation should be part of a positive mechanistic BEFM. To incorporate this, two aspects must be considered: on the one hand the nature of the innovation itself and on the other hand, the attitude of the farmers (Wossink and Renkema, 1994) (see Section 2.5.1).

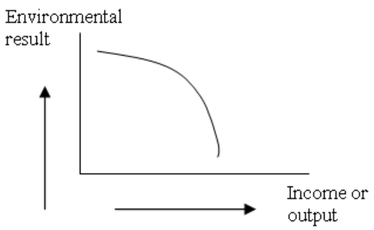
A shortcoming identified by McCown (2001) in the use of mechanistic BEFMs in advising farmers, is that a gap exists between the normative economically and technologically efficient advice given to farmers and the situation on the farm, in which the farmer finds himself. McCown (2001) proposes participatory approaches based on dialogue between farmers and researchers instead of design approaches to bridge this gap as, for example, done by Schilizzi and Boulier (1997).

A) Indicators, Response Multipliers and Elasticities

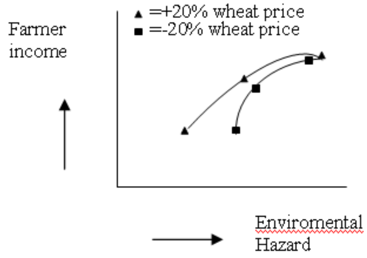
	Indicator of nitrogen leaching (kg N/ha)	Response Multiplier[1] of farmer income[2]	Price Elasticity[3]
Soft wheat	50	+ 5%	2
Potatoes	25	+10%	1

[1] = the relative change in the objective variable caused by discrete parametric changes of an input parameter (Kruseman and Bade, 1998)
[2] = Response Multiplier due to 20% change in product price
[3] = Change in production of an output or use of an input divided by the price change of the input or output causing this change in production.

B) Trade off curve (adapted from Zander and Kächele (1999))



C) Frontier Analysis (adapted from Falconer and Hodge (2001))



D) A spider diagram based on indicators

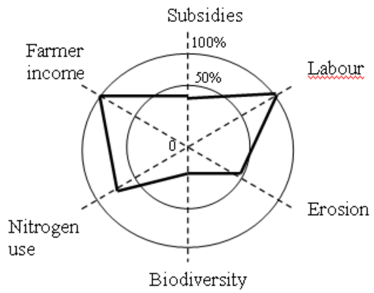


Figure 2.1. Examples of different results from BEFMs: A. Indicators, Response Multipliers and Elasticities, B. Trade off curves, C. Frontier Analysis, D. Spider diagram based on indicators.

It is possible to use mechanistic BEFMs in a positive approach (e.g. Deybe and Flichman (1991); Vatn et al. (2003)), for example through the use of Positive Mathematical Programming (PMP). Positive Mathematical Programming (Howitt, 1995) is a methodology that ensures that the model outcomes in the base run calibrate exactly on what is found in reality and that counters the tendency for over-specialization of LP models by adding quadratic cost terms to the objective function. Positive mechanistic approaches are more suitable for predictions of the effects of policy changes and technological innovations in the medium to short term. Strengths of BEFMs in policy assessment are that they have the potential to identify the possible trade-offs between economic and environmental objectives (Ruben et al., 1998) and that they include important aspects often disregarded in the policy making process, i.e., they follow a holistic approach and environmental effects at lower spatial scales (Pacini, 2003) and allow assessment of policies based on coercion (direct regulation or command and control, e.g. quotas, income support) or exchange (e.g. taxes, subsidies, cross-compliance policies, agri-environment schemes), but cannot handle polices based on persuasion, e.g. education and information (Falconer and Hodge, 2000).

Results of a BEFM can be presented in different ways, depending on the interest of policy makers, farmers or other stakeholders. Means of presentation of results (Fig. 2.1) are response multipliers (Kruseman and Bade (1998)), indicators (Pacini (2003) and Zander and Kächele (1999)), elasticities (Pannell (1997) and Falconer and Hodge (2000)), trade-off curves (Rossing et al. (1997), Zander and Kächele (1999), Ten Berge et al. (2000) and Weersink et al. (2004)), frontier analysis (Falconer and Hodge (2001)) and cost-effectiveness index of two policies (Falconer and Hodge (2001)).

2.3. Farmer decision making

2.3.1. Profit maximization versus multiple criteria approaches

The objective function of the mechanistic BEFM states which goals the farmer wants to achieve and the activities selected simulate how the farmer could achieve these goals. The end use of the mechanistic BEFM has large implications for the complexity of the objective function used. A simple formulation can be used for showing a farmer the financial or environmental effects of a change in farm technology in a normative approach, while a more complex formulation is needed for showing policy makers the possible response of farms to policy changes or a group of stakeholders with differing objectives the land use to strive for, e.g. using utility functions or multiple criteria approaches.

Farmer decision making can be classified as operational, sequential and strategic decision making, with an increasing time horizon of the decision at stake (Bouma et al., 1999). Operational decisions are the day-by-day management decisions during the growing season (Bouma et al., 1999), such as deciding whether to mow a pasture or spray a crop depending on the weather forecast. Sequential or tactical decision making relates to decisions within a growing season and to the fact that decisions on crop choice and technology are of a sequential nature. For example, a farmer may decide to use relatively more inputs on his onions during the growing season than foreseen at the start of the growing season, if he notices during the growing season that onion prices are increasing. Strategic decision making has an impact on the structure of the farm over many years, such as the choice between conventional and organic farming and investment decisions.

In the 42 different models used in 48 model studies, farmer decision making was modelled in different ways in the objective function: 23 used a simple measure of profit (income, net revenue etc) maximization, 5 a measure of profit maximization minus some risk factor (e.g. risk as avoidance of income variability), 5 an objective function that maximized expected utility (e.g. by including long term goals or measuring utility by interviewing respondents) and 9 studies used an objective function based on different objectives (multi-criteria approaches). If a farmer is assumed to be a rational profit-maximiser, his production decisions are influenced mainly by the relative prices of inputs and products (Falconer and Hodge, 2000) and the production of products, on which the farmer is assumed to have perfect knowledge (McCown, 2001). As Dent et al. (1995) state "common sense suggests that not all farmers or farm households within any given farm type are similar, and it is becoming increasingly apparent that few individuals maximize financial gain." In reality decisions of farmers are motivated by multiple, often conflicting, objectives of which profit maximisation is only one (McCown, 2001; Wallace and Moss, 2002). Personal, family and farm business objectives and attitudes are not independent and need to be considered jointly, and farmers' behaviour reflects a combination of personality factors as well as lifestyle and economic objectives

(McCown, 2001; Wallace and Moss, 2002; Bergevoet et al., 2004). Here off- or non-farm income might play a crucial role, as farmers with such extra income, do not necessarily need to make a profit and can easily pay of loans, as discussed by Wallace and Moss (2002) and Garrett (1984). In a positive approach, this diversity of farmer objectives should be considered, if the BEFMs are to approximate observed behaviour in reality.
Multi Criteria Decision Making (MCDM) methods refer to approaches that take more than one objective into account (Rehman and Romero, 1993), where the different objectives are often conflicting and derived from different dimensions, i.e., economic, environmental, biophysical and social. In this chapter, we do not consider approaches based on risk or maximization of expected utility to be MCDM approaches as these are considered adaptations of the profit maximization approach. MCDM methods are normally used in multi-stakeholder negotiation processes, where each stakeholder has his/her specific objectives. MCDM methods show what the space of possible solutions is given the objectives of each of the stakeholders (Van Latesteijn and Rabbinge, 1992; Rossing et al., 1997; Meyer-Aurich, 2005). MCDM methods are thus generally used in a normative approach. The MCDM models either incorporate the multiple objectives in the objective function or optimise one objective while using the other objectives as constraints. Or, they optimise farm profit, while taking the other objectives as externalities of the maximization of profit. There are many different MCDM approaches (Table 2.2), which made Rehman and Romero (1993) conclude that "the choice of a particular MCDM approach is in itself an MCDM problem!".

2.3.2. Risk

A farmer faces risk and uncertainty about the economic and environmental consequences of his actions due to his limited ability to predict e.g., weather, prices and biological responses to different farming practices (Pannell et al., 2000). Modelling with average data assumes risk neutrality (Thompson, 1982), as variation due to weather and price fluctuations is the source of risk and uncertainty.

Table 2.2. An overview of MCDM approaches used in BEFMs to simulate farm behaviour

Methodology	Abbre-viation	Reference	Description
Multiple Goal Linear Programming	MGLP	Rossing, et al. (1997); Zander and Kächele (1999); Ten Berge, et al. (2000)	a number of optimization rounds, in each of which one goal is optimized , while the constraints on the other goals are increasingly tightened
Weighted Goal Programming	WGP	Oglethorpe (1995);Wallace and Moss (2002); Weersink et al. (2004)	for each of the goals targets are specified and the overall objective is to minimize deviations from those targets
Lexicographic Goal Programming	LGP	Dorward (1999)	a form of WGP, but instead of the weights being relative as in WGP the weights are absolute or pre-emptive.
Modelling environmental effects as externalities		Pacini (2003)	one goal is maximized, other objectives are captured in indicator values resulting as joint outputs from agricultural production
Nearly Optimal Linear Programming	NOLP	Jeffrey et al. (1992)	solutions are produced that are not optimal with respect to any one objective, but instead are ‘nearly’ optimal for all objectives
Compromise Programming	CP	Yu (1973); Zeleny (1973)	solution closest to the Ideal Point is sought: the Ideal Point is the optimum value of different objectives given the constraints of the model and the preference of the decision maker
Multi-attribute Utility Functions	MAUT	Keeney and Raiffa (1976)	multi-attribute utility function is assumed for the decision maker or elicited from the decision maker, which is used to rank a set of finite alternative solutions
Outranking		Strassert and Prato (2002)	a procedure of several steps in which activities are compared in their achievement of several objectives

A distinction is often made between embedded and non-embedded risk (Dorward, 1999). Non-embedded risk is related to the uncertain yield and price levels beyond control of the decision maker, so without the possibility for the decision maker to respond to these uncertain yield and price levels and reducing the final risk. In the case of embedded risk the decision maker has the opportunity to exercise some control by sequential decision making (Dorward, 1999), thereby influencing the final risk he runs, for example by decreasing labour or pest control resources for potatoes if during the season the potato price drops.

Non-embedded risk is often defined as income variance. Pannell et al. (2000) found that farmer welfare was only reduced to a small extent with a large reduction in income variance. Therefore they argue that it is often not worthwhile to model non-embedded risk when assisting farmer decision making, as it is relatively less important in determining optimal farmer welfare than the correct representation of underlying biophysical relationships and the incorporation of tactical decision making. However, even if this is true in some conditions the agricultural activities and farm intensity selected by a model depend on whether or not non-embedded risk is incorporated (Oglethorpe, 1995; Pannell et al., 2000) as the model will select activities with a low variance in income, when non-embedded risk is avoided. We therefore think that in policy assessments it is useful to incorporate non-embedded risk if prices or yields do vary significantly. An example of an objective function to take account of non-embedded risk is (adapted from Freund (1956)):

$$\text{Max } u = e - \varphi \lambda \qquad (1)$$

with u as expected utility, e as expected income, φ as a exogenously determined risk aversion coefficient indicating to what extent the farmer avoids non-embedded risk and λ as the variance of income according to states of nature. This variance of income is calculated on the basis of the deviation of the expected income for each state of nature, where each state of nature has different weather and price conditions. Effects of weather variation can also be investigated by running the BEFM with technical coefficients derived from non-average weather data e.g. by modelling good and bad years (Gutierrez-Aleman et al. (1986b)), and assessing whether income can be maintained when extreme weather events occur (Gibbons et al., 2005). This indicates to what extent income and environmental effects are weather dependent.

Farm behaviour as related to embedded risk and sequential (often tactical) decision making depends on access to resource markets and opportunities the farmer has to adjust his decisions as information becomes available (Dorward and Parton, 1997). Thus, when using BEFMs for policy assessment and assisting farmer decision making, it may not be extremely relevant to model embedded risk in cases in which farmers have access to input markets for labour and short term capital (Deybe and Flichman, 1991) as these farmers will be able to maintain 'ideal' production activities by hiring in resources from outside the farm in case of unfavourable conditions occurring (Dorward, 1999). The construction and calculation of models incorporating embedded risk (stochastic programming models, Fig. 2.2, e.g. Apland (1993) and Dorward (1999)) are data and labour intensive as the size of a sequential decision problem increases rapidly (Hardaker et al., 1997), also termed curse of dimensionality (Bellman, 1957) so the extra effort and costs should be worthwhile (Dorward, 1999).

2.3.3. Time

Most BEFMs do not explicitly take account of time, i.e., they model a period with one time step. Dynamic models take account of time explicitly to capture some of the decision variables as functions of time (Blanco Fonseca and Flichman, 2002). A subdivision of dynamic models (Blanco Fonseca and Flichman, 2002) can be made in recursive models, intertemporal models and dynamic recursive models (Fig. 2.2). Recursive models are run over several periods; for each period the starting values are the end values of the last period (Wallace and Moss, 2002). Optimization is carried out for each period separately. Inter-temporal models optimize an objective function over the whole time period and allow for inter-temporal trade-offs between the time periods. For example, an objective function maximizes farm income over the whole time period, while considering the relative preference for current income above future income through a discount rate and the inter-temporal allocation of resources through a set of constraints (Pandey and Hardaker, 1995). Dynamic recursive models optimise over the whole period, while explicitly accounting for the dynamic interactions across years by using for each year starting values as the end values of the previous year (Louhichi et al., 1999). An example of a dynamic recursive model is used by Barbier and Bergeron (1999).

Stochastic programming models subdivide one year into several sub-periods (Fig. 2.2). They deal with information becoming available during the growing season and embedded risk (Section 2.3.2) by using a distribution of the values of technical coefficients at each time step. They can be said to be a type of dynamic model as they sub-divide one time step into several smaller time steps.

Out of 48 models studies considered, 37 used static models and 11 dynamic models. The static BEFMs ignore firstly the feedback on yields of adverse environmental effects (such as depletion of soil organic matter) on the longer term. However, static models can monitor what the environmental effects are of certain practices, for example with respect to soil organic matter (Dogliotti et al., 2005). Secondly, static BEFMs ignore the strategic decision making by farmers over many years, e.g., whether or not to build a new shed or incorporate a new enterprise in the farm system, and thirdly, they overlook the changing farm family objectives over time. Farm family objectives do change, as the farm family goes through a process of generation, maturation, decline and regeneration (Wossink and Renkema, 1994; Wallace and Moss, 2002).

Strategic decision making affects the farm system in the long term. According to several authors (Csaki, 1977; Hardaker et al., 1997; Wallace and Moss, 2002) it is of vital importance for the performance of the farm system that the farmer gets ‘big’ investment decisions right. Following Csaki (1977), investment depends on the availability of capital (dependent on the capital market and the capital position of the farm) and the need to invest. Once the investments are made, the increased fixed costs have to be paid and a farmer cannot easily move away from his investments. Only two model studies (Barbier and Bergeron, 1999; Wallace and Moss, 2002) incorporated strategic objectives into the objective function. Given the importance and relative absence of strategic decision making and investment in BEFMs, BEFMs could benefit from more attention to these aspects.

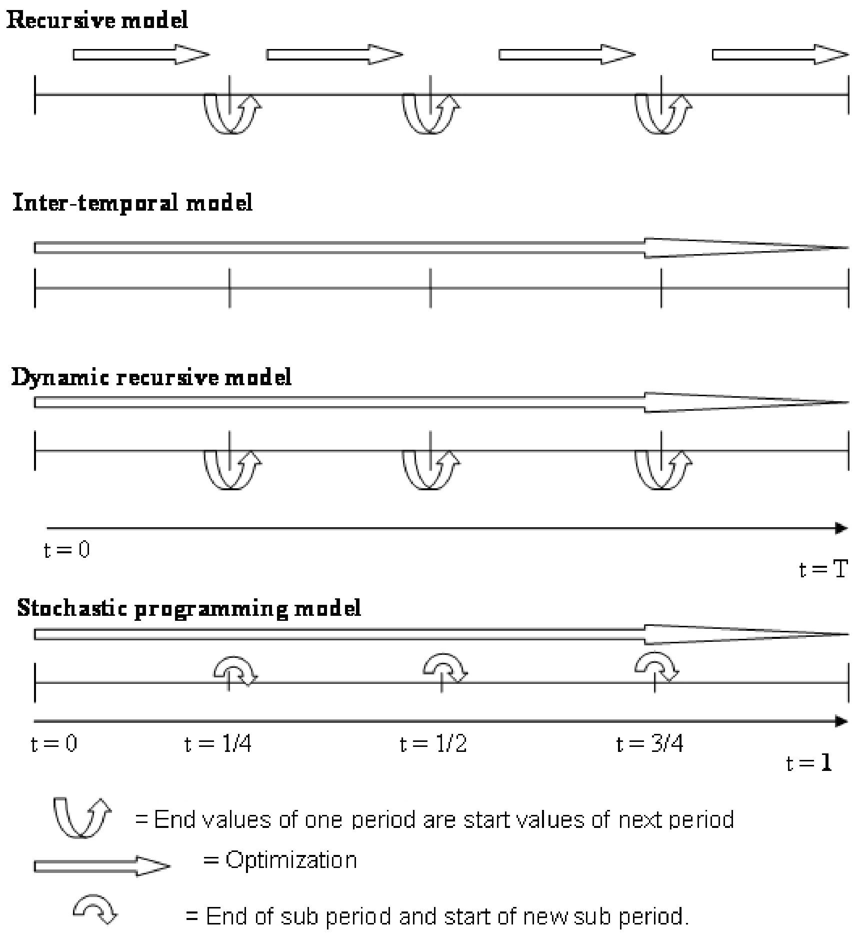


Figure 2.2. Different types of dynamic models (Blanco Fonseca and Flichman, 2002) and stochastic programming models

In conclusion, how to model farmer decision making, and whether or not to incorporate embedded or non-embedded risk and time depends first and foremost on the issues at stake (Weersink et al., 2004), as it can be a complicated task in terms of data requirements and model complexity. Our understanding of farm decision making is still limited, which is a hindrance in positive BEFMs more than in normative BEFMs. Incorporating non-embedded risk is justified when the interest is in the activities selected by the BEFMs rather than in the objective values only, while embedded risk needs to be incorporated when the farmer has poor access to resources and resource markets (labour, capital, inputs) to supplement his scarce resources during the season. If a large number of objectives, periods and risks are considered, BEFMs can become "bushy messes" (Hardaker et al., 1997), requiring large amounts of data and long solution time.

2.4. Agricultural activities

2.4.1. Activities to represent interactions between inputs

An agricultural activity consists of an enterprise, e.g. maize-wheat-potato rotation, sugar beet crop, dairy cows or beef cows, and a production technique describing the management of the activity (the inputs). An agricultural activity in BEFMs is described through the technical coefficients (TCs) or input-output coefficients. These technical coefficients are discrete estimates stating the amount of inputs needed to achieve certain outputs and the associated economic and environmental effects. Technical Coefficient Generators (TCGs) (Hengsdijk and Van Ittersum, 2003b) can then be defined as algorithms to translate data information into coefficients that represent the input and output relationships for each discrete activity.

The biophysical and economic rules that determine the transformation of inputs into outputs for a given activity are generally non-linear (Ten Berge et al., 2000). These non-linearities and the non-linearity of the production functions should ideally be embedded in the technical coefficients by defining several agricultural activities. Each agricultural activity then represents a point on the non-linear production function. Through the use of technical coefficients synergy between inputs and outputs can be taken into account. Agricultural activities are constructed according to a Leontieff production function (Leontief, 1986), in which inputs are used in fixed proportions, which is one of the core advantages of BEFMs compared to econometric methods using continuous production functions. Substitution between inputs is captured by formulating different agricultural activities in which different ratios of inputs are used (Hazell and Norton, 1986). Technical coefficients take account of the non-convexities of production and pollution as explained by Flichman and Jacquet (2003). Responses of a crop yield to a single input are usually concave, while responses of pollution to a single input are usually convex. However, if several inputs are considered jointly, yield and pollution curves may be, respectively, non-concave or non-convex (De Wit, 1992).

Current and alternative activities can be discriminated. Current activities are those being practiced on farms and can be derived from observed data or from experts with knowledge of the current situation. Alternative activities (alternatives in the remainder of this article) are not currently practised by specific farmers, but might be a suitable alternative for the future, often representing technological innovations or newly developed cropping or husbandry practices. These technical coefficients are usually generated and assessed using different sources of information, such as literature, census data, national statistics, farm management handbooks, expert knowledge, field trails and research farms. From the 48 model studies reviewed 13 did not mention their data sources for their technical coefficients.

Two approaches of estimating technical coefficients can be taken. The input-oriented approach implies that inputs serve as a basis for the calculation of outputs, which together form the technical coefficients. In the output-oriented approach the production target (output) is set dependent on the most limiting growth factor and on the objectives of the agricultural activity and then the most efficient set of inputs to realize this target is defined (Van Ittersum and Rabbinge, 1997; Hengsdijk and Van Ittersum, 2002). The latter method is particularly apt for alternatives.

2.4.2. Alternative activities

In assessment of technological innovations a 'very' large number of alternatives needs to be included, as this is the only way a BEFM can find the most promising alternative cropping and husbandry techniques from economic, social or environmental viewpoint (Hazell and Norton, 1986; Falconer and Hodge, 2000; Ten Berge et al., 2000; Hengsdijk and Van Ittersum, 2002). This 'very' large number of alternatives represents the technological innovation (e.g. precision weeding) in combination with other management aspects (e.g. irrigation and fertilization) of the farm that might influence the uptake of this particular technological innovation. In policy assessment the number of the alternatives can be relatively low; the alternatives defined should capture already identified promising techniques that are used by progressive farmers or broad categories of technologies that might be picked up due to the policy change. Alternatives must be feasible from a biophysical and technical point of view; whether or not they are socio-economically viable will be assessed in the BEFM (Hengsdijk and Van Ittersum, 2002)

Immense numbers of activities can potentially be incorporated, for example over 100,000 crop rotations can be generated if potentially 15 crops can be grown on a certain farm. The number of activities is commonly reduced to a feasible number based on expert judgement. This dependence on expert judgement poses the risk of missing out on activities that experts could not think of, thus limiting the solution space and feeding arbitrariness (Dogliotti et al., 2003). This risk is also noted by Hengsdijk and Van Ittersum (2002), who found that many land use studies hardly discuss or ignore completely the underlying concepts and data used for the description of activities and choices concerning the type of activities that are considered are not made explicit. Of the 48 model studies reviewed 18 mentioned and described the alternatives included in their model, while 18 model studies included only currently used activities. Of the remaining 12 model studies it could not be derived from the publication whether also alternatives or only currently used activities were included. To counter this risk of missing out on promising activities, Dogliotti et al. (2003) developed a tool (ROTAT) to generate all possible activities, in this case crop rotations and then reduce them to a feasible number of activities by the use of explicit filters. Generally speaking, alternatives should not be an arbitrarily selected set, but must be selected according to well-thought and explicit agronomic and socio-economic rules.

2.4.3. Level of analysis

Agricultural activities might be quantified or generated at different hierarchical levels, for example at crop, rotation, herd or livestock unit-scale. Whether to model at the rotation/herd or individual crop/livestock unit level largely depends on how the model takes account of time. A potential advantage of offering rotations and herds instead of crops and animals as activities to a static BEFM is that non-linear temporal interactions across crops and management alternatives can be captured outside the linear programming frame (Dogliotti et al., 2003). Hence, the structure of the BEFM remains simpler as interactions between crops and animal classes do not need to be modelled within the LP by adding rotational or herd constraints in the LP that restrict, for example, crops to be grown after other crops, or crops to grow in too high frequency within a rotation, or root crops to be grown too frequently. Static linear programming models can only capture temporal interactions by adding extra constraints with integer and binary variables. In a dynamic model it is probably easier to model at rotation/herd level in terms of model complexity, but the model can also be constructed at the crop level. In the studies reviewed, modelling at crop and livestock unit level was more popular than modelling at rotation scale: 24 out of the 48 were at crop and livestock unit level, while 4 were at herd and rotation level (in the other 10 model studies it is not explained at which level the activities are modelled). The models at crop level often ignored temporal interactions in the cropping system.

Interactions between plant and animal production, can be well captured within LP models (cf. Berentsen (2003)). A thorough discussion of the possibilities of modelling interactions between crops and livestock is provided by Thornton and Herrero (2001).

Next to temporal interactions and interactions between plant and animal enterprises, spatial interactions occur between adjacent agricultural fields or systems. Input-output relationships of agricultural systems could be defined as a function of output of soil and hydrological processes of adjacent agricultural fields or systems (Hengsdijk and Van Ittersum, 2002). For example in the case of erosion or run off, there is an input into an adjacent agricultural or non-agricultural field or system. To simplify and study these spatial interactions, Vatn et al. (2003) introduce the concept of partitioning to include lateral interactions.

2.5. Comprehensiveness

Obviously, a farm is organised differently than science: in a farm the social, economic, agronomic, environmental and institutional aspects are fully integrated and dependent on each other. In science these aspects are generally studied from different disciplinary perspectives. A BEFM that is weak in one of the disciplines is likely to lead to biased analyses. Constructing a BEFM thus requires integration in an inter/multi/trans-disciplinary set up. In principle a strong point of BEFMs is that they allow such integration of disciplines. In this Section BEFMs are assessed on their ability to accurately model all the different aspects of the farming system. Three general aspects which we consider important in the construction of a comprehensive BEFM for policy and technology assessment will be further discussed in the next paragraphs: 1. social milieu; 2 environmental impacts; and 3 new functions of agriculture.

Table 2.3 provides a comprehensive overview and indicates which aspects of farming systems have been often incorporated through the activities or constraints in 48 model studies. The analysis shows that some aspects are more popular than others; for example aspects related to nitrogen are often incorporated, but generally far little attention is paid to pests and diseases, off- and non-farm income, soil fertility as a constraint, soil organic matter, landscape quality, and biodiversity and nature.

Table 2.3. Different aspects of a mechanistic BEFM and the number of model studies that explicitly addressed these aspects via inclusion of activities or constraints in the mathematical programming model. (* = mainly nitrogen; ** = linked to rotational constraint)

	Aspects	Number of studies
Activities	input expenses	39
	nitrogen balance	25
	farmer time allocation	25
	capital availability	24
	run off/leaching	16
	soil type/soil depth	15
	weather variability	11
	emissions	10
	phosphorus balance	9
	erosion	9
	potassium balance	6
	biodiversity and nature	6
	Soil Organic Matter***	5
	off- and non-farm income	4
Constraints	rotational constraints	27
	machinery availability	26
	labour/planning	23
	nutrient availability	14*
	capital availability	14
	regulations/laws and subsidy schemes	14
	water supply	13
	production quotas	11
	use of inputs	11
	pests and diseases	8**
	soil fertility	5
	run off/leaching	5
	transport	4
	emissions	2
	slope	2
	erosion	1

*** = mainly nitrogen; ** = linked to rotational constraint; *** = includes C-sequestration**

2.5.1. Social milieu

A farmer often does not decide independently on how to react to a policy and a potential technological innovation: he is influenced by his social milieu (Anderson et al., 1985). The utility of BEFM for policy assessment is limited by their lack of understanding of the dynamics of the farm household and of the impact of psychological and cultural values on farmer decision making (Dent et al., 1995), as no mechanistic BEFMs were found that incorporated the social milieu of the farmer. As Anderson et al. (1985) noted: "If FSR (Farming Systems Research) is not 'major-crop' biased, the farmer of relevance in many cases will be a woman. Since the preferences of women are likely to be different from men, omission of the women's viewpoint is likely to lead to misspecified models." Presumably the most important factors of this social milieu are the other members of the farm family (Edwards-Jones and McGregor, 1994; Dent et al., 1995; Ruben et al., 1998) and farm families living in the neighbourhood (Berger, 2001). Wossink et al. (1992) propose to distinguish different categories of family farms as to their financial and technical status. But not only the economic background of the farm families is important, also social parameters like attitudes, values, traditions, peer group pressure and culture should be considered (Dent, 1990). Potentially suitable objective functions for the incorporation of the social milieu are an additive utility function, which adds up the utility functions of the individual household members or a constrained objective function which is constrained by certain basic goals other family members have and that are entered as constraints. The use of so-called farming styles (Van der Ploeg, 1994) to distinguish groups of farms with different strategies due to farm internal and external factors might be useful for application of BEFMs to different farm types.

2.5.2. Environmental impacts

Environmental impacts are a result of agricultural practices. This is also termed 'joint production' of agricultural outputs and environmental effects (Falconer and Hodge, 2001). Those environmental impacts should be incorporated, that have a clear relation to agricultural practices. Most often environmental effects are modelled through indicators. Effect-based indicators are indicators based on the results in terms of environmental effects, while means-based indicators are based on changes in agricultural practices that could lead to a better environmental performance (Payraudeau and Van der Werf, 2005). Effect-based indcators are preferred above means-based (Payraudeau and Van der Werf, 2005), because they characterise the environmental risk more directly and are easier to validate.

2.5.3. New functions of agriculture

New functions of agriculture, according to the European Union (EC, 1999), are preservation, management and enhancement of the rural landscape, protection of the environment and a contribution to the viability of rural areas. Some of these new functions can be modelled and quantified by including extra activities a farmer might incorporate on his farm, for example a recreation-activity, others can be modelled by quantifying the positive or negative externalities of activities, for example effects of cows on landscape quality. Quantification of the production of 'rural landscape' and 'protection of the environment' is far more difficult than the quantification of extra income from, for example, farm shops. Production of 'rural landscape' can be understood as maintenance of biodiversity and the provision of a pleasant landscape. Biodiversity can be measured with indicators, for example crop diversity indicator, livestock diversity indicator, herbaceous plant biodiversity indicator (Pacini, 2003), or indicators for wild plants, partridges or amphibians (Meyer-Aurich et al., 1998). These indicators, however, either focus only on the agro-biodiversity, or focus on the single species rather than on the complex interactions in food webs underlying biodiversity. It is also challenging to find indicators for the provision of a pleasant rural landscape as pleasantness of landscape is largely subjective. Potential indicators regarding landscape issues which might be further explored are presence, size and amount of landscape elements like field margins, hedges, pools, wetlands, etc (Hendriks et al., 2000; Groot et al., 2007).

2.6. Evaluation of BEFMs

A challenge in the use of BEFMs is to ensure that its results can be trusted as sensible and reliable and that the model can be re-used. This Section therefore discusses model evaluation and transferability of BEFMs.

2.6.1. Model evaluation

A broad definition of model evaluation is given by Jansen (1997): the major method of showing the reliability of a model for a purpose. The BEFM and its outcomes should closely match reality (Gutierrez-Aleman et al., 1986b). Of the 48 BEFMs reviewed 23 carried out some form of comparison with actual farming practices, of which 8 BEFMs fitted their simulated results to observed data by (automatically) adjusting model parameters as part of a calibration procedure. Only Thompson (1982), Schilizzi and Boulier (1997), Ramsden et al. (1999) and Vatn et al. (2003) describe the comparison between their model outcomes and actual farming practices quantitatively, while others only briefly mentioned the fit with observed data without discussing the quality of fit. The gap between model outcomes and actual farming practices (Wossink et al., 1992) gives an indication of the ability of the model to come close to reality (Thompson, 1982). This gap varied from 5 to 10% in land use at farm level and input coefficients at activity level for Thompson (1982), from 1 to 65% in land use at farm level and input coefficients at activity level for Vatn et al. (2003), from 15 to 40% in income for Schilizzi and Boulier (1997) and from 7% in total production at farm level to 40% in input-output coefficients at activity level for Ramsden et al. (1999). Model outcomes contain a number of different variables, so a model may match closely actual farming practices for one variable, while a large gap exists for other variables. The fact that only four references were found, which explicitly discuss their thorough model evaluation, demonstrates the urgent need for more work in this area.

Four reasons can cause the gap between model outcomes and actual farming practices (Wossink et al., 1992), i.e., poor specifications of objective functions, missing dynamic aspects, poorly defined activities, and an incomplete models. This gap can be minimized by making the model more specific and comprehensive and hence complex. Obviously, a trade-off between simplicity and greater accuracy exists (Thompson, 1982) (Fig. 2.3).

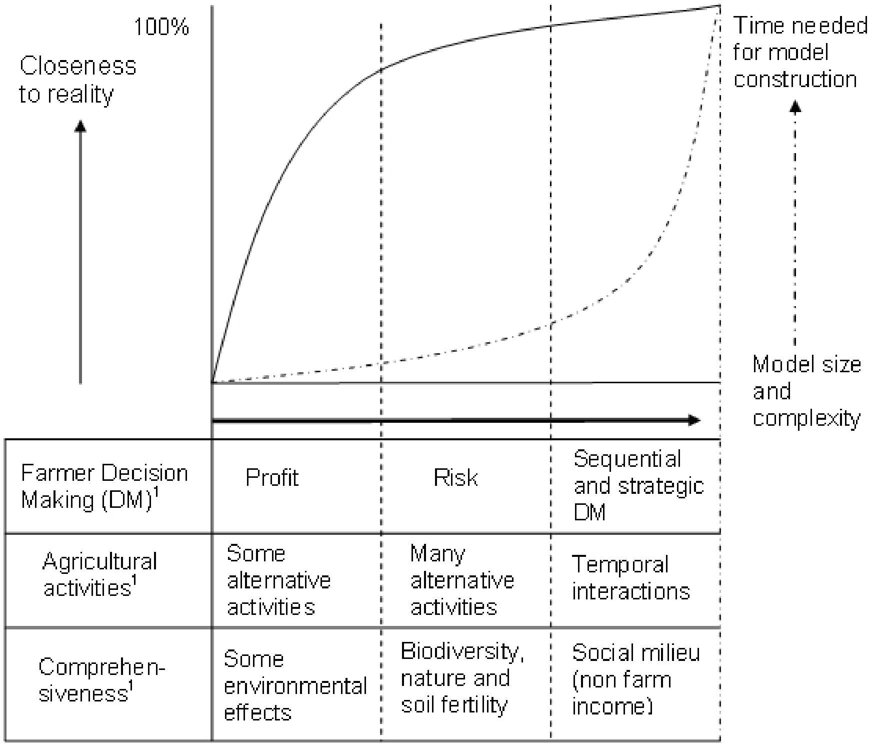


Figure 2.3. The relation between model size and complexity, development time needed and closeness to reality and some factors part of farmer decision making, agricultural activities and comprehensiveness, that make models increasingly complex and time consuming. ([1] = the factors included here are only an example)

Even if the model is robust to changes in parameters, it still does not mean that the model structure itself is correct (Pannell, 1997). As part of model evaluation a model can be validated by introducing a new dataset and assessing whether the model can without changes in structure and parameters adequately reproduce the new observed values (Thomann and Muller, 1982). This new dataset could refer to a different year

with a policy change or a technological innovation or to a similar farm in another region. Such a validation exercise was explicitly mentioned by only 5 out of the 48 model studies.
Sensitivity analysis of an LP model to parameter values represents a special case according to Rossing et al. (1997) and Makowski et al. (2001), as different technical coefficients may have little effect on the realisation of objectives, but leads to very different activities selected by the optimization procedure. Eight model studies worked with sensitivity analysis.

2.6.2. Transferability

BEFMs in principle provide the ability to replicate assessments for a vast range of spatial conditions and farming practices. A BEFM that is easy to transfer between locations or farm types is called generic. Although some of the model studies claim that their model is easily transferable, no evidence from the literature has been found trying to transfer one model between several locations and farm types. This could be due to on the one hand these models being very specific for a location or farm type or on the other hand modellers preferring to build their own model rather than re-using existing models. Data needs, size and structure of a BEFM can limit its transferability. A simple, small, easily manageable model with a clear structure is probably easier to transfer, but it requires time and effort to make a BEFM generic. The lack of generic BEFMs could limit the uptake of BEFM as a tool for assessments of policies and technological innovations on a larger scale outside the research domain, as the development and use of BEFMs remains a time and resources consuming exercise requiring specialist knowledge of researchers.

2.7. Good practices and research agenda

Bio-Economic Farm Models enable assessment of policy changes and technological innovations as claimed by a large number of studies, which carried out such an exercise. In previous Sections different aspects of BEFMs were explored, thereby drawing up a state-of-the-art in terms of strengths and shortcomings of present

BEFMs. In this section, we identify some good practices in the use of BEFMs and draw up a research agenda for the coming years based on our assessment of methodological shortcomings or limitations.

The following good practices were identified based on the review of 48 model studies. First, authors should clearly state or discuss the end-use of their BEFM, i.e. policy assessment, assisting farmer decision making or methodology development, and link their end-use to the assumptions they are making in their modelling. The modelling purpose has strong implications on the details of the BEFM as argued in this chapter.

Second, the definition of agricultural activities that form the input to the model should be explicit and documented in any publication, as the inputs determine the outputs of any model. Too many model studies do not mention sources of their data on technical coefficients, while many others did not explicitly discuss assumptions in formulating their current and/or alternative activities.

Third, model evaluation is a vital part of any application of a BEFM and should be explicitly and comprehensively presented in any BEFM publication, as it is the only way of conveying that the assumptions made during the modelling exercise are valid and acceptable. Few of the cited studies explicitly address model evaluation. As a result, the reader of these model studies cannot objectively judge the quality of the BEFM, and the discussion of the results looses grounding.

Fourth, all constraints incorporated in the model across the different scenarios should be explicitly mentioned and discussed. Given that it is very difficult and often not even accepted to provide all modelling details in the form of mathematical equations or LP tableaus in a scientific paper, the models together with their documentation should be made available for download.

Our review suggests several shortcomings in current research. As a first item on a research agenda, it would be essential to develop a consistent and widely accepted model evaluation procedure, comprising steps of checking the correspondence with observed values, calibration and validation. Secondly, it should be further investigated how strategic farmer decision making, including possible effects of the social milieu, can be more adequately represented in BEFMs, especially when targeting policy assessments.

As a third research challenge, in the previous Sections some suggestions were made of aspects that could be incorporated in a BEFM, which have so far only received limited attention e.g. investment decisions, pests and diseases, biodiversity and landscape quality and temporal effects of soil fertility.

Finally, we suggest the development of an easily transferable BEFM with a generic and modular structure, as currently many BEFMs exist, amended to specific locations or purposes. Existing BEFMs are rarely re-used and the newly developed models and their applications do not add a lot of new features or approaches to the body of literature. An easily transferable BEFM with a generic and modular structure could enable a group of researchers to work jointly on one model, extending it with new features and allowing re-use across data-sets, farm types and locations.

Chapter 3. A generic approach to identify alternative agricultural activities for future studies

Abstract

In science different future-oriented methodologies (e.g. mono-disciplinary simulation modelling, integrated assessment, prototyping agricultural systems) have been developed for assessment of the sustainability of agricultural systems. Alternative activities are crucial in future-oriented studies. Alternative activities are technically feasible, alternative production options, often technological innovations or newly developed cropping practices. This chapter describes a theoretical framework to specify alternative activities and its implementation in the Agricultural Management Model (AMM). The theoretical framework and the AMM are based on an heuristic approach of filtering alternatives from possible permutations of agricultural activities. With respect to four challenges in defining alternative activities, it is argued that (i) the theoretical framework is generic across locations, farm types and data sets, (ii) the AMM offers a typology of management to consistently define parameter values for innovations, (iii) the theoretical framework forces an explicit consideration of all assumptions by using an heuristic filtering approach, and (iv) the AMM is able to manage large numbers of alternative activities. The process of identification of feasible and plausible alternative activities may be more valuable than the alternatives finally generated.

Sander Janssen, Huib Hengsdijk, Roelof Oomen, Ioannis N. Athanasiadis, Edwin van der Maden and Martin van Ittersum, A generic approach to identify alternative agricultural activities for future studies. To be submitted to European Journal of Agronomy

3.1. Introduction

3.1.1. Future-oriented studies and agricultural management

Agriculture in Europe faces a continuously changing environment due to trade liberalization, globalization, evolving societal needs and climate change. This changing environment requires agriculture to adjust the means and methods of production. Multi-scale innovation and adaptation is required to enhance the sustainability of agricultural systems and their contribution to sustainable development at large. Integrated assessment of these innovations may contribute to the increased efficiency and effectiveness of decision-making (Bland, 1999; Van Ittersum et al., 2008). In science different methodologies (e.g. mono-disciplinary simulation modelling, integrated assessment, prototyping agricultural systems) have been developed for such ex-ante assessments.

Different disciplines have developed different types of simulation models for ex-ante assessments, such as cropping systems models to simulate crop growth, productivity and externalities in response to climate change and management (Van Ittersum and Donatelli, 2003); bio-economic farm models linking farmers' resource management decisions to current and alternative production possibilities described by input-output relationships and associated externalities (Wossink et al., 1992; Donaldson et al., 1995); partial and general equilibrium models based on optimization techniques representing (part of) the agricultural economy in terms of markets or trade (Hertel, 1997; Heckelei and Britz, 2001); and land use change models simulating the competition for land between different land uses subject to specific allocation rules (Verburg et al., 2008).

Integrated assessment (IA) is the interdisciplinary process of combining different strands of disciplinary knowledge to coherently represent complex societal problems of interest to decision makers (Rotmans and Asselt, 1996; Harris, 2002; Parker et al., 2002). Recently, various IA projects on European agriculture and land use have been conducted (Rounsevell et al., 2005; Helming et al., 2008; Van Ittersum et al., 2008), largely using the linkage of disciplinary simulation models.

Prototyping farming systems (Vereijken, 1997) is a method to design alternative farming systems through a step-wise procedure. This procedure consists of defining current shortcomings and objectives for improvement, designing a theoretical prototype fulfilling the objectives, test the theoretical prototypes on pilot farms, and if successful, disseminate the prototype to farm practice.

Simulation modelling, integrated assessment and prototyping have in common that they require some kind of specification of agricultural activities carried out on farms. An agricultural activity is a coherent set of annual or perennial crops or animals plus the operations with associated inputs resulting in, for example, the delivery of a marketable product, the restoration of soil fertility, or the production of feedstuffs for on-farm use (Van Ittersum and Rabbinge, 1997; Ten Berge et al., 2000). In practice farmers have many options to convert inputs into outputs and the ultimate choice for agricultural activities is based on a number of factors comprising farmer objectives and endowments, available technology and knowledge, prevailing weather, policies and the economic environment.

We distinguish current and alternative agricultural activities. Current activities represent the current means of production in a given region which can be derived from observed data, either through surveys (Zander et al., 2009), statistics (EC, 2008a; Eurostat, 2008; FAO, 2008) or expert knowledge (Zander et al., 2009). Alternative activities are technically feasible alternatives, often technological innovations or newly developed cropping practices not yet practiced at a wide scale in a region under study (Van Ittersum and Rabbinge, 1997; Hengsdijk and Van Ittersum, 2002). In future-oriented studies focusing on short term predictions (e.g. 5-10 years) (Van Ittersum et al., 1998) alternative activities might be closer to the current activities that are already practiced on other farms in the region or elsewhere. Alternative activities may deviate considerably from current activities in studies that are long term explorations (i.e. >15 years) (De Koning et al., 1995; Lu et al., 2004).

Ignoring these alternative activities in future-oriented studies occurs frequently (Hengsdijk and Van Ittersum, 2002; Janssen and Van Ittersum, 2007), which is equivalent to assuming that agricultural activities in the future will be exactly the same as current activities. Alternative activities are crucial in future-oriented studies, as they allow for explicit consideration of potential changes that can occur in agricultural practices.

3.1.2. Four challenges in alternative agricultural activities

We identified four challenges in the definition and quantification of alternative activities on the basis of previous research. First, alternative activities need to be build following an explicit conceptual framework with consideration of production ecological principles (Van Ittersum and Rabbinge, 1997). A framework based on goal orientation and agro-ecological engineering has been proposed (Hengsdijk and Van Ittersum, 2002; Hengsdijk and Van Ittersum, 2003b) (Challenge 1). In this framework a production target or objective for the activity is set, after which inputs to reach this production target are calculated (Van Ittersum and Rabbinge, 1997; Hengsdijk and Van Ittersum, 2002; Hengsdijk and Van Ittersum, 2003a), the so-called output-oriented approach. In contrast, in an input-oriented approach inputs serve as a basis for the calculation of outputs.

Second, parameter values need to be estimated for new and innovative agricultural technologies for which only limited data is available (Hengsdijk and Van Ittersum, 2003a) (Challenge 2). These new and innovative agricultural technologies have often only been applied at research stations or pilot farms, and limited data or only expert knowledge might be available. Parameter values for inputs and outputs can be calculated according to different outlooks through a simple model describing biophysical or economic processes (Hengsdijk and Van Ittersum, 2003a; Dogliotti et al., 2004).

Third, a comprehensive set of alternative activities must be considered and not a limited subset to avoid the risk of arbitrary subsets (Hengsdijk and Van Ittersum, 2002; Dogliotti et al., 2003) (Challenge 3). Dogliotti et al. (2003) and Bachinger and

Zander (2007) developed a tool (e.g. ROTAT and ROTOR) to generate alternative rotations by generating all possible rotations and filtering those to a set of feasible rotations. Similarly, Klein Haneveld and Stegeman (2005), Castellazzi et al. (2008) and Detlefsen and Jensen (2007) modelled the rotation generation problem through Linear Programming or network flow approaches.

Fourth, potentially large numbers of alternative activities exist for any given situation, which need to be managed (Wossink et al., 1992) (Challenge 4). In defining alternative activities, many different aspects can be varied, either in isolation or jointly, which is related to the uncertainty on the future. Combinatorial explosions (Wossink et al., 1992; Dogliotti et al., 2003) occur, leading to immense numbers of alternatives, that cannot be scrutinized manually. Tools are required to manage these immense numbers

This chapter contributes to the further improvement of methods to deal with these challenges in defining alternative activities. With respect to Challenge 1, the chapter proposes a generic theoretical framework, that does not limit itself to goal setting or generation of rotations. The proposed theoretical framework consistently builds on steps of generating and filtering alternatives and includes both strategic decisions on production enterprises and tactical decisions on annual management of those enterprises. This framework is implemented in a set of components for arable systems, that are able to manage very large numbers of alternatives (cf. Challenge 4). Through an application of these components, the impact of rules and assumptions on the set of feasible alternatives is demonstrated (cf. Challenge 3). Finally, with respect to Challenge 2, this chapter proposes some different methods of specifying parameter values building on established methods (Van Ittersum and Rabbinge, 1997; Hengsdijk et al., 1999; Hengsdijk and Van Ittersum, 2003a; Van de Ven et al., 2003).

The next section briefly introduces important concepts and definitions of agricultural activities in more detail. Section 3.3 introduces the theoretical framework to generate alternative activities, and Section 3.4 presents the components based on this theoretical framework. Section 3.5 presents an application for the Flevoland region in the Netherlands. Finally, Section 3.6 offers a discussion of the framework and some conclusions.

3.2. Concepts and definitions

Production functions describe the relationships between outputs that can be produced from a given set of inputs (Coelli et al., 1998). Two types of production functions are relevant for modelling agricultural activities, i.e. discontinuous production functions (e.g. Leontief (Leontief, 1986), Liebig, Liebscher and Mitscherlich (De Wit, 1992)) and continuous production functions (e.g. Cobb-Douglas production functions (Cobb and Douglas, 1928)). In agricultural activities constructed according to discontinuous Leontief production functions, inputs are used in fixed proportions and corresponding outputs are quantified. Proportionally increasing the quantities of all inputs leads to a proportional increase in the quantities of outputs. Continuous production functions establish a continuous mathematical relationship between the use of inputs and production of outputs.

In agricultural production at farm level, continuous production functions are difficult to construct due to non-linearities (Hazell and Norton, 1986; Ten Berge et al., 2000) and non-convexities (Flichman and Jacquet, 2003) occurring in the biophysical and economic process that determine the conversion of inputs into outputs for a given activity. These non-linearities and non-convexities are caused in the interaction between many different inputs used (Flichman and Jacquet, 2003) and by changing biophysical conditions. Therefore, the use of Leontief-production functions to describe agricultural activities is preferred and this chapter focuses on activities based on such functions. There have been attempts to derive continuous production functions from Leontief production functions (e.g. Ruben and Van Ruijven, 2001). An activity according to Leontief production function is characterised by a set of coefficients (Technical Coefficients, TCs, or input-output coefficients) that express the activity's contribution to the realisation of user-defined goals (or objective in modelling terms) (Ten Berge et al., 2000). Hence, the characterization of activities

through technical coefficients must take into account any non-linearity and non-convexity.
A production orientation is defined as a set of value driven aims and restrictions of the agricultural activity that affect the input and output levels (Van Ittersum and Rabbinge, 1997). Aims and limitations of agricultural activities are based on the (often implicit) preferences of stakeholders and on the purpose of the study. The concept of production orientation requires that common terms such as ‘integrated’, ‘organic’, ‘conventional’ or ‘labour-saving’ are made explicit in terms of TCs. For example, conventional activities commonly aim at profit maximization with a high cropping frequency of the most profitable crop and the use of external inputs.

3. Theoretical framework

3.3.1. Overview

In our view, for a conceptualization of agricultural activities to be generic, it must meet the following requirements: (i) applicable across locations and data sets; (ii) flexible in terms of detail depending on the purpose, i.e. multi-scale applicability; (iii) comprehensive in representing all possible management options (iv) consistent across locations and datasets, (v) transparent in its application to different locations, datasets and policy questions, (vi) decomposable into smaller modules, that can be independently applied. The proposed framework to develop alternative activities must fulfill these requirements. To fulfill these requirements, we propose a theoretical and abstract framework that be implemented and adjusted to specific locations, research questions and available data sets.
Our framework links production enterprises to production techniques in such a way that the combination can be assessed in different types of simulation models (e.g. cropping system models, bio-economic farm models and budgeting models). A production enterprise is the description of the temporal structure of annual cropping within a field, i.e. a crop rotation. A production technique is a complete set of agronomic inputs characterized by type, level, timing and application technique

(Van Ittersum and Rabbinge, 1997). Together a production enterprise and production technique define the inputs of an agricultural activity.

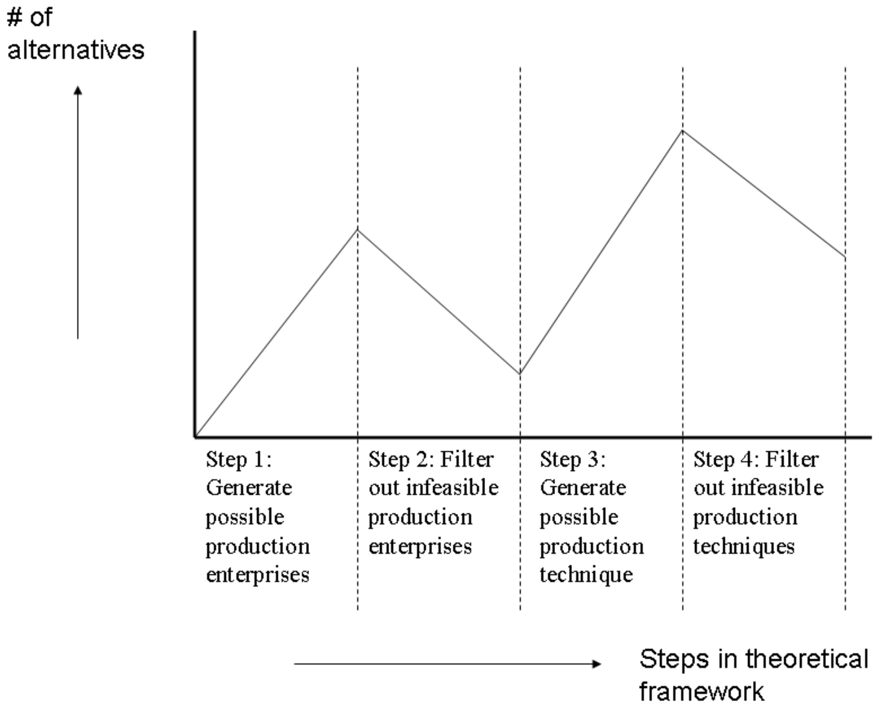


Figure 3.1. Steps to identify sets of feasible production enterprises and techniques.

The proposed framework consists of four steps (Fig. 3.1). The first step is to generate all possible combinations of crops as production enterprises. In the second step these possible production enterprises are filtered using rules-of-thumb and more knowledge-based rules to arrive at a set of plausible production enterprises. The third step is to generate all possible combinations of production techniques for the production enterprises. Finally, in the fourth step a selection of the production techniques is made based on all the possible combinations and another set of rules-of-thumb. This theoretical framework is based on first generating all possible alternatives as permutations and subsequently filtering out impossible, implausible and impractical combinations on the basis of scientific and expert knowledge.

3.3.2. Permutation generation

The steps to generate possible production enterprises (step 1) and production techniques (step 3) are based on the generation of permutations (Sedgewick, 1977). A permutation is an ordered sequence of elements from a set of elements, in which not necessarily all elements of the set have to be used. For example, given a set of four elements A, B, C and D, permutations consisting of three elements are, ABA, CDB and DCA. Generating possible production enterprises or production techniques from a given set of crops and possibilities in crop management, implies generating permutations of these production enterprises or production techniques. Two cases can occur in generating permutations for alternative agricultural activities, which is related to activities being conceptualized as a combination of production enterprise and a production technique.

In the first case only production enterprises are considered and possible production enterprises follow the generation of cyclic permutations (Athanasiadis et al., 2007a). These exclude cyclic equivalents of a permutation (i.e. a possible rotation). In other words, the rotation 0-1-2 is the same as rotation 1-2-0, where 1, 2 and 0 represent three different crops, and one rotation can be excluded. Given a set of *n* crops, and *r* the rotation length in years, each permutation can be uniquely identified by a single integer *i*, where $i \geq 0$, and $i < n^r$. Each permutation can also be identified by a unique sequence of *r* digits *d*, in which each digit represents a position in the permutation (e.g. a year in a rotation) in reverse order. The value of the single integer *i* of a sequence of r digits "$d_r \cdots d_k \cdots d_2 d_1$" in the n-base system is given by the equation:

$$i = \sum_{k=1}^{r} d_k \cdot n^{k-1} .$$

As an example, the four-digit permutation 1-0-1-2 based on the set {0,1,2} with 3 elements represents the number $1 \cdot 3^3 + 0 \cdot 3^2 + 1 \cdot 3^1 + 2 \cdot 3^0 = 27+0+3+2=30$. This permutation 1-0-1-2 has (at most) r cyclic equivalents, with index *i*:

$$i_m = \sum_{k=1}^{r} d_{k-m} \cdot n^{k-1} \text{, where } m=1 \ldots r.$$

In the second case, both production enterprises and production techniques are considered. In this case, restrictions apply in the combination of production enterprises with production techniques. Consider the following example, we have a production enterprise specified as a two year rotation with maize in year 1 and peas in year 2. For maize two production techniques have been specified with different levels of nitrogen fertilization (M1 and M2). For peas, also two different production techniques have been specified with different levels of phosphate fertilization (M3 and M4). For the maize and peas rotation, the four possible combinations are: maize-M1 and peas-M3; maize-M2 and peas-M3; maize-M1 and peas-M4; maize-M2 and peas-M4. For this example with restrictions on the combination of production enterprises and production techniques, generating the permutations represents a case of vector multiplication, in which each year is represented by one vector. The result of such a vector multiplication is a matrix, containing all permutations. If *d* denotes the number of elements in subset *K* available for each *c* in set *C*, than the total number of possibilities is equal to $d_r * d_{r-1} * \ldots d_2 * d_1$ for an activity of *r years.* The matrix with all permutations can be constructed by looping iteratively through each of the elements of subsets *K* and combining the elements.

3.3.3. Knowledge based filters

On the basis of production ecological and economic knowledge and insights, the permutations of production enterprises generated in step 1 and of production techniques generated in step 3 are filtered in the steps 2 and 4, respectively. These filters combine, scientific knowledge from peer-reviewed publications and expert knowledge from advisory handbooks or crop specialists. The filters exclude permutations by identifying impossible, implausible or impractical crop combinations or crop and production technique combinations given the available knowledge. For example, frequent repetition of potato in a rotation is not possible due to soil nematodes or the available irrigation equipment on a farm does not allow to irrigate both the maize and soybean crop (e.g. an impractical agricultural activity). The filters can be more or less specific, for example referring to crop group (e.g.

root crops) or referring to a specific crop (e.g. wheat). The more filters are specified, the more permutations can be excluded.

3.4. Implementation of the framework: the agricultural management model

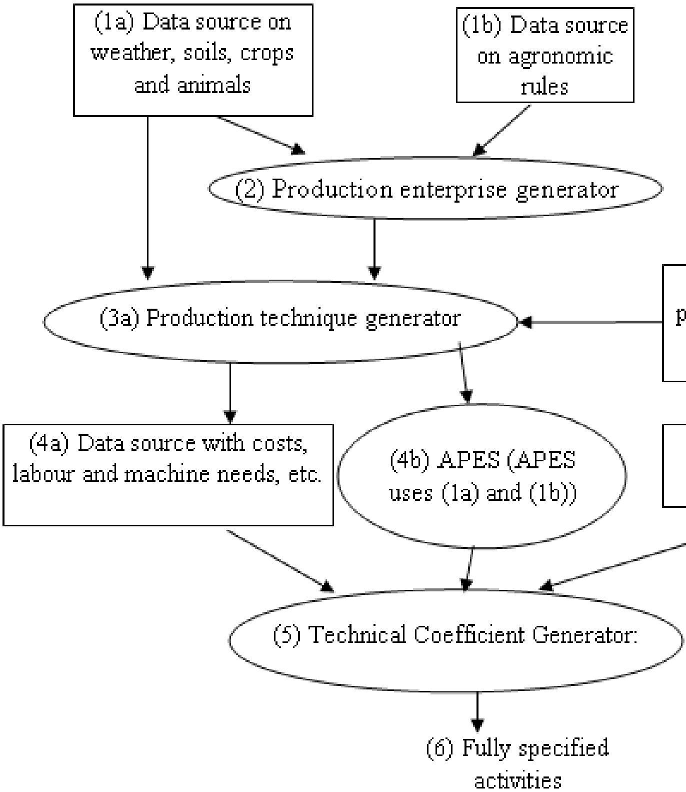


Figure 3.2. Agricultural Management Model and its components: algorithms, databases and connections. Ellipses are the components, squares are the data sources and arrows indicate input-output relationships. The Agricultural Production and Externalities Simulator (APES) (Donatelli et al., 2009) is an example of a cropping system model that has been linked to the AMM.

3.4.1. Introduction

The theoretical framework for alternative agricultural management is implemented as the Agricultural Management Model (AMM). The aim of AMM is to describe current activities, generate alternative activities and quantify the activities through all the required technical coefficients. These activities can be evaluated by a cropping system model and can be used in bio-economic farm, agricultural sector models or other future-oriented studies. AMM is a part of the bio-economic farm model FSSIM (Louhichi et al., 2009) in the SEAMLESS model chain (Van Ittersum et al., 2008). In this chapter we only focus on the components of the AMM to generate alternative activities.

The AMM has a component-based set up, e.g. it can be dissected in distinct autonomous parts that communicate with other components or provide services to other components. The AMM consists of three main components: (i) Production Enterprise Generator generating production enterprises, (ii) Production Technique Generator generating and specifying the production techniques of production enterprises, and (iii) Technical Coefficient Generator quantifying, collecting and formatting the technical coefficients (Fig. 3.2).

3.4.2. Production Enterprise Generator (PEG)

PEG is an extended version of ROTAT (Dogliotti et al., 2003) and is thus a tool to generate feasible sets of farm production enterprises using suitability filters based on crop, soil and climate characteristics. The PEG aims to design production enterprises in a coherent, transparent and reproducible way. The PEG contains a number of crop and rotation suitability filters that limit in an early stage the number of crop rotations for which production techniques need to be defined. The crop and rotation suitability filters consist of sets of pre-defined criteria to exclude options. The crop filter procedure has great similarities with the guidelines for land evaluation matching crop requirements and land qualities (FAO, 1976). Many of the diagnostic criteria relate to biophysical characteristics of land and crops (Table 3.1), but may be extended to include other type of criteria. For example, the unavailability of machinery may constrain the production of certain crops, and these crops can be excluded from the final generated set of rotations in the PEG. Different from the

FAO-evaluation procedure is that the PEG does not identify suitability ratings but it determines whether a soil and climate are either suitable or unsuitable.

Table 3.1. Examples of crop suitability filters that can be adapted to a situation.

Name	Description
Slope[1]	The steepness of the slope
Clay content[1]	The percentage of clay in the soil
Rooting depth[1]	The depth to which roots can grow in the soil
Roughness[1]	The roughness or size of stones and crumbs in the soil
Salinity[1]	The percentage salt in the soil
Alkalinity[1]	The acid neutralizing capacity of the soil
Drainage[1]	The capacity of the soil to transport water
Minimum Temperature sum[2]	The minimum accumulated daily mean temperature above a crop-specific threshold temperature required to complete a full phenological crop cycle from emergence to maturity
Maximum Altitude[3]	Altitude may be an appropriate filter in mountainous regions to account for low temperatures, risks of climatic hazards (e.g. excess of water) and lack of suitable land in such areas.
Maximum rainfall[3]	High rainfall limits product quality and, for example, results in (cereal) grains with too high moisture content, and reduces the number of workable field days during harvest.

[1] Based on Reinds and Van Lanen (1992)

[2] Based on the Crop Growth Monitoring System (Alterra and INRA, 2005)

[3] Based on Russel (1990) and Wolf et al. (2004)

After having filtered out impossible crops for a given biophysical situation, the PEG generates on the basis of the feasible set of crops crop rotations by cyclic permutations (Section 3.3.2) while applying suitability filters for rotations (Table 3.2). The filter procedure comprises diagnostic criteria to exclude rotations that are not feasible from a agronomic point of view or less desirable from a phyto-sanitary point of view (Table 3.2). For example, growing potato as a mono-crop is not desirable from a phyto-sanitary view, due to soil borne diseases. Growing winter oil seed rape after sugar beet is not possible from an agronomic point of view as the sugar beet has not been harvested when the winter oil seed rape must be sown. The filters in the PEG are the same as those in the ROTAT-tool developed by Dogliotti et al. (2003).

Table 3.2. Set of rotational suitability filters in the Production Enterprise Generator based on Dogliotti et al. (2003)

Name	Description
Sowing harvesting filter	A timing constraint indicating that a crop cannot be planted before the previous crop is harvested.
Minimum intercrop period	The minimum period required in days to prepare the field for the next crop after harvesting
Crop sequence filter	Certain crops are not possible as predecessors to other crops due to soil borne pest and diseases.
Crop frequency filter	The maximum frequency of a crop in a rotation
Crop group frequency filter	The maximum frequency of a group of crops in a rotation
Crop repetition filter	The minimum number of years before repetition of crop is allowed
Crop group repetition filter	The minimum number of years before repetition of crop from the same crop group is allowed
Rotation length	The maximum and minimum rotation length
Number of different crops	The maximum number of different crops

3.4.3. Production Technique Generator (PTG)

The PTG describes production techniques of agricultural activities for the feasible set of production enterprises. The PTG characterizes production techniques and identifies infeasible production techniques for well-defined production orientations. For arable cropping activities, we sub-divided production techniques into water management, nutrient management, weed, pest and disease management, conservation management (including tillage) and general management (e.g., field inspection, planting and harvesting). Each of these management practices consists of several aspects (Table 3.3). For management practices involving the use of external inputs these aspects comprise the method and timing of application, and the type and amount of input applied,. In generating alternative production techniques, each aspect can be varied to create alternatives. In our definition, conservation management includes both soil conservation management and biodiversity and landscape management and is the broadest category in terms of management aspects (Table 3.3). Table 3.3 provides examples of management aspects and does not aim to be comprehensive. The management aspects of relevance to be included depend on the research question, data availability and on the models involved.

Table 3.3. Management practices with examples of relevant management aspects

Management practice	Management aspects
Water	+ Method of application: drip, furrow, sprinkler; + Level of application: amount or rule; + Timing: soil water threshold values, or fixed number per cropping season.
Nutrient	+ Level of application: full replacement of crop needs or input oriented approach; + Type of nutrient: inorganic fertilizer (e.g. ammonium sulphate, NPK fertilizers), organic manure (e.g. pig slurry, farm yard manure) or green manure (e.g. mustard, legume); + Method of application: broadcast or drilled; + Dose/timing: timing rule and number of applications, either all at once or in splits.
Weed, pest and disease	+ Chemical control dose rates, frequency and timing; + Mechanical control: implement, frequency and timing; + Prevention: prevention measures and timing.
Conservation	+ Tillage: implement, working depth and mixing + Biodiversity and landscape management: mowing regimes, buffer strips, tree strips, etc.
General	+ Sowing: sowing depth, implement, timing; + Harvesting: implement, timing, yield loss; + Clipping and pruning; + Field inspection.

Filters in the PTG are based on production orientations and their aim is to develop internally consistent management practices, i.e., high irrigation water input is not combined with a low nutrient input, as the yield-increasing effect of (costly) irrigation water would be off-set by low nutrient availability. Production orientations (e.g., integrated, highly innovative, conventional) are used as guidelines to assess whether management practices are internally consistent and reflect the normative character of activities.

The PTG develops alternative arable activities in four steps. First, variants of management practices for specific crops are generated on the basis of changes in management aspects. Second, vector multiplication (Section 3.3.2) is used to generate the possible set of production techniques on the basis of variants for each management practice. Third, these production techniques are filtered on the basis of production orientations. Fourth, vector multiplication constructs agricultural activities by combining rotations with the possible production techniques for each crop of the rotation.

The PTG uses an output-oriented approach to quantify inputs of production techniques. For example, based on a target yield level of 8 t wheat/ha the amount of nutrients is calculated that is needed to realise this yield. Such an estimated amount of nutrients using the output-oriented approach can be used to initialise cropping systems models to simulate the yield level that is attainable under given conditions. Such simulated yields can and will most likely differ from the target yield because of weather conditions occurring in the growing season.

Currently, the PTG is implemented for water and nutrient management. The crops for which alternative water and nutrient management have to be made, need to be specified. For water management, three predefined sets of water aspects (e.g. maximum number of applications, amount of water per application, time window in which irrigation is possible, soil plant available water threshold) have been specified: demand based irrigation, potential irrigation and user defined irrigation. Demand based irrigation is assumed to provide just the amount of water necessary to refill from a defined soil plant available water content (e.g. 80% for water-sensitive crops and 60% for less-water sensitive crops) to field capacity. Potential irrigation is assumed to provide ample water by specifying 20 irrigation events throughout the year and starting irrigation whenever plant available water content drops below 95%. User- defined irrigation implies that the user defines all four irrigation aspects based on his knowledge and the research question.

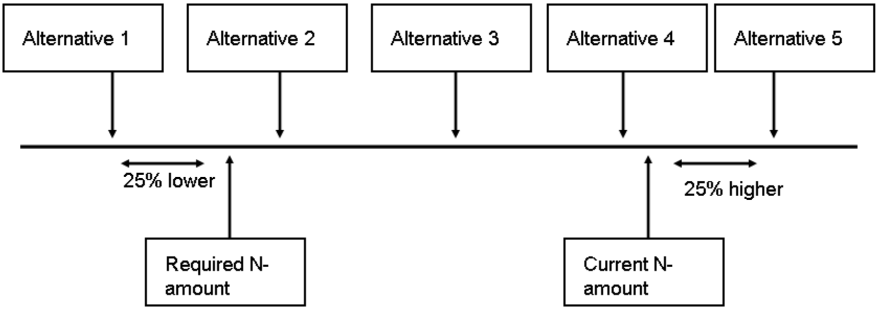


Figure 3.3. Schematic representation of nutrient fertiliser rates based on current and calculated N-amounts. In this example, the current nitrogen application is higher than that needed based on theoretical nitrogen requirements.

The current implementation of alternative nutrient management aims to create alternative fertiliser rates around the current fertiliser rate. First, the theoretical required nitrogen amount to achieve the current yield is calculated using a partial nitrogen balance approach and assuming a certain nitrogen use efficiency of the crop and indigenous soil nitrogen supply. Second, based on a user-defined percentage of variation (25, 33 or 40%), a range around the current and calculated N-amounts is calculated (Figure 3.3). Maximum and minimum rates are 25, 33 or 40% (user-defined) higher and lower, respectively, than the current and calculated N requirements. Third, on the basis of this range, 2 to maximum 5 equidistant N rates are calculated.

3.4.4. Technical Coefficient Generator (TCG)

An activity is characterised by a set of coefficients (Technical Coefficients (TCs) or input-output coefficients) that express the activity's contribution to the realisation of user defined goals (or objective in modelling terms) (Ten Berge et al., 2000). Technical Coefficient Generators (TCGs) (De Koning et al., 1995; Hengsdijk et al., 1999; Ten Berge et al., 2000; Ruben and Van Ruijven, 2001; Hengsdijk and Van

Ittersum, 2002; Ponsioen et al., 2006) can then be defined as algorithms to translate data information into coefficients that represent the input and output coefficients for each discrete activity. The Technical Coefficient Generator (TCG) in AMM links the agronomic input and output coefficients generated by the PEG, PTG and the cropping system model to socio-economic inputs and outputs by simple calculations. The TCG quantifies other or remaining inputs of each crop in each agricultural activity, i.e. the inputs not simulated through the cropping system model. These inputs, for example, refer to all inputs associated with management operations not considered critical for the performance of crop activities (e.g. harvesting operations), and labour and machinery requirements associated with management operations. The TCG functions as a wrapping component between different models to translate inputs and outputs from one model to the other model, which is especially relevant in linking models.

3.4.5. Software Design

The PEG, PTG and TCG have been implemented in Java^{tm} programming language and designed as components to facilitate replacement by other components. Objectives during software development were to separate algorithms, data and user interface to facilitate linkage to other databases and user interfaces, to modify and expand algorithms easily, and to increase the transparency and comprehensibility of the software. Software design patterns (Metsker, 2002) were used to allow for easy extensibility of PEG and PTG with new filters. The abstract factory pattern (Metsker, 2002) was found to be especially relevant as it allows to register filters in a catalogue-like source file that and as it separates the source code file of the filter from the executing file and catalogue file. More information on the software design can be found on in Athanasiadis and Janssen (2008) and Wien et al. (2009).

3.5. Application

3.5.1. Study area and data

The aim of the applications is to show the impact of different assumptions (e.g. generation algorithms, filters and input data) on the number and type of activities included. The applications concern Flevoland, a province in the Netherlands. Flevoland has been reclaimed from the sea in the 1960’s and its young and fertile soils are very productive. It is very homogenous in terms of soil (e.g. average 20% clay, 45% sand and 35% silt) and climate (e.g. average min. and max. temperature 6.2 and 14.6°C; average annual precipitation 617mm). Still, Flevoland has been subdivided in five agro-environmental zones (i.e. unique combinations of soil and climate) based on a biophysical typology (Hazeu et al., 2009). One of these five agro-environmental zones covers 80% of the area of Flevoland. Flevoland is an annual cropping area in the Netherlands, with a focus on cash crops like onions, (seed) potatos, bulb production, carrots and sugar beets. For Flevoland, current activities were collected through a survey (Zander et al., 2009), which can be used as a basis to make alternative activities.

3.5.2. Application of PEG

The application of the PEG consists of two parts. First, a list of crops was compiled that could be applied with the crop suitability filters. Table 3.4 contains the list of crops, that resulted from the application of the crop suitability filters, although it was not fully possible to apply the crop suitability filters according to Table 3.1 as originally envisaged. Given the fertile, well drained soil and flat topography of Flevoland, the crop suitability filters related to soil and altitude properties did not exclude any infeasible crops. The two climate-related filters, e.g. temperature sum and maximum rainfall (Table 3.1), were difficult to apply due to a lack of data to determine crop thresholds (Van der Maden, 2007). The minimum required temperature sum and the tolerance for high rainfall for a successful growing season are only available for common crops like wheat, sugar beet and rice. For potentially new crops for Flevoland (e.g. sunflower, soybean, elephant grass) it was very difficult to find thresholds for the minimum required temperature sum and tolerance for high rainfall and no comprehensive list of parameters could be derived from literature. Therefore, an expert-based process was followed to examine the characteristics of these potentially new crops vis-à-vis the soils and climate in Flevoland. In this expert-based process it was concluded that sunflower (Table 3.4) may become a relevant crop in Flevoland in about 10 years, since the border for growing sunflower in Europe was moving further north (Van der Maden, 2007). The reliance on the expert-based process indicates shortcomings in the applicability of the conceptual approach adopted in the PEG.

For the second part of the PEG, the list of crops was run through the rotation generation algorithm, that included all filters from Table 3.2. Different sets of crops and maximum and minimum rotation length were considered to examine the sensitivity of the number of alternative rotations generated. Increasing the maximum rotation length and including fallow lead to many more rotations being generated, while decreasing the number of crops or increasing the minimum rotation length had much smaller effects. There are two causes of the extremely high number of alternatives with a higher maximum rotation length (e.g. 6 years) and more crops (e.g. 16 or 17 crops). The first cause is that no crops are excluded from all rotations at higher rotations lengths. For example, tulip, fibre flax and peas are excluded from rotations at a rotation length of 4 or 5, because there need to be at least 5.5 years between repetitions (Table 3.5). The second cause is that at higher rotation length many more permutations can be generated. At a rotation length of 4 years 174 (= 83.521) possibilities exist, while at a rotation length of 6 years 176 (=24.137.569) possibilities exist. Additional normative filters can be added to reduce the number of rotations and to become more case specific.

Table 3.4. Crops considered for Flevoland, with some examples of relevant properties for rotation generation (Van der Maden, 2007).

Crop	Sowing date (julian day nr.)	Harvest date (julian day nr.)	Growth Period (days)	Min inter-crop period (days)	Max frequency (# yr^{-1})	Min period before repetition (yr)
Sugar beet	91	278	187	14	0.25	3.5
Potato	105	263	158	10	0.25	3.5
Winter wheat	-72[2]	222	294	5	0.50	1.5
Onion	84	258	174	5	0.20	4.5
Forage maize	115	274	159	4	0.50	1.5
Chicory	130	293	163	14	0.25	3.5
Spring wheat	74	227	153	5	0.50	1.5
Spring barley	74	222	148	5	0.50	1.5
Carrot	130	293	163	14	0.25	3.5
Tulip	-72[2]	182	254	10	0.16	5.5
Pea	74	196	122	5	0.16	5.5
Grass seed	-91[2]	213	304	5	1.00	0.5
Fibre hemp	121	274	153	5	1.00	0.5
Fibre flax	91	213	122	5	0.16	5.5
Sunflower	105	274	169	5	0.25	3.5
Winter oil seed rape	-131[2]	217	348	5	0.20	4.5
Fallow[1]	121	227	106	1	1.000	0.5

[1] Fallow represents in this application a non-productive cover crop, that is planted due to a policy obligation or to maintain soil structure and fertility.

[2] A negative number indicates a wintercrop that is sown in the year before the year in which it is harvested.

3.5.3. Application of PTG

In the application of the PTG the current activities specified in a survey (Zander et al., 2009) were used as a starting point to generate alternative activities. These activities contained the crops potato, winter wheat, spring wheat, sugar beet, maize and onion and concerned 7 rotations. Combining these 7 rotations with the five unique soil-climate combinations in Flevoland leads to 35 current activities. The management of the crops in these 35 activities can be varied through the PTG to create alternative activities.

Table 3.5. Number of alternative rotations based on different sets of crops and rotation lengths.

Number of crops[1]	Include fallow (Yes or No)	Max. rotation length (years)	Min. rotation length (years)	Number of alternative rotations
17	Y	6	1	218,665
16	N	6	1	93,615
17	Y	5	1	9,648
16	N	5	1	3,495
17	Y	4	1	994
16	N	4	1	411
10	Y	5	1	2,325
9	N	5	1	537
10	Y	5	3	2,285
9	N	5	3	525

[1] Crops included are: 17 = all crops from Table 3.5; 16 = all crops from Table 3.5 excl. fallow; 10 = spring wheat, potato, winter wheat, sugar beet, spring barley, carrot, maize, grass, chicory and fallow; 9 = spring wheat, potato, winter wheat, sugar beet, spring barley, carrot, maize, grass and chicory.

Four parameters of management were varied to investigate the number of alternatives generated. First parameter is that a new water management for maize was specified, which is the predefined option of demand-based irrigation. In the current activities, none of the crops were irrigated. Second parameter is the number of fertilizer rates included, which were either 3 or 4 (Fig. 3.3). The range for fertilizer rate is chosen as 33%. Third parameter entails varying the number of crops for which alternative nutrient management is calculated between either four (i.e. potato, sugar beet, maize and onion) or all six crops.

The fourth parameter concerns the assumptions of combining different alternative managements available for each crop. Under one assumption, alternative managements were combined according to vector multiplication (Section 3.3.2). Under another assumption the concept production orientation is used implying that intensive management of a crop can only be combined with intensive management of another crop, but not with extensive management of that crop. For example, in a two year maize-wheat rotation, three alternative nutrient fertilization levels (e.g. high, medium, low) are specified for each crop. Three alternative activities can be made, e.g. maize-high fertilization with wheat-high fertilization, maize-medium fertilization with wheat-medium fertilization and maize-low fertilization with wheat low-fertilization.

Table 3.6. Number of alternatives generated for different parameter settings of PTG component based on 35 current activities.

Assumption for combining managements alternatives[1]	Number of crops for alternative nutrient management	Number of fertilizer rates	Irrigated maize (Yes/No)	Number of alternatives
1	6	4	Y	175
1	6	4	N	175
1	6	3	Y	140
1	6	3	N	140
1	4	4	Y	175
1	4	4	N	175
1	4	3	Y	140
1	4	3	N	140
2	6	4	Y	5,747
2	6	4	N	5,747
2	6	3	Y	2,636
2	6	3	N	2,636
2	4	4	Y	1,751
2	4	4	N	1,751
2	4	3	Y	1,100
2	4	3	N	1,100

[1] 1= fixed combinations of managements; 2 = vector multiplication

On top of these parameter changes, it is assumed that alternative managements of a crop are not combined with current management of other crops, except if no alternative managements for a crop have been made. Also, for maize it is assumed that alternative water management and alternative nutrient managements are exclusively combined to form a production technique, if alternative water management for maize is available. This assumption implies that for maize alternative water management is not combined with current nutrient management and current water management is not combined with alternative nutrient management, if alternative water management is available.

Table 3.7. Current and alternative fertilizers rates as calculated by the PTG

	Current fertilizer rate (kg nitrogen /ha)	Theoretical required fertilizer rate (kg nitrogen /ha)	Number of alternative fertilizer rates	Calculated fertilizer rates in PTG (kg nitrogen/ha)
Potato	125	98	4	166 141 115 90
Potato	125	98	3	166 141 115
Spring wheat	140	177	4	236 201 165 129
Spring wheat	140	177	3	236 201 165

Table 3.6 shows the number of alternative activities generated while varying the values of the 4 parameters of the PTG. Due to the assumptions to restrict combinations of alternative water management and nutrient for maize, adding irrigation management of maize does not lead to more alternatives, as either all maize alternatives are irrigated or all maize alternatives are not integrated. If these assumptions are less restrictive (e.g. alternative water management of maize can be combined with current nutrient management), then more alternative activities will be generated. Changing the assumption for generating alternatives has a strong impact on the number of alternatives generated. If only fixed combinations of managements in a rotation are allowed, then only 5 to 6 times as many alternative (i.e. 140 or 175) as current activities (i.e. 35) are defined. If alternatives are made using vector

multiplication, then 31 to 164 times as many alternatives (i.e. 1100 to 5747) as current activities (i.e. 35) are defined. The PTG generates alternative fertilizer rates (fig. 3.3) based on the current fertilizer rate and a theoretical required fertilizer rate to fulfill crop needs, as can be seen in Table 3.7.

Table 3.8. Three examples of alternative activities made through PEG and PTG.

	Year 1	Year 2	Year 3	Year 4
Activity 1	Winter wheat	Potato	Sugar beet	Maize
Sowing week (week nr.)	-9[1]	11	30	17
Harvest week (week nr.)	33	40	41	40
Fertilizer rate (kg N/ha)	301	305	235	214
Number of irrigations	0	0	0	3
Annual irrigated amount (mm/ha)	0	0	0	33
Variable costs (€/ha)	715	1808	1294	1175
Activity 2	Winter wheat	Potato	Sugar beet	Maize
Sowing week (week nr.)	-9	11	30	17
Harvest week (week nr.)	33	40	41	40
Fertilizer rate (kg N/ha)	301	373	304	259
Number of irrigations	0	0	0	0
Annual irrigated amount (mm/ha)	0	0	0	0
Variable costs (€/ha)	715	1849	1352	1168
Activity 3	Winter wheat	Maize	Spring wheat	
Sowing week (week nr.)	-9[1]	17	11	
Harvest week (week nr.)	33	40	35	
Fertilizer rate (kg N/ha)	262	214	165	
Number of irrigations	0	3	0	
Annual irrigated amount (mm/ha)	0	33	0	
Variable costs (€/ha)	669	1175	526	

[1] A negative number indicates a wintercrop that is sown in the year before the year in which it is harvested.

Ultimately, the joint application of the PEG and PTG leads to fully quantified alternative agricultural activities (Table 3.8) that can be simulated in a cropping system model (e.g. APES (Donatelli et al., 2009), CropSyst (Stöckle et al., 2003) or APSIM (Keating et al., 2003)) or a simple budgeting approaches (e.g. Hengsdijk and Van Ittersum, 2003a; Dogliotti et al., 2004) to calculate desired outputs (e.g. crop yields, nitrate leaching, run-off or erosion). Some examples of alternative activities can be found in Table 3.8.

3.6. Discussion and Conclusions

3.6.1. Challenge 1. Explicit conceptual framework

In this chapter we propose a theoretical framework for generating alternative activities (Section 3.3) and we demonstrate its use in an implementation in the AMM for arable farming systems. The theoretical framework consists of an iterative cycle of generating permutations and filtering these through heuristic filters. The theoretical framework combines the agro-ecological engineering approach proposed by Van Ittersum and Rabbinge (1997) and Hengsdijk and Van Ittersum (2003b) with the approach to generate feasible alternatives as proposed by Wossink et al. (1992), Dogliotti et al. (2003; 2004) and Bachinger and Zander (2007). Although this combination itself is not new, the theoretical framework proposed here makes it more generic by an explicit distinction between production enterprises, production techniques, management practices and management aspects and by explicitly separating steps of generating and filtering, in which agro-ecological principles can be used. In principle, our procedure can be supplemented with a third or fourth cycle of generating and filtering of alternatives, for example, to allow for alternative financing strategies of new production enterprises or production techniques.
The theoretical framework is applicable irrespective of the research question or location, as it is abstract and sets general principles. The AMM currently considers only arable activities and focuses on cropping system models and bio-economic farm models. Similarly, components can be developed for livestock and perennial cropping systems and for other types of models (e.g. partial or general equilibrium optimization models). The theoretical framework is applicable to different types of farming (e.g. arable, livestock or perennial), as also in livestock or perennial cropping systems a distinction can be made between production enterprises (e.g. the herd or tree structure) and production techniques (e.g. feeding strategies or weed,

pest and disease management). The different types of permutation generation (e.g. cyclic permutation and vector multiplication; Section 3.3.2) can also be used in specifying permutations of production enterprises or techniques for livestock or perennial cropping systems. Heuristic filers can be specified according to similar principles for livestock and perennial cropping systems as for arable cropping systems.

3.6.2. Challenge 2. Parameter values for new and innovative technologies

Our approach classifies agricultural activities in production enterprises and production techniques, which in turn can be classified in management practices and management aspects. This multi-level classification of agricultural activities helps to explicitly define assumptions for each management aspect and for the combination of management aspects into production techniques and production techniques with production enterprises (Section 3.5.3). Some new procedures to calculate parameter values have been suggested in this chapter to define alternative nutrient and water management, and these can be combined with or replaced by procedures from earlier research (Hengsdijk and Van Ittersum, 2002; Van de Ven et al., 2003).

Other aspects that still deserve more attention in the conceptualization of agricultural management are (i) preventive management, like green manures and cover crops, which are often ignored or only one possible preventive management action is taken into account; (ii) definition and quantification of conservation management, which serves other objectives than pure profit maximization; (iii) temporal interactions between the specification of different inputs of crops in one rotation, as supposedly a wheat crop preceded by a sugar beet crop will require more nitrogen as a wheat crop preceded by a grass-clover mixture.

3.6.3. Challenge 3. Comprehensive set of alternative activities

A comprehensive set of plausible alternative activities is formed through specifying heuristic filters. These require knowledge that may be location or research question specific. This knowledge takes two forms. First, it is a rule that has to be followed. Second, it is a parameter to set a threshold in the rule. For example, if the frequency of potato in a rotation is lower than 0.25, then the rotation is not feasible. The threshold parameter is 0.25 and the rule is constructed through the 'if-then' statement. The rules cannot easily be derived from statistics or public data-sets, and have to be inferred from scientific publications and publications in farm management handbooks. Constructing a good set of filters requires time to translate the knowledge into sensible rules. Clearly the more time is spent on identifying filters, the more filters can be found in the literature and the more alternatives can be excluded (Fig. 3.4). This does not necessarily mean that all filters are equally effective in reducing the number of alternatives, and presumably a trade-off exists between the time spent and the effectiveness of the filters.

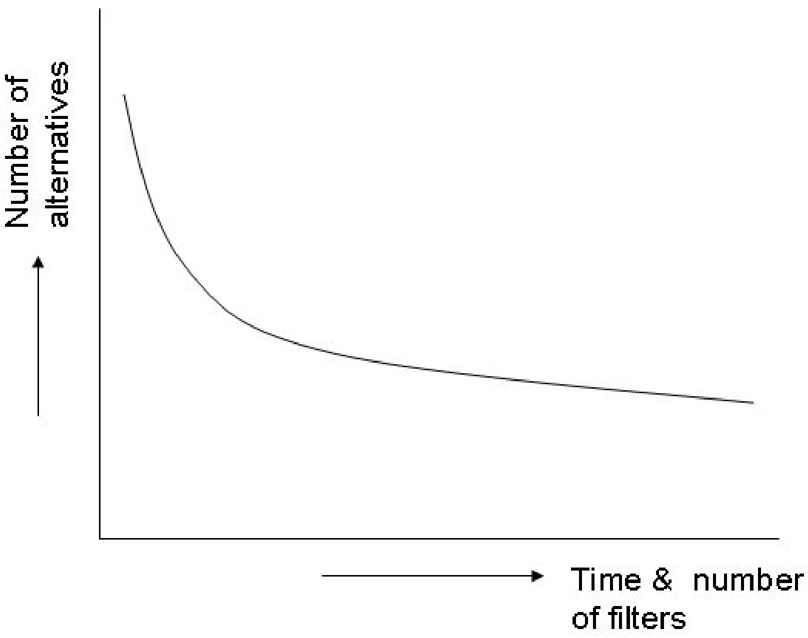


Figure 3.4. Trade-off between the number of alternatives and the time spent on formulating suitable filters.

The parameters to set thresholds in the rules can be difficult to find in the literature or from non- published data and might introduce some arbitrariness. For the crop suitability filters as part of the PEG it was difficult to identify suitable parameters for several innovative or currently less grown crops (Van der Maden, 2007). Crop suitability filters developed in the PEG (Table 3.1) could not be made operational, as we originally envisioned. Still, these crop suitability filters may function and their applicability may be tested on a large scale with crops for which data is available. A more promising route may be to converted the filters into gradual suitability ratings, that are not as discriminatory as the filters. The data limitations in parameterization of the crop suitability filters indicate that the selection of crops for a future-oriented study cannot easily be based on rigorous and generic procedures. All filters in the PEG and PTG have a heuristic nature and an implicit requirement of common sense. Future research must focus on compiling more robust set of filters, that are proven to perform in various environments.

3.6.4. Challenge 4. Managing large number of activities

Combinatorial explosions occur in generating alternative activities, as noted before by Wossink et al. (1992) and Dogliotti et al. (2003). Also in our applications these combinatorial explosions occur, when increasing the maximum rotation length (Section 3.5.2), incorporating variation in more management aspects (Section 3.5.3) and using vector multiplication to create alternative production techniques instead of fixed combinations of production techniques (Section 3.5.3). Another combinatorial explosion occurs if one combines the rotations generated by the PEG with the production techniques generated by the PTG. For example, combining the 218,665 rotations at a maximum rotation length of 6 with vector multiplication of 4 fertilizer rates leads to millions of alternative activities.

As proposed in this chapter and by Wossink et al. (1992), filters according to production orientations can help to reduce the number of activities. To substantially reduce the numbers quite strict filters reflecting strong assumptions are required. Examples of strict filters are leaving out management aspects (e.g. not specifying an irrigation method) or allowing only fixed combinations of managements (e.g. in the PTG application of Section 3.5.3). It depends on the purpose and research question of the future-oriented study, what constitutes a valid strict filter or assumption. Our theoretical framework helps to make these assumptions explicit and document them in the publication of the study. The combinatorial explosions represent a scientific challenge for subsequent future-oriented research to identify ways to manage them instead of ignoring them.

For the current implementation of the AMM the large numbers of alternatives are manageable. Generating 218,665 rotations takes less than 10 minutes in the PEG and making 5747 alternative activities in the PTG takes less than 10 seconds (on a computer with dual core 1.8GHZ processor and 2GB ram). ROTAT by Dogliotti et al. (2003) had a maximum limit of 250.000 rotations it could generate, which took several hours. Mono disciplinary simulation models might have more problems to analyse such large numbers of activities. A cropping system model like APES (Donatelli et al., 2009) currently runs a simulation of 25 years for one activity in about 30 seconds (on a computer with dual core 1.8GHZ processor and 2GB ram), requiring it 227 days to run all 218.665 rotations, while a bio-economic farm model like FSSIM (Louhichi et al., 2009) may not be able to keep a matrix with 218.665 rotations in memory. Super computing provide a solution to run cropping system models and bio-economic farm models. Alternatively, advanced and efficient algorithms to analyse large sets of agricultural activities may be more helpful. Such algorithms need to combine combinatorics to generate alternatives with, for example, optimization techniques according to Multi-Criteria approaches (Rehman and Romero, 1993) or heuristic search techniques (MengBo Li and Yost, 2000) to analyse alternatives.

3.6.5. Conclusion

The presented heuristic approach to systematically specify and identify relevant alternative activities is a necessity for future-oriented research. The process of identification of feasible and plausible alternative activities may be more valuable than the alternatives finally generated, as it (i) forces to be explicit about assumptions and deliberate omissions of activities, (ii) provokes discussion about assumptions and threshold values of filters and (iii) helps to identify knowledge gaps. Focus for future research must be on methods to analyse large number of alternatives on their merits as the current simulations models might not be up for the task. Hence, with a robust set of filters and tested generation and analysis algorithms available the explicit and coherent inclusion of alternative activities in future-oriented research can become a common practice instead of an exception.

Chapter 4. A generic bio economic farm model for environmental and economic assessment of agricultural systems

Abstract

Bio-economic farm models are tools to evaluate ex-post or to assess ex-ante the impact of policy and technology change on agriculture, economics and environment. Recently, various BEFMs have been developed, often for one purpose or location, but hardly any of these models are re-used later for other purposes or locations. These BEFMS stayed mostly in the research domain. The Farm System Simulator (FSSIM) aims to overcome this specificity by providing a generic framework enabling the application of BEFMs under various situations and for different purposes. This chapter introduces FSSIM and its components. FSSIM is a model that simulates farm responses to policy changes, technological innovations, societal or biophysical trends and it calculates agronomic, economic and environmental performance indicators of farming systems. Five criteria are introduced that generic BEFMs should satisfy, i.e. they should be useful for different (1) biophysical conditions and (2) farm types, (3) suitable for both technology and policy assessments, (4) allow the use of different levels of detail in input or output data and (5) the linking to other models at different scales. The generic nature of FSSIM is evaluated on the basis of these five criteria by examining various recent applications. The model is available for applications to other conditions and research issues, and it is open to be further tested and to be extended with new components, indicators or linkages to other models.

Sander Janssen, Kamel Louhichi, Argyris Kanellopoulos, Peter Zander, Guillermo Flichman, Huib Hengsdijk, Eelco Meuter, Erling Andersen, Hatem Belhouchette, Maria Blanco, Nina Borkowski, Thomas Heckelei, Martin Hecker, Hongtao Li, Alfons Oude Lansink, Grete Stokstad, Peter Thorne, Herman van Keulen and Martin K. van Ittersum, A generic bio economic farm model for environmental and economic assessment of agricultural systems submitted to Environmental Management

4.1. Introduction

Agriculture uses more than 40% of the land in the European Union (EU) and agricultural activities have a great impact on the environment and countryside through resource use, labor demand, environmental externalities and landscape layout. Farmers in the EU are under increasing pressure to consider the economic outputs of their activities, but also the environmental and social outcomes, as stipulated in European Commission policy documents, such as the Nitrates Directive (EC, 1991; EC, 2002) and the Water Framework Directive (EC, 2000; EC, 2007). Bio-economic farm models have been frequently proposed by research as tool to assess the impact of agricultural emissions on the environment (Vatn et al., 1997; Falconer and Hodge, 2001; Wossink et al., 2001) and of agriculture on landscape and biodiversity (Meyer-Aurich et al., 1998; Oglethorpe and Sanderson, 1999; Schuler and Kachele, 2003). Bio-economic farm models have also been proposed to assess the performance of different farming systems (De Buck et al., 1999; Berentsen, 2003; Pacini, 2003) or to evaluate the Common Agricultural Policy of the EU (Donaldson et al., 1995; Topp and Mitchell, 2003; Onate et al., 2007). Here a Bio-Economic Farm Model (BEFM) is defined as a model links farms' resource management decisions to current and alternative production possibilities describing input-output relationships and associated externalities. BEFMs can be useful to evaluate ex-post or to assess ex-ante the impact of policy and technology change on agriculture and environment (Chapter 2). In our review on the usefulness of BEFMs (Chapter 2), we identified a lack of re-use of these BEFMs, i.e. most models are used for the specific purpose and location only. They also largely stayed in the research domain and are not used for policy assessment. Applications of the same model for other purposes or locations are rare. An exception is the German model MODAM that has been applied during the last decade in different German and a number of European regions (Meyer-Aurich et al., 1998; Zander and Kächele, 1999; Kachele and Dabbert, 2002; Uthes et al., 2008). Another exception is the MIDAS model (Morrison et al., 1986; Kingwell and Pannell, 1987) that has been repeatedly used through the last decennia on sheep-arable farms in South-West Australia (Kingwell et al., 1995; Gibson et al., 2008; Kopke et al., 2008). In contrast, the re-use of cropping system models for diverse purposes and locations is far more wide-spread. For example, application of the Agricultural Production Systems sIMulator (APSIM) model has resulted in 102 publications (Keating et al., 2003). Also the CropSyst model (Stockle et al., 1994) has been applied for different crops and environments (Pala et al., 1996; Confalonieri and Bocchi, 2005; Wang et al., 2006). An example of an economic model that has been repeatedly used for different policy and trade questions is the Global Trade and Analysis Project (GTAP) model (Hertel, 1997).

To stimulate re-use with the option for new developments at each application, we propose to develop a generic BEFM that is suitable for many different applications. It is clear that required resources for development and maintenance as well as the level of abstraction will increase with more general applicability. Therefore, the question in reality will not be "generic or not", but rather relate to an optimal degree of being generic with some remaining restrictions on applicability. Still we believe that for scientific progress the challenge is to understand and model the "generic" processes, i.e. to identify and model those processes relevant to many purposes, research questions, locations and scales. Trying to shift the balance from the current emphasis on specific BEFMs to more generic BEFMs seems correct from a scientific and efficient from an application point of view.

In our view, there are several advantages of a generic BEFM, with one common and accepted concept and implementation achieved by a community of scientists. First, applications of BEFMs are easily repeatable and reproducible by a larger community, which makes consistent and large scale applications to a great diversity of agricultural systems possible. Second, a generic model could facilitate interdisciplinary research, as research groups can cooperate more efficiently. It allows to focus on innovations and extensions in science instead of each time "inventing the wheel" for each application, which saves time and resources. Synergies in building the model across research groups may occur, each bringing their own specialization and features to the model. Third, a generic BEFM makes peer review easier and more transparent as referees are more likely to be familiar with the common concept of the model. Fourth, it is easier to communicate with stakeholders (e.g. end-users and researchers in other domains) about the model and to achieve stakeholder acceptance of and confidence in the model results, when only one generic concept and model needs to be explained instead of explaining a new model with every application. Fifth, the extensive data requirements of BEFMs can be standardized and managed efficiently (Chapter 5).

There may also be disadvantages of a generic model. First, it may be more difficult to maintain an overview of the model, as new features and extensions are added over time and are developed by somebody else. It will become necessary to invest in maintenance instead of repeated development. Manuals and peer reviewed publications are required for adequate documentation and accessibility. Second, the level of detail of processes modelled or data used in a generic model may not be appropriate for a specific application. A generic model might be less suited less than a specifically developed model for a research question. Third, there are risks related to the implementation in source code, i.e. lock in effects, path dependency and legacy code. Lock in effects mean that inferior programming solutions are kept, while superior solutions exist. Path dependency means that potential progress depends on the path being followed, while alternative paths exist that yield more progress. Legacy code (Feathers, 2004) is a working source code for a purpose with assumptions on its use, that is subsequently used for other purposes under different assumptions. Tests and documentation are unavailable for these new purposes and different assumptions, which makes the source code difficult or impossible to maintain, improve or use. These risks of lock in effects, path dependency and legacy code can be mitigated by initially developing the model for a range of purposes, with a clear description of assumptions made, by using version management with a description of changes between versions and by adopting a software architecture that supports replacement and extension of components without affecting the other components.

The Farm System Simulator (FSSIM) has been developed as a generic BEFM. The aim of this chapter is to define a set of criteria for a generic BEFM, to introduce FSSIM, to describe its components and to demonstrate its generic features through describing different applications. Finally, the chapter discusses whether FSSIM satisfies the criteria for a generic model. These criteria are elaborated in the next Section, where also the philosophy underlying a generic model is introduced. Section 4.3 presents the underlying concept and some specific features of FSSIM and Section 4.4 describes the components of FSSIM in more detail. The technical implementation of FSSIM is presented in Section 4.5. Section 4.6 describes applications of FSSIM in relation to the criteria for generic models. Finally, Section 4.7 discusses whether FSSIM meets the criteria to be characterised as a generic model, and provides more information on the availability, maintenance and extension of FSSIM.

4.2. Criteria for a generic BEFM

Several criteria have to be met in our view to by a BEFM to be classified as generic. The first criterion is that it should be relevant for a range of agri-environmental zones. An agri-environmental zone is a homogenous combination of soil and climate types, that covers parts or whole administrative regeions. An example of such a definition of agri-environmental for the European Union can be found in Hazeu et al. (2009). These biophysical conditions strongly affect the current farm structure, the farming possibilities and potential in a location and thus the specification of a BEFM. For example, for a highland area with only grazing a different configuration of the BEFM is required than for a lowland fertile area.

The second criterion is that the BEFM should be applicable to a range of farm types, for example, arable, livestock and mixed farming systems, and low and high intensity systems as defined by a farm typology. Different farm types can be identified on the basis of a farm typology (Andersen et al., 2007a). The BEFM should have the capability to handle these different farm types consistently and without bias.

The third criterion is the capability to address different purposes, e.g. assessments of technological innovations or policy issues (Chapter 2). Assessments of policy issues have usually a short-term horizon and require realistic and validated modelling of farm responses, while assessments of technological innovations are explorative and often based on postulated optimizing farm responses (Flichman and Jacquet, 2003).

The fourth criterion is the capability to handle applications at different levels of detail in input or output data. Data availability will differ with the application, scale and location. For example, for regional studies often other type of data compared to an application at national or continental level. Moreover, accurate assessment of some indicators (e.g. landscape, biodiversity and greenhouse gas emissions) requires more detailed data on agricultural activities and their effects than that of other indicators (e.g. farmer income, total costs, revenues). The data requirements depend also on the purpose of the application and perspective of the researcher.

Finally, the fifth criterion is that the model should be capable to link to different types of models simulating processes at different scales. Linking could be required to assess the impact of simulated land use changes on markets, bio-physical, structural or aesthetical parameters of landscapes, and on emissions to water and air. For example, the up-scaling of farm responses to market level models is relevant for assessments of high level policies, while for assessments of biodiversity and landscape impacts of farming linking to landscape models is relevant. The BEFM should not be constrained in its linking to one specific type of model, but instead be capable to exchange input and output data with each of these model types in a flexible way.

4.3. Overview of FSSIM

FSSIM has been developed as part of SEAMLESS (System for Environmental and Agricultural Modelling; Linking European Science and Society), which was an

Integrated Assessment and Modelling research project (Van Ittersum et al., 2008) that developed a computerized framework to assess the impact of policies on the sustainability of agricultural systems in the EU at multiple scales. This aim is achieved by linking models across scales, disciplines and methodologies (Van Ittersum et al., 2008), and combining these models with qualitative judgements and experiences (Ewert et al., 2009).

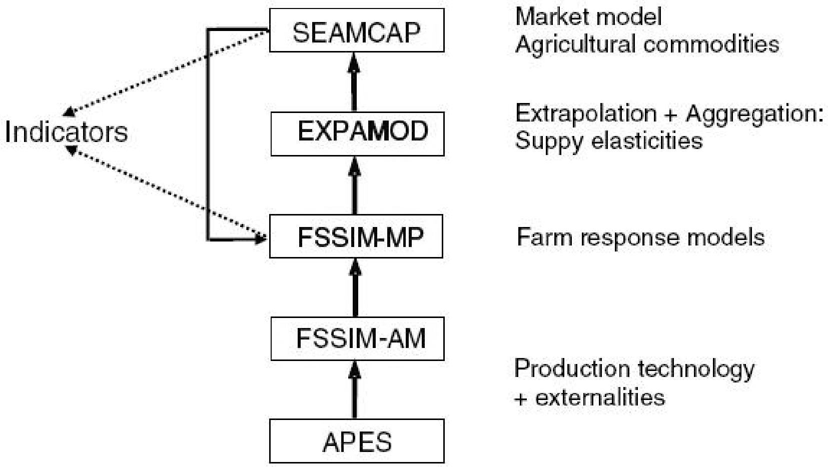


Figure 4.1. The model chain[a] in SEAMLESS (Van Ittersum et al., 2008).
[a]APES: Agricultural Production and Externalities Simulator; FSSIM-AM: Farm System SIMulator-Agricultural Management; FSSIM-MP: FSSIM-Mathematical Programming; EXPAMOD: EXtraPolation and Aggregation MODel and SEAMCAP: SEAMLESS version of the Common Agricultural Policy Regional Impact Analysis model.

Conceptually, FSSIM serves two main purposes. The first purpose is to provide supply-response functions for so-called NUTS2-regions (EC, 2008b) that can be upscaled to EU level. NUTS stands for Nomenclature of Territorial Units for Statistics and the second level corresponds to provinces in most countries. For this purpose, FSSIM is linked to an econometric extrapolation model (EXPAMOD), as its aggregate behaviour is needed as input to a partial equilibrium market model (Fig. 4.1). The second purpose is to enable detailed regional integrated assessments of agricultural and environmental policies and technological innovations on farming practices and sustainability of the different farming systems. For this purpose, FSSIM is linked to a cropping system model (APES) to quantify agricultural activities in terms of production and environmental externalities (Fig 4.1.) The consequence of this dual purpose of FSSIM is that some of its applications are more data intensive than other applications.

BEFMS are usually based on mathematical programming (MP) techniques. In MP the farm is represented as a linear combination of farm activities. The concept of activity is specific to mathematical programming and incorporates the idea of "a way of doing things" (Dorfman et al., 1958). An activity is a coherent set of operations with inputs resulting in the delivery of corresponding marketable products or products for on-farm use and externalities, e.g. nitrate leaching, pesticide run-off and biodiversity (Ten Berge et al., 2000). An activity is characterised by a set of technical coefficients (TCs, or input-output coefficients) expressing the activity's contribution to the realisation of defined goals or objectives in modelling terms (Hengsdijk and Van Ittersum, 2003a). Constraints are included to express farm level minimum or maximum quantities of input use or output marketing restrictions. Optimal activity levels are obtained by maximising an objective function reflecting user-specified goals, for example profit maximization, subject to the set of constraints (Hazell and Norton, 1986). Standard mathematical formulations of MP models can be found in Hazell and Norton (1986). FSSIM consists of two main components, FSSIM-Mathematical Programming (MP) and FSSIM-Agricultural Management (AM) (Fig. 4.2). FSSIM-AM comprises the activities in the BEFM, while FSSIM-MP describes the available resources, socio-economic and policy constraints and the farm's major objectives (Louhichi et al., 2009b). Both components are jointly configured to simulate a mathematical problem of resource allocation depending on the farm type, agri-environmental zones, research question and data availability.

The aim of FSSIM-AM is to describe current activities, generate alternative activities and quantify the activities through all the required technical coefficients. Alternative activities are new activities or activities currently not widely practiced in

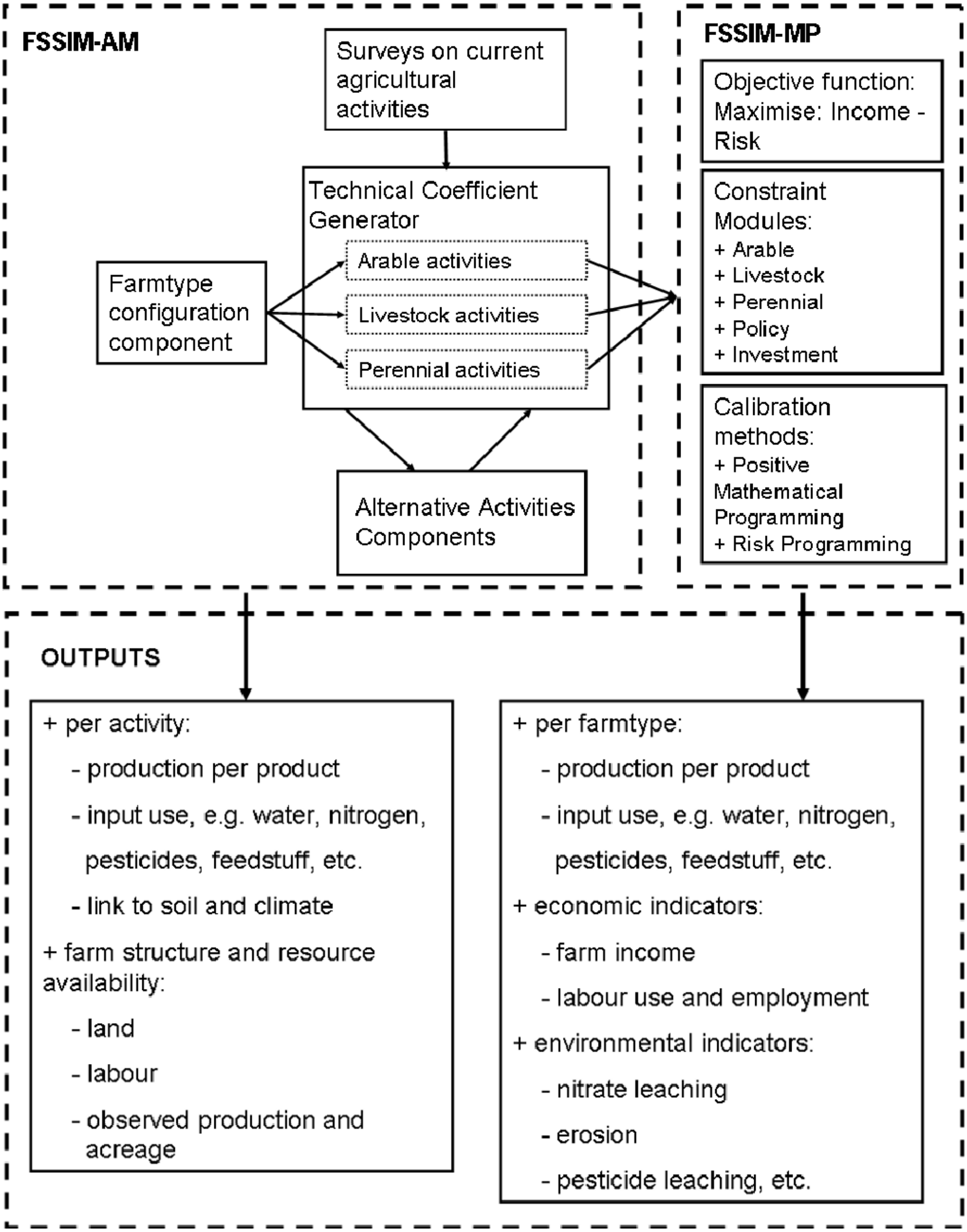


Figure 4.2. The structure of FSSIM[a] and its outputs.

[a] FSSIM consists of two main parts, i.e. the FSSIM-Agricultural Management (AM) component representing activities of the BEFM and the FSSIM-Mathematical Programming (MP) component representing the objective function and constraints of the BEFM.

the study area, and include technological innovations and newly developed cropping or husbandry practices (Van Ittersum and Rabbinge, 1997; Hengsdijk and Van Ittersum, 2002). Based on the farm typology, the Technical Coefficient Generator (TCG) quantifies inputs and outputs for arable, livestock or perennial activities or combinations of activities (Section 4.4.2.3/4.4.2.4). These activities can be simulated by a cropping system model such as the Agricultural Production and Externality Simulator (APES; Donatelli, et al., 2009) in terms of production and environmental effects. The quantified activities in terms of inputs and outputs are assessed in FSSIM-MP with respect to their contribution to the farms and policy goals considered (Fig. 4.2)

The outputs of FSSIM at farm scale are allocated areas with crop, grassland and perennial activities, or numbers of animals with livestock activities depending on the farm type considered. On the basis of optimal activity levels, different types of indicators can be calculated such as economic indicators for income, gross production and the share of subsidy in income, and environmental indicators for nitrate and pesticide leaching and erosion. Currently, over 100 indicators have been specified at both activity and farm level (Alkan Olsson et al., 2009).

In order to perform with/without assessment of technological innovations, policies or societal trends, a base year, baseline and one or more counterfactual experiments have to be specified for simulating a research question with FSSIM. Historic production patterns (e.g. land use and animal levels) of the base year are used to calibrate the model, e.g. ensuring that observed production patterns can be reproduced. Different calibration procedures have been incorporated (Kanellopoulos et al., 2009). Subsequently, a future baseline experiment is run using accepted and implemented policies. Results of this baseline experiment are used as benchmark for results of counterfactual experiments with the same time horizons. By using such calibration procedures and experimental set up, the overall aim of FSSIM is achieved, which is to simulate the actual farm responses through realistic and validated (e.g. positive) modelling (Flichman and Jacquet, 2003).

4.4. Components of FSSIM

4.4.1. FSSIM Mathematical Programming

FSSIM-MP (Louhichi et al., 2009b) is a model maximising a farm's utility function subject to a number of resource and policy constraints. The model can be characterised as a static positive, risk programming approach. A positive model means that its empirical applications simulate realistically the observed behaviour of economic agents. A static model does not include a time step in the model. Although the model is static, the input and output coefficients of the agricultural activities take temporal interactions into account as "crop rotations" and "dressed animal" instead of individual crops or animals (Hengsdijk and Van Ittersum, 2003b). The risk programming is the Mean-Standard deviation method in which expected utility is defined as expected income and standard deviation of income (Freund, 1956; Hazell and Norton, 1986).

Table 4.1. Components and their role in FSSIM-MP.

Title component	Role and functionality
Integrative Component	Solve the components together and manage model execution
Annual crops	Resource constraints (e.g. land, labour, irrigable land), production, revenues and income from arable activities
Livestock	Resource constraints (e.g. feed availability and requirement, stable size, concentrate purchases, labour), production, revenues and income from livestock activities
Perennial activities	Resource constraints (e.g. replacement and investment, land, labour, irrigable land), production, revenues and income from perennial activities
Policy	Price and market support-policies, set-aside schemes, quota schemes, production and income support policies, tax and penalties, cross-compliance and agri-environmental measures
PMP	Different PMP variants for exact calibration
Risk	Risk as aversion from yield and price variation
Trend	Yield and prices trends between base year and baseline experiment

FSSIM-MP consists of components (i.e. blocks of equations) that capture the agricultural activities (e.g. arable, livestock, perennials) and components for inclusion of alternative policies, calibration procedures (Positive Mathematical Programming (PMP)), risk and trend analysis (Table 4.1). The principal policy instruments implemented in FSSIM-MP (Louhichi et al., 2009b) are price and market policies, set-aside schemes, quota schemes, production and income support policies, taxes and levies, cross-compliance and agri-environmental measures (Table 4.1). Policy instruments in FSSIM-MP are modelled either as part of the objective function (e.g. premiums as monetary incentives), or by including them as constraints (e.g. set-aside and quota schemes).

These components are solved simultaneously and they are managed by an integrative component (Table 4.1), containing the objective function and the common constraints. Thanks to its modularity, FSSIM-MP provides the capability to add and remove components (and their corresponding constraints) in accordance with the needs of the simulation experiment and to control the flow of data between the database and the software tools. FSSIM-MP has been programmed in the General Algebraic Modelling System (GAMS, 2008).

4.4.2. FSSIM Agricultural Management

4.4.2.1. Current agricultural management

A detailed knowledge of current agricultural management is required to reproduce production patterns in the base year and to assess the impact of short term policy changes, where farmers response is based on their current technologies. Current agricultural management serves as input for the definition of alternative activities. These current activities represent the inputs and outputs of actual farming practices for average weather conditions (Borkowski et al., 2007). Diversity in actual farming practices, and thus in inputs and outputs of activities is large. This diversity in activities can either be captured by average or typical current activities. Average activities represent the mean of activities carried out on a representative sample of farms, while typical activities are described on the basis of representative activities such as available in farm management handbooks or extension brochures. Information on current activities can be based on observed data or expert knowledge. In the SEAMLESS project, a lack of data and information on agricultural activities at European level was identified, especially with respect to non-economic data. For example, the Farm Accountancy Data Network (EC, 2008a) provides aggregate costs and aggregate input use for the whole farm but not specified per crop or animal type and without a temporal distribution. Therefore, two dedicated surveys were developed as part of the SEAMLESS project (Borkowski et al., 2007). A detailed survey carried out in five EU regions (Brandenburg, Andalucia, Midi-Pyrénées, Flevoland and Zachodniopomorskie) collected data for typical current arable activities on input quantities, timing of input use, crop rotations, machinery and labour use, and associated costs. The detailed survey was conducted by regional experts, who work regularly with farmers. A so called "simple survey" was conducted to collect a reduced data set in 16 EU regions for arable, livestock and perennial activities comprising economic variables (e.g. product costs and prices), yields, composition of rotations and some aggregate physical variables describing input use (e.g. nitrogen use per crop and total medicine costs per animal) (Borkowski et al., 2007). The simple survey does not contain information on detailed management variables, i.e. frequency and timing of input use. The regions were selected to represent the diversity of farm types in different bio-physical endowments across EU-25. The regions selected are administrative regions, but the information in the surveys is linked to different agro-management-zones (1-5 per region) within a region. The simple survey was conducted by scientists working in the region supported by statistical data and farm management handbooks.

4.4.2.2. Alternative Agricultural Management

Few BEFM applications include technically feasible alternative activities and if they are used they are based on expert judgment with the risk of missing out suitable alternatives (Hengsdijk and Van Ittersum, 2002; Dogliotti et al., 2003; Chapter 2). In

FSSIM two specific components are available generating systematically alternative crop rotations and crop management options. The Production Enterprise Generator (PEG) is a version of ROTAT (Dogliotti et al., 2003) that generates crop rotations based on best agronomic practices formalised in crop and rotation suitability criteria, for example the maximum frequency of specific crops in a rotation to avoid the build up of soil born diseases. The Production Technique Generator (PTG) generates alternative crop management for entire rotations based on user-defined rules for water, nutrient, conservation, weed, pest and disease management. For example, the amount of nitrogen fertilizer is based on expected crop requirements to realize current yields instead of the amount of nitrogen fertilizer in the simple survey (Section 4.4.2.1). The methods to generate alternative activities developed in the PEG and PTG may be extended for livestock and perennial activities.

4.4.2.3. Technical Coefficient Generator for Arable Activities

Technical coefficient generators (TCGs) (Hengsdijk and Van Ittersum, 2003b) are algorithms to process data and information into technical coefficients directly usable by a mathematical programming model (e.g. FSSIM-MP) and cropping system models (e.g. APES). The Current Activities component (CAC) of the TCG processes survey data into compatible inputs for FSSIM and links them to regional farm types, while calculating an average over several years for the observed cropping pattern, product price and yield variability for these farm types using data from the FADN-based farm typology (Andersen et al., 2007a). The Simple Management Translator (SMT) of the TCG processes simple survey data (Section 4.4.2.1) into sets of inputs required for running APES based on expert-based management rules (Oomen et al., 2009). In the SMT, the aggregated physical input use from the simple survey is converted into a number of crop management events characterized by amounts, timing rules, machinery usage and working depths. Expert crop-specific management rules have been developed for sowing, harvesting, tillage, nutrient and water management. For example, if the simple survey data indicates that 150 kg N/ha is applied in a wheat crop, the management rule determines that this amount is applied in three splits, i.e. 30% in the first split at beginning of tillering, 40% in the second at ear initiation and 30% of the total in the last split at development of the last leaf. When detailed crop management data is available, for example through the detailed survey (Section 4.4.2.1), the conversion of the simple management data through expert rules in the SMT is not needed.

4.4.2.4. Technical Coefficient Generator for Livestock Activities

The TCG also prepares quantified livestock activities for dressed animals (Thorne et al., 2008), i.e. a combination of a mother-animal and its replacement in the form of a number of young animals. The types of livestock considered are dairy and beef cattle, sheep and goats. TCG for livestock activities characterizes livestock activities in terms of energy, protein and fill units requirements (Jarrige et al., 1986) according to the French feed evaluation system (Jarrige, 1988; Jarrige, 1989; Beaumont et al., 2007). These energy, protein and fill unit requirements of livestock activities have to be met in FSSIM-MP with the energy, protein and fill units of the feed resources available at farm, such as grass fodder, grass-silage, hay and feed production on arable land (e.g. fodder maize and fodder beets). Energy, protein and fill unit contents of feed resources are either based on Jarrige (1988; 1989) or calculated according to static relationships with on the one hand grassland yields and associated nitrogen input levels and on the other hand energy, protein and fill unit contents (Thorne et al., 2008).

4.4.3. FSSIM Graphical User Interface

One of the features that could stimulate the use of generic BEFMs by a larger community and that benefits from the modular set up is an easy to use and accessible graphical user interface (GUI), which is specifically developed for FSSIM (Meuter et al., 2009). This FSSIM GUI is a user-friendly interface allowing users to initialize, run and modify data for simulations with FSSIM (Meuter et al., 2009). The functionality is primarily targeted at users with less experience in the use of BEFMs.. In the GUI, the user specifies model experiments to select and configure the components available in FSSIM, because usually not all available components are needed for a specific experiment. Depending on the selected components, the components are further configured on the basis of the study region, farm type, available agri-environmental zones and crops. In addition, parameter values for prices and policy instruments need to be set by the user. The FSSIM-GUI is web-based, which makes the application easily accessible for the research and user community and allows the application to keep track of its users. Outputs from the model experiments can be downloaded for further processing.

4.5. Technical design of FSSIM

An adequate technical design is required to achieve a conceptually generic model, that is relatively easy to use, maintain and extend. The technical design of FSSIM is based on the theory of software components, semantically aware components and multi-tiered application. The division of a model in software components supports the modularity of FSSIM in the conceptual components presented in Section 4.3 and 4.4 (Fig. 4.3). The components are made semantically aware. Semantically aware components use a common "dictionary" of shared data types to ensure meaningful, consistent and explicit exchange of information between FSSIM components. Finally, multi-tiered applications help to separate common operations such as data storage and access, visualization and execution of the model from the implementation of the model in source code, thereby allowing modelers to focus on model implementation (Evans, 2003; Knapen et al., 2007)(Fig. 4.3). The implementation based on these three theories, i.e. software components, semantically aware components and tiered applications, ensure that the FSSIM model can be divided into parts that can be developed, maintained and extended simultaneously with an adequate data-exchange between these parts.

Software components (Szyperski et al., 2002) means that a model (or program) can be dissected in distinct autonomous parts (e.g. a component) that communicates with other components in the model and provides services to other components or a model. For something to be called a software component, it must have a clearly defined interface, be able to communicate with other components, encapsulate its inner workings, be non-context specific and independently re-usable in other situations (Szyperski et al., 2002). FSSIM is divided into two main components, i.e. FSSIM MP and FSSIM AM which each are divided into smaller components, for example, the livestock component of FSSIM-AM generating livestock activities and the policy component of FSSIM-MP that models agricultural EU policies. This design allows to use, replace and improve FSSIM components independently facilitating model development and maintenance of the model by different modelers. The interfaces of FSSIM components, i.e. the inputs and outputs of a component are annotated and described explicitly in an ontology (Athanasiadis and Janssen, 2008). In computer science, an ontology is considered the specification of a conceptualization (Gruber, 1993). Such a conceptualization is expressed in a machine readable format, for example the Web Ontology Language (McGuinness and Van Harmelen, 2004). The use of an ontology facilitates clear definitions for loosely integrated models in an open software environment (Li et al., 2007; Rizzoli et al., 2008). The ontology with the component interfaces functions as a common dictionary and ensures consistent definitions of concepts and data types across components. The ontology helps to link internal FSSIM components and to link FSSIM to models from other domains. Component modelers have to interact to clarify the interfaces of each of the components.

The tiers in FSSIM consist of presentation tiers, a data tier, an application tier and domain tier. The presentation tier is the graphical user interface (GUI), which obtains user-input and presents the model results. Two different presentation tiers are linked to FSSIM, the SEAMLESS-Integrated Framework GUI and the FSSIM-GUI. The FSSIM-GUI (Section 4.4.3) is used to operate FSSIM as standalone model independent of other SEAMLESS models. In SEAMLESS-IF, FSSIM is integrated with other models and is run as part of a model chain managed by the SEAMLESS-IF GUI (Fig. 4.1). The FSSIM application tier manages the interaction between different tiers, especially the model execution from the presentation tier. FSSIM forms its own domain tier. The data tier handles data requests by the application tier or domain tier and communicates with the SEAMLESS database to retrieve this data. Finally, the domain tier consists of the components of FSSIM and offers to functionality of FSSIM to the other layers. Advantage of a tiered application is the separation of roles and modularity, as changes in one tier do not directly have to affect other tiers.

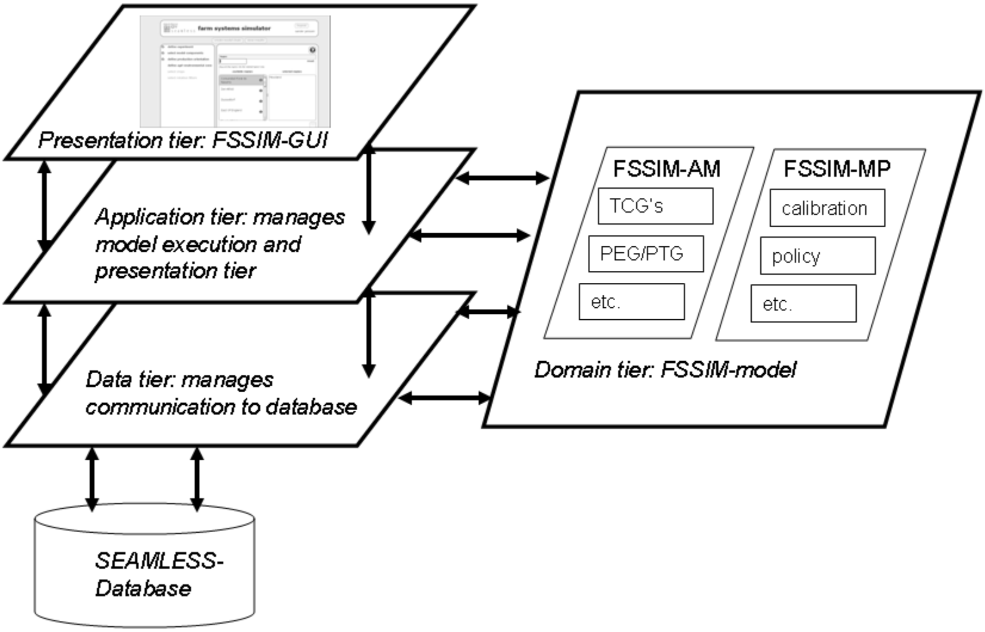


Figure 4.3. Technical design of FSSIM[a].

[a]The tiers presented in rectangles and the FSSIM component inside the domain tier. Arrows indicate information exchange among tiers.

4.6. Applications

FSSIM has been applied in a number of cases over the last years by different research groups for two purposes, i.e. micro-macro analysis (Section 4.6.1) and regional integrated assessment (Section 4.6.2).

4.6.1. Application for micro-macro analysis

FSSIM was used to provide input to supply-response functions at NUTS2 level that were upscaled to EU level. In this context, FSSIM was applied to 13 regions and 55 arable farm types throughout the EU to obtain values for price elasticities of different crop products. A price elasticity is the percentage change in supply as a results of one percent change in price. Table 4.2 provides an example of FSSIM results in the form of price elasticities for soft wheat in five regions. In Kanellopoulos, et al. (2009), a description of the application to two of these 13 regions can be found, i.e. Flevoland in the Netherlands and Midi-Pyrénées in France. FSSIM is used according to a standardized and automated procedure in each region. First, data are retrieved from FADN (EC, 2008a) and from the simple survey on agricultural management (Section 4.4.2.1) for each farm type in a region. Second, these data are processed in an automated way through FSSIM-AM to prepare the technical coefficients, e.g. specifications of relevant activities and farm and policy parameters. Subsequently FSSIM-MP optimizes the objective using the region-specific and farm-specific sets of activities and constraints with an automated calibration procedure (Kanellopoulos et al., 2009). By using a standardized and automated procedure, the application is repeatable and consistent over different farm types. Case-specific characteristics of farm types and regions beyond those implied by the standard data sources, technology generation differentiated by biophysical conditions and FADN based farm type resources could not be taken into account.

Table 4.2. Price elasticities for soft wheat for five different regions as derived from simulations by FSSIM (Kanellopoulos et al., 2009; Pérez Domínguez et al., 2009).

Region (land)	Price Elasticity for Soft wheat
Andalucía (Spain)	0.22
Midi-Pyrénées (France)	4.37
Poitou-Charentes (France)	2.36
Brandenburg (Germany)	0
Flevoland (Netherlands)	2.26

4.6.2. Applications for regional integrated assessment

Six detailed regional assessments have been done using FSSIM involving different farm types (e.g. arable and livestock), different scales (e.g. individual farm types, catchments and regions), different geographical locations (e.g. North, East, Western and Southern Europe, Africa) and using different components to estimate yields and environmental effects of activities (e.g. models and expert knowledge). In some of these applications, adjustments to FSSIM-AM or alternative procedures to estimate technical coefficients have been made dependent on the availability calibrated cropping system models and detailed data for regions or farm types.

In one application, FSSIM and CropSyst (Stöckle et al., 2003) were jointly applied to assess the impacts of the Nitrates Directive (EC, 1991) on three arable farm types in the French Midi-Pyrénées region (Louhichi et al., 2008). Table 4.3 provides values for the indicators farm income and nitrate leaching for the baseline and a counterfactual "Nitrates Directive" experiment as an example of results of a regional integrated assessment. Louhichi et al. (2008) also applied FSSIM-MP in four farm types in the Sikasso region (Mali) evaluating the impacts of improved cropping practices and introduction of organic cotton. Majewski et al. (2009) applied FSSIM to several arable farm types in the Zachodniopomorskie region in Poland to investigate the impacts on economic indicators and cropping pattern due to changes in farm quotas and the introduction of biofuels. In a catchment in Scotland, Mouratiadou, et al. (2009) used outputs of the process-based nitrogen simulation model NDICEA (Van der Burgt et al., 2006) in FSSIM-MP to assess impacts of EU's 2003 reform of the Common Agricultural Policy (EC, 2003) on economic and water quality indicators of two farm types. The application to livestock farming is an assessment of the consequences of an increase in milk quota and concentrate prices on dairy farms in Auvergne, France and in Flevoland, Netherlands (Louhichi et al., 2009a).

Table 4.3. Farm income and nitrate leaching of three farm types in the Midi-Pyrénées region in France[a].

Farm type	Farm income (k€/year)		Nitrate leaching (kg $N\text{-}NO_3^-$/ha)	
	Baseline	Nitrates directive[b]	Baseline	Nitrates directive[b]
Large scale-medium intensity-arable cereal	72	71	41	25
Large scale-medium intensity-arable fallow	77	76	36	36
Large scale-medium intensity-arable (others)	74	73	34	26

[a] A regional integrated assessment of the nitrate directive (adapted from Louhichi, et al. (2008))

[b] Experiment based on Nitrates directive (EC, 1991)

The impacts of alternative irrigation and nutrient management on crop allocation, farm income and environmental indicators is investigated using FSSIM-MP, FSSIM-AM and APES for Flevoland in the Netherlands, which is partly described in Chapter 3. In this application, a standardized and automated procedure processes data of arable activities from the simple survey and FADN to create inputs for the cropping system model APES of which the results are subsequently provided to FSSIM-MP. This procedure can be used for other regions and thus allows to combine applications on regional integrated assessments and micro-macro analysis.

4.7. Is FSSIM generic, usable and extensible?

The applications of FSSIM (Section 4.6) are evaluated using the criteria defined for generic BEFM introduced in Section 4.2 (Table 4.4). For criterion 1, FSSIM has been applied for different climate zones, e.g. Atlantic, Continental, Mediterranean, Lusitanian and Alpine and soil types e.g. sandy and clay soils (Criterion 1). FSSIM has been applied to a range of different farm types (criterion 2) with different specializations (e.g. specialised crops based arable, cereal-based arable, livestock and mixed farms), different intensities (e.g. extensive and intensive farms) and sizes (e.g. small sized farms in Mali and Zachodniopomorskie to large sized farms in Midi-Pyrénées). FSSIM needs to be extended to be able to simulate farm types with perennial, intensive horticulture and intensive livestock systems. Intensive livestock (e.g. pigs, poultry) and horticulture (e.g. greenhouse production) systems are characterized by capital intensive, often soil-less and high external input use activities and these systems are often not bound to land and labour resource constraints. FSSIM needs to be extended with constraints related to capital availability and an adapted definition of capital intensive activities for these farming systems. Extending FSSIM for perennial farming systems requires incorporating the temporal changes in perennial crops from a growing and established crop to a productive crop.

With respect to criterion 3, in most applications FSSIM has been used to assess the effects of policy changes and in two applications to assess the impact of technological innovations (Louhichi et al., 2008; Chapter 3). FSSIM may be applied in the future to assess the impacts of societal or physical trends, for example the effects of climate change and increases in energy prices on farm performance.

In the various applications, different data sources, level of detail (e.g. criterion 4) and model configurations have been used. In the application for micro-macro analysis the level of detail in data was lowest, as only regional data sources could be used that were standard available (Kanellopoulos et al., 2009). In the regional integrated assessment studies, more detailed data and specifications could be used, often by incorporating ad hoc procedures (Louhichi et al., 2008; Louhichi et al., 2009a; Majewski et al., 2009; Mouratiadou et al., 2009).

Criterion 5 required that a generic BEFM can be linked to different types of models. FSSIM has been linked to economic models (i.e. EXPAMOD and CAPRI) (Pérez Domínguez et al., 2009) for up-scaling of its supply responses, but also to an environmental externality simulation model (i.e. NDICEA) and different cropping system models (i.e. CropSyst and APES). A useful extension of the model linking is to link FSSIM to a landscape model, that allows to visualize or analyse the results of FSSIM at the landscape level.

Not all components of FSSIM have been used in each application (Table 4.4). In some applications, both FSSIM-AM and MP were used. In other applications only FSSIM-MP was used in combination with other models and methods then FSSIM-AM to quantify farm activities. Individual components of FSSIM may also be used as stand alone tool, for example, the integrative component, crop component and risk component of FSSIM MP (Table 4.1) to assess the response to changing risk on an arable farm or the livestock component of FSSIM-AM to calculate yearly feed requirements in terms of energy, protein and intake capacity of different farm animals. FSSIM needs to be configured depending on the data availability, research question and location. For example, to identify improved nitrogen fertilization techniques, FSSIM-AM components for both current and alternative activities need to be configured in order to allow for a realistic farm response by including all potentially relevant production activities. FSSIM-MP can be configured without the PMP-based calibration procedures and instead risk calibration procedures can be used, because the aim is to identify more optimal nitrogen fertilization techniques as the current practices and not to simulate in a realistic and validated farm responses. Alternatively, if the research question is to assess the short term effects of the abolishment of the EU set-aside policy, then the PMP calibration procedure of FSSIM-MP is required, but components for alternative activities in FSSIM-AM may not be needed as in the short term agricultural management is less likely to change significantly.

FSSIM is available for use and extension, for new purposes, locations and scales, either through its GUI or by working directly with the source code of the model. FSSIM will be maintained and extended during the next four years as part of the SEAMLESS association (www.seamlessassociation.org). FSSIM would benefit from extensions to model biodiversity, landscape and conservation indicators, from procedures for sensitivity and uncertainty analysis, from a more detailed modeling of the relationships between livestock density, grassland and manure production and from further calibration and validation of the model to new locations and research questions.

Table 4.4. Applications of FSSIM used to evaluate the generic nature of FSSIM according the criteria defined in Section 4.2.

Reference for application	Region	FSSIM-component	Purpose[a]	Climate	Soil type Classified carbon content	Farm types	Type of Assessment[b]	Level of detail + = low to +++ = high	Link to other models
Kanellopoulos et al., 2009	13 in Europe	AM and MP	1	13 across Europe	Very low to very high	Arable	2	+	EXPA-MOD CAPRI
Louhichi, et al. 2008, Midi-Pyrénées	Midi-Pyrénées, France	MP	2	Mediter-ranean Lusita-nian	Very low to very high	Arable	1 and 2	++	Crop-Syst
Louhichi, et al. 2008, Mali	Sikasso, Mali	MP	2	Tropical	Very low	Arable	1 and 2	++	Expert
Louhichi et al., 2009a	Flevoland (NL), Auvergne, (FR)	AM and MP	2	Atlantic	Low to very high	Live-stock	2	+	--
Majewski et al., 2009	Zachodnio-pomorskie (PL)	MP	2	Conti-nental	Very low to very high	Arable/ Live-stock	2	++	--
Mouratiadou et al., 2009	Catchment in Scotland, UK	MP	2	Atlantic	Moderate to very high	Arable	2	++	NDICEA
Chapter 3	Flevoland, Netherlands	AM and MP	2	Atlantic	High to very high	Arable	1	+++	APES

a: 1 = up-scaling of supply responses, 2 = regional integrated assessment; b: 1= technological innovation; 2 = policy changes

The conceptual and technical integration of the different FSSIM components has proved to be a challenging and time-consuming task due to the complex data-types (or data-structures) being exchanged between components, the large data amounts and diverse data sources required to run FSSIM. The required investment in conceptual and technical integration might be a barrier to the initial development and maintenance of a generic BEFM. The division of FSSIM in components and tiers (Section 4.5) was useful to separate and group functionality, without lumping all functionality in one monolithic piece of source code with data. Making these components semantically aware (e.g. annotating them in an ontology) helped to clarify the data types exchanged between components, to integrate the different data sources, to create data repositories to manage these data in an adequate way and to link FSSIM in a transparent and explicit way to other models. To integrate new components into FSSIM, the following explicit integration procedure is proposed, which already has been used to integrate the livestock parts of FSSIM:

1. Conceptual development, implementation and testing of stand-alone component;
2. Enter component interfaces (e.g. inputs and outputs of the model) in an ontology and link to other ontologies;
3. Enter and check data in database based on the ontology;
4. Develop and test the wrapper of the component with the rest of FSSIM. The wrapper acts between components to translate data from one programming paradigm into another;
5. Make the tested and integrated component available in FSSIM;
6. Apply the integrated component to more regions, locations and experiments with new datasets.

This integration procedure can now be used to extend FSSIM with new components, e.g. for perennial activities, multi-functionality or intensive livestock or horticultural systems. A technical barrier to the use of the FSSIM is the different programming paradigms used in components. Researchers are usually specialised in one programming paradigm. Training, simple user interfaces and documentation may help to overcome this barrier and generalists, who technically overlook FSSIM and its components, are required to maintain an overview.

In conclusion, FSSIM has been developed as a generic BEFM that targets wide applicability and models “generic” processes instead of specific processes to one research question, location or data source. FSSIM is a product from a joint development of agronomists and economists. This lead to a balanced definition between different types activities, policy instruments and technological innovations, without emphasizing any in particular, to enable use of FSSIM for different study objectives. FSSIM can be easily maintained and extended, as it comprises individual components that can be extended and maintained independent from other components. Although a truly generic model might not be possible, FSSIM represents a first step in the development of a BEFM as a library of components and functionality that can be adapted to the purpose, scale, location and linked to other models.

Chapter 5. A database for integrated assessment of European agricultural systems

Abstract

A major bottleneck for data-based policy making is that data sources are collected, managed, and distributed by different institutions, residing in different locations, resulting in conceptual and practical problems. The use of dispersed data for agricultural systems research requires the integration of data sources, which means to ensure consistency in data interpretations, units, spatial and temporal scales, to respect legal regulations of privacy, ownership and copyright, and to enable easy dissemination of data. This chapter describes the SEAMLESS integrated database on European agricultural systems. It contains data on cropping patterns, production, farm structural data, soil and climate conditions, current agricultural management and policy information. To arrive at one integrated database, a shared ontology was developed according to a collaborative process, which facilitates interdisciplinary research. The chapter details this process, which can be re-used in other research projects for integrating data sources.

Sander Janssen, Erling Andersen, Ioannis N. Athanasiadis and Martin K. van Ittersum, 2009. A database for integrated assessment of European agricultural systems. Environmental Science & Policy In press,

5.1. Introduction

5.1.1. Problem definition

Statistics and indicators based on data are essential to inform policy (Niemeijer, 2002; AbouZahr et al., 2007). Governments benefit from specialized statistical agencies for data collection, such as FAOSTAT (FAO, 2008), EUROSTAT (Eurostat, 2008) and national bureaus of statistics. Effectiveness and efficiency of policies can be evaluated through processing data on potential impacts, either after a policy is implemented (ex-post), or before a policy is implemented (ex-ante). For this purpose, different methodologies can be used, for example indicators and typologies derived from primary data, or indicators derived from quantitative modeling.

Indicators and typologies are means that can be used to process datasets to provide new insights. Both provide with summarized information about complex issues (Andersen et al., 2007a). Indicators synthesize relevant data and indicate the change or define the status of something (Gallopin, 1997), while a typology is a stratification of data that is homogeneous according to specific criteria relevant to policy, such as environmental and economic performance (Andersen et al., 2007a). Relying directly on available data, indicators and typologies may be used to (a) identify or justify needs for policy intervention, and (b) assess ex-post the impact of previous and current policies. Indicators are established for achieving both uses, as for example in the IRENA initiative on agri-environmental indicators (EEA, 2005), and in the assessment of the impact of the rural development programs of the European Union (EC, 2006). These uses are less acknowledged in relation to typologies. Recently Andersen et al. (2007b) argued that the criteria for a European farm typology may influence the assessment of policy changes.

Another technique to process data and inform policy making is Integrated Assessment and Modelling (IAM), which is used to assess the impacts of policies, technologies or societal trends on the environmental, economic and social sustainability of a system (Parker et al., 2002). IAM is a methodology that combines quantitative models representing different aspects of sub-systems and scales into an overall framework for Integrated Assessment (Parker et al., 2002). Quantitative models used in an IAM study originate from a different discipline, operate on different spatial and temporal scales, and require diverse (and sometimes, overlapping) data-sources. Model integration within an IAM project requires that all input and output data of each model have to be integrated. Prominent examples of IAM relate to the assessment of climate change impacts (Weyant et al., 1996; Cohen, 1997) or water quality in catchment areas (Turner et al., 2001).

There are technical, conceptual and institutional barriers to the effective use of data for policy making (AbouZahr et al., 2007). Examples of technical barriers are missing data, i.e. missing values in a time series (Britz et al., 2007), uncertain data, i.e. noisy data (Refsgaard et al., 2005), and non-available data i.e. no data sources available (Niemeijer, 2002). Conceptual barriers refer to different interpretations of data, while institutional bottlenecks include issues related to data management policies and conflict of interests between the hosting institutions. The use of dispersed data in IAM studies requires the integration of data sources, both in conceptual and technical terms. Here integration means to define shared concepts, to ensure consistency in data interpretation, units, spatial and temporal scales and to respect legal regulations of privacy, ownership and copyrights.

5.1.2. Integrated database

There have been several efforts in different application domains to bring various data sources together. For example, in the field of medical research, Ali et al. (2007) made an inventory of data sources available to assess the environmental conditions that could affect the frequency of chronic diseases in Pittsburgh, Pennsylvania. In the field of environmental sciences, Gobin et al. (2004) connected different data sources together to assess indicators on the European scale relevant to soil erosion, while Refsgaard et al. (2005) integrated data on the Water Framework Directive of the European Union. Herrero et al. (2007) developed a generic household-level database to store data on crop-livestock systems in developing countries. Villa et al. (2007) have demonstrated how artificial intelligence tools can be used for developing next-generation “intelligent databases” for the transparent and sound valuation of ecosystem services. The INSPIRE initiative (INSPIRE, 2008) of the European Commission targets the creation of a European spatial information infrastructure that improves the interoperability and the availability of spatial data across the EU. Refsgaard et al. (2005), Herrero et al. (2007) and Villa et al. (2007) reported on the availability of an integrated database to store the datasets, so that these datasets can be re-used easily for policy assessments.

As data sources on agricultural systems are distributed across institutions, scientists, who are required to integrate data, typically extract data from the original data sources in an ad hoc manner. This practice is certainly prone to errors and a paradigm shift is needed to overcome technical, conceptual and institutional problems. To support policy evaluation and policy impact assessment through indicators, typologies and models, there is a need for an integrated database on agricultural systems, which consistently combines data from different sources and which ensures easy availability of data. SEAMLESS (System for Environmental and Agricultural Modelling; Linking European Science and Society) is an IAM research project (Van Ittersum et al., 2008), which aims to provide a computerized framework to assess the impact of policies on the sustainability of agricultural systems in the European Union at multiple scales. This aim is achieved by combining micro and macro level analysis, addressing economic, environmental and social issues, facilitating the re-use of models and providing methods to conceptually and technically link different models (Van Ittersum et al., 2008). SEAMLESS provides a framework for policy assessment in agriculture by integrating relationships and processes across disciplines and scales and combining quantitative analysis with qualitative judgments and experiences (Ewert et al., 2009). In SEAMLESS, models of different kinds, designed for specific purposes and scales, are integrated for achieving the overall project objectives. Part of the integration activity is related to the extensive data requirements of the models. Data need to be collected and made consistent and available for serving dynamic biophysical models, static bio-economic farm models and partial equilibrium market models, with the ultimate goal to provide multi-scale assessment capability as to agricultural systems (Fig. 5.1 and Section 5.2.2). To achieve this goal, it is required to integrate several data-sources related to European agriculture, including economic, biophysical, climatic data, model simulation input and output data, scientific workflow configurations and calculation of indicators into a single relational database schema. By data integration in this chapter we mean both data alignment across different sources, so that a unified schema is defined with references to shared concepts and scaled data structures, and data homogenization, by creating one single database that can simultaneously hold data from different sources.

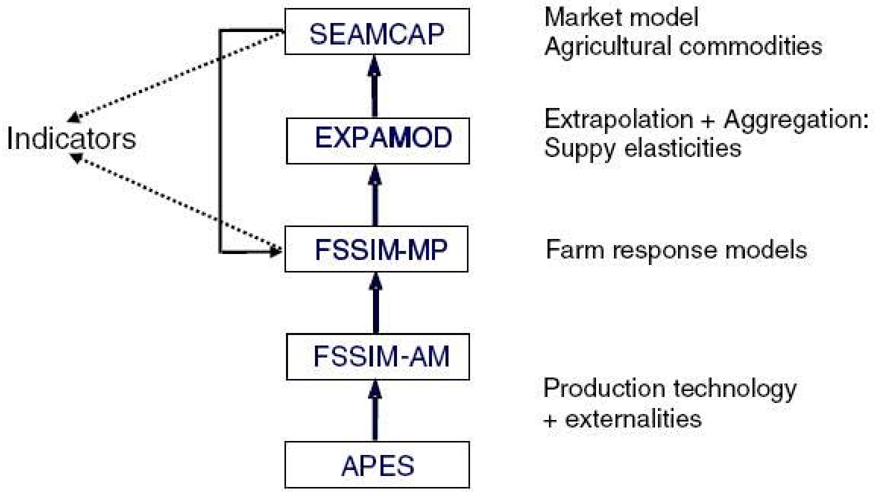


Figure 5.1. The models in SEAMLESS (after Van Ittersum et al., 2008). APES: Agricultural Production and Externalities Simulator; FSSIM-AM: Farm System SIMulator-Agricultural Management; FSSIM-MP: FSSIM-Mathematical Programming; EXPAMOD: EXtraPolation and Aggregation MODel and SEAMCAP: SEAMLESS version of the Common Agricultural Policy Regional Impact Analysis model.

The present chapter describes the SEAMLESS integrated database on European agricultural systems and demonstrates the use of the data in the database for calculating indicators and for model inputs in IAM. The chapter also describes the process of development of the SEAMLESS database, and the human factors involved in the process of reaching consensus across peers with clashing requirements and needs. To consistently define concepts across the different data sources, we adopted a structured process using an ontology as a means to arrive at one integrated database serving a set of models from different disciplines. We argue that this process is re-usable for other IAM projects, whereas we aim to make the end result (i.e. the database) freely available for non-commercial purposes in agricultural systems research and policy evaluations or assessments carried out in Europe. The chapter illustrates how the development and use of a shared ontology facilitates interdisciplinary research through development of an integrated database.

Section 5.2 describes the relevant data sources and models of the SEAMLESS project. Section 5.3 presents the background and the process of ontology engineering. Subsequently, the results are presented in the Section 5.4. The database on European agricultural systems is described, along with examples of the data present in the database and the process used to construct this database with a group of researchers. Section 5.5 offers a discussion of the database, the maintenance and support of the database and some reflections on the process of database development. Finally, conclusions and recommendations are provided.

5.2. Data sources and their use in models

5.2.1. Data sources

5.2.1.1. Farm Accountancy Data Network

The Farm Accountancy Data Network (FADN) (EC, 2008a) is an instrument for evaluating the income of agricultural holdings and the impacts of the Common Agricultural Policy. It consists of an annual survey carried out by the Member States of the European Union. The Member States of the European Union collect every year accountancy data from a sample of the agricultural holdings (EC, 2008a). The sample only covers 'professional' farms, which means that small, part-time, and hobby farms are poorly represented. Data collected per farm include physical and structural data, such as location, crop areas, livestock numbers, labour force, and economic and financial data, as the value of production of different crops, sales and purchases, production costs, production quotas and subsidies. Data on farm management and externalities are not collected. Due to legal disclosure rules, data from FADN can only be displayed as averages of more than 15 sample farms, as data from individual farms should not be traceable for reasons of privacy.

5.2.1.2. European Soil Database

The European Soil Database (ESDB) (ESBN, 2008) provides a harmonised set of soil parameters, covering Europe (the enlarged EU) and bordering Mediterranean countries, to be used in agro-meteorological and environmental modelling at regional, national, and/or continental levels. It is 1x1km raster data and it contains the Soil Geographical Database of Eurasia, PedoTransfer Rules Database, Soil Profile Analytical Database of Europe and Database of Hydraulic Properties of European Soils (ESBN, 2008). These soil data have been supplemented with selected variables from the SINFO project (Baruth et al., 2006), which improved the soil parameters, pedo-transfer rules and the soil classification for use in a yield-forecasting tool. Finally, the map of organic carbon content in topsoils in Europe (Jones et al., 2005) was crucial for the development of the agri-environmental zones used in SEAMLESS (Section 5.4.1.3).

5.2.1.3. European Interpolated Climate Data

The European Interpolated Climate Data (EICD) (JRC, 2008) provide interpolated daily data for a grid of 50x50 km covering Europe and Maghreb (period 1975 - today). The majority of the original observations originate from around 1500 meteorological stations across the European continent, Maghreb countries and Turkey. The observations at station level are not available in the dataset, only spatially interpolated data are (JRC, 2008). The interpolation is a simple two-step procedure in which the first step is the selection of up to 4 suitable meteorological stations for the determination of the representative meteorological conditions for a grid cell. The actual interpolation, the second step, is a simple average for the meteorological parameters, corrected for an altitude difference in the case of temperature and vapour pressure (Van der Goot, 1997).

5.2.1.4. Surveys on Farm Management

Farm management data have been collected through dedicated surveys as part of the SEAMLESS project (Borkowski et al., 2007). In the SEAMLESS project, a lack of European data on agricultural management was identified. Agricultural management data are the use of inputs (fertilizers, pesticides, irrigation) and the timing of input use at crop level, which are crucial to bio-economic farm models and biophysical crop growth models. FADN only provides aggregated farm level input data often expressed in monetary terms. Two different surveys (Borkowski et al., 2007) were developed as part of the SEAMLESS project: a detailed and a simple one. In the detailed survey, only data for arable systems were collected including timing and amounts of inputs, crop rotations, machinery, labour requirements and costs. It has been carried out in five regions in Europe (Brandenburg, Andalucia, Midi-Pyrenees, Flevoland and Zachodniopomorskie). The detailed survey was completed by regional experts, who in their day-to-day work provide advice to farmers or work regularly with farmers, and thus describes an average farmer behavior. The detailed survey aims to meet the input requirements of biophysical crop growth models.

The simple survey was applied to a larger sample of 16 regions in Europe. It collects data on arable, livestock and perennial agricultural systems. These 16 sample regions aim to cover the range of biophysical conditions and farm types present in the European Union. The simple survey differs from the detailed survey as only a sub-set of the variables from the detailed survey is collected, including economic variables (e.g. costs, product prices), production, rotations and some aggregate variables describing input use (e.g. total nitrogen use for a crop or total medicine costs per livestock unit). The simple survey was completed by scientists working in the region with the help of farm management handbooks, which are used by farmers for advice.

5.2.1.5. COCO/CAPREG

The COCO/CAPREG dataset (Britz et al., 2007) is based on NewCronos (Eurostat, 2008) and FAOSTAT (FAO, 2008). Missing values, missing time series and incorrect values from NewCronos and FAOSTAT were estimated and adjusted through statistical estimation procedures. COCO/CAPREG is the dataset linked to the SEAMCAP market model (a market equilibrium model detailed in Section 5.2.2.3). This dataset provides the data on agricultural policies and prices in the 27 Member States from 1985 and 2004, e.g. subsidies given to farmers for different regions, cuts of subsidies given to farmer, coupling degrees and prices per Member State, subsidized exports and tariff agreements between European Union and trading blocks.

5.2.1.6. Relevance of typologies

The datasets from the FADN, ESDB and EICD have been categorised into farm and regional typologies (Metzger et al., 2005; Andersen et al., 2007a; Hazeu et al., 2009) to enable modelling in homogenous spatial units and to allow for characterization of the variation in the environment, e.g. climate, soil and farms. This is useful for sampling purposes. For example, farm management data were not available and they cannot easily be collected for all regions across Europe due to budget and time restrictions. Instead, based on classification in typologies, representative regions were selected for the simple and detailed surveys (Section 5.2.1.4). Typologies are used to combine data, to provide a flexible and manageable data structure and to respect disclosure rules. Further regional typologies have been developed which characterize regions to provide contextual information for the assessments. Examples of regional typologies are livestock density, share of area in nitrate vulnerable zones and degree of rurality.

The data sources have been aligned with the existing administrative categorization of the EU territory, like the Nomenclature of Territorial Units for Statistics (NUTS) (EC, 2008b). In SEAMLESS, the NUTS-level of relevance is NUTS-2 (except for United Kingdom, where level 1 is used) and when reference is made to NUTS-regions in this chapter, NUTS-2 regions are intended. EU25 has 270 NUTS-2 regions, which typically correspond to provinces, or constituent states/cantons.

5.2.2. Models using the data

5.2.2.1. APES: a dynamic crop growth simulation model

APES is a cropping system model estimating the biophysical processes of agricultural production systems, at point level, in response to weather, soils and different options of agro-technical management (Van Ittersum and Donatelli, 2003). APES is a modular simulation model targeted at estimating the biophysical behaviour of agricultural production systems taking into account the interaction among weather, soil and crop characteristics and different options of agricultural management. Biophysical processes are simulated in APES with deterministic approaches which are mainly based on mechanistic representations of biophysical processes (Donatelli et al., 2009).

5.2.2.2. FSSIM: a bio-economic farm model

The Farm System SIMulator (FSSIM) is a bio-economic farm model developed to assess the economic and ecological impacts of agricultural and environmental policies and technological innovations. A bio-economic farm model is defined as a model that links formulations describing farmers' resource management decisions, to formulations that describe current and alternative production possibilities in terms of required inputs to achieve certain outputs (both yield and environmental effects) (Chapter 2). FSSIM consists of a mathematical programming model (FSSIM-MP), and an agricultural management module (FSSIM-AM) (Louhichi et al., 2009).

5.2.2.3. SEAMCAP: a market level model

SEAMCAP is a version of the model Common Agricultural Policy Regional Impact Analysis (CAPRI) (Heckelei and Britz, 2001) integrated in SEAMLESS. CAPRI is a partial equilibrium model for the agricultural sector. SEAMCAP makes use of non linear mathematical programming tools to maximise regional agricultural income with explicit consideration of the Common Agricultural Policy instruments of support in an open economy where price interactions with other regions of the world are taken into account. It consists of a supply and market module, which interact iteratively.

5.2.2.4. EXPAMOD: a regional upscaling model

EXPAMOD is an econometric model describing price-quantity responses of farms given specific farm resources and biophysical characteristics that are available EU-wide (Pérez Domínguez et al., 2008). It provides an aggregation procedure to make the regional supply modules of CAPRI behave like the aggregate of the farm (FSSIM) models of the same region – apart from additional aspects entering the market supply such as regional land or political constraints (premium ceilings). All available FSSIM models run for ranges of exogenously fixed prices, computing multi-dimensional price-quantity response surfaces. Thus, the econometric model is estimated using simulated price-response data for farm types in regions for which farm type models exist and then applied to project supply responses of other farm types and regions (Pérez Domínguez et al., 2008).

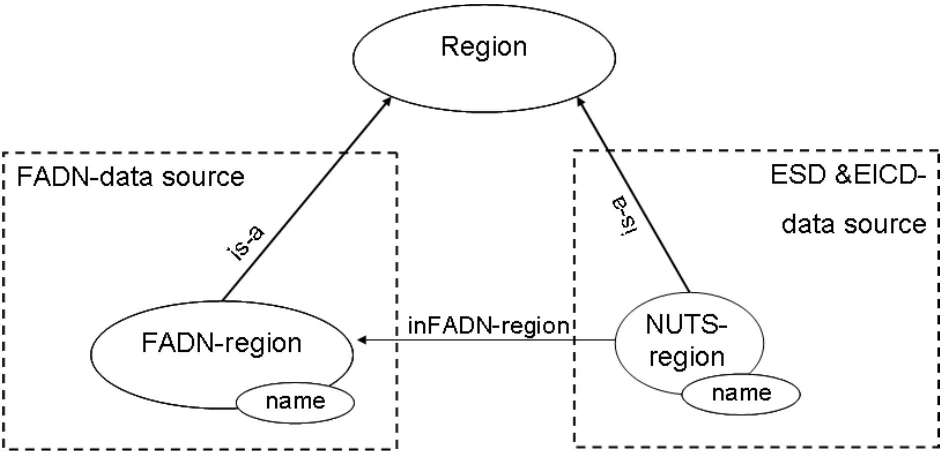
5.3. Database development, data consistency and integration

5.3.1. Process of database development

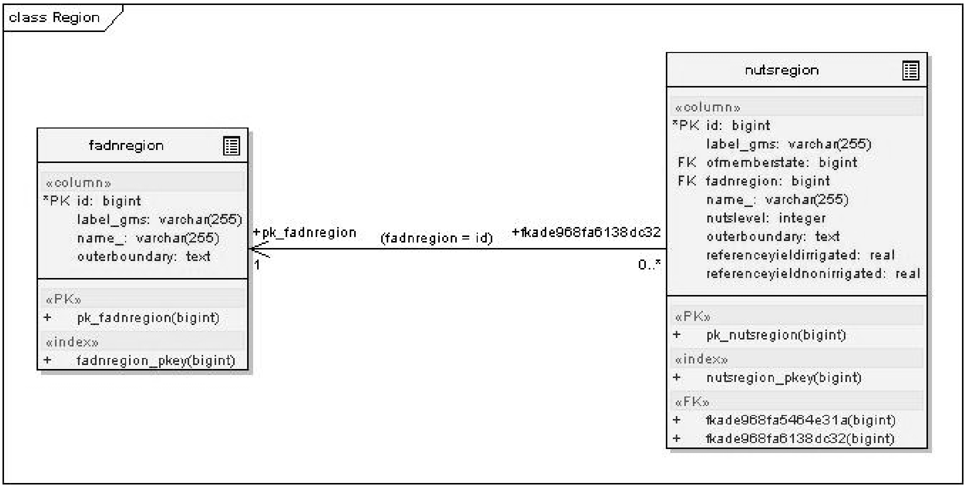
Data integration across the sources presented above requires to take into account complex conceptual problems, related to the terminology adopted, the scale of information, and the heterogeneity of the original database schemas. For example, FADN, ESDB and EICD all refer to a "Region" entity. In the case of FADN, the definition of regions is different than those of ESDB and EICD. ESDB refer to soil mapping units and EICD refer to 50x50km grid, which were both linked to NUTS regions (Fig 5.2a), when preparing the data for the database. FADN uses a delineation of regions that is specific to FADN, and these regions are referred to as FADN-regions in this chapter. In integrating the data sources in one database schema, these data sources have to be adapted to shared concepts, to respect geographical entities and to be aligned in time, e.g. covering overlapping time periods. Integrating the data sources into one database is a time consuming and challenging task that requires collaboration of scientists from agricultural economy, environmental science, agronomy and computer science, with dissimilar education and research experience.

To tackle the heterogeneity of the constituent data schemas, we developed an overall ontology, covering the union of the constituent data sources and domains. An ontology is the appropriate tool for defining a shared conceptual schema, as ontologies consist of a finite list of concepts and the relationships between these concepts (Antoniou and Van Harmelen, 2004), and they are expressive enough for defining equivalent entities, hierarchies, complements, unions or intersections, based on description logics. This was particularly useful for marking and resolving ambiguities across the original schemas.

A shared ontology is an ontology that is jointly developed between a group of individuals, in this case researchers. A collaborative approach was adopted for developing a shared ontology about the different data sources in SEAMLESS. Our development was ‘a joint effort reflecting experiences and viewpoints of persons who intentionally cooperate to produce it’ and thus requires a consensus-building mechanism (Holsapple and Joshi, 2002). Part of our effort was based on an inductive approach (Holsapple and Joshi, 2002), where the shared ontology was developed by examining and analyzing the initial data sources and extracting relevant properties or discussing the relationships between concepts in these data sources.



a. The different definitions of the concept Region between data sources



b. The representation of the relationships between FADN-region and NUTS-region in a relational database

Figure 5.2. The different types of Regions in the integrated database in an ontology schema (a) and a relational database schema (b). The same relationship is represented in a. and b. between NUTS-region and FADN-region, with the difference that the relationship in the ontology schema (a) has a name ('inFADN-Region') and definition, while this is not the case in the relational database (b). ESDB = European Soil Database; EICD = European Interpolated Climate Data

Semantic modelling and ontologies are more powerful for domain modelling than conventional relational data schemas, and this is why we adopted ontologies for defining the integrated schema, First of all, ontologies are richer in their representation of relationships between concepts than relational database schemas (Fig. 5.2). In an Web Ontology Language (OWL) (McGuinness and Van Harmelen, 2004) ontology relationships have a direction and can be shared across concepts, restricted with logical constraints, and form hierarchies. Also, ontologies have a strong inter-operability background, as they are in line with the Semantic Web initiative (Berners Lee et al., 2006). There are much more tools and techniques for ontology alignment and integration. For example, two ontologies developed in separate efforts can easily be linked to each other by investigating the semantic relationships between their concepts (El Gohary and El Diraby, 2005). Furthermore, OWL ontologies can be connected by a reasoner that is based on description logics and thus data can be validated against logical constraints. Finally, an ontology may be considered as distinct product for capturing knowledge, which can be re-used in the future for building other systems.

5.3.2. Technical implementation

The shared ontology was subsequently translated into a relational database schema. A relational database schema provides the structure of the database, in which the data from the different data sources can be entered. This translation from ontology to relational database schema was done based on the conventions of the Semantic-Rich Development Architecture (SeRiDA) (Athanasiadis et al., 2007b; Athanasiadis et al., 2007c), which acts as a bridge between different programming paradigms, e.g. object-oriented programming, relational databases and ontologies (Athanasiadis et al., 2007b). Object-oriented programming is used in SEAMLESS for model and application development, relational databases for persistent storage of data and ontologies for defining and storing knowledge.

The integrated database is running on a PostgreSQL database server (PostgreSQL, 2008). The models are linked to the database through Hibernate (JBOSS, 2008) and the exchanged datatypes are implemented with JavaBeansTM. The database is linked to a spatial database that provides geographical information, exploiting the

PostGIS capabilities (PostGIS, 2008). The spatial information is also available through Web Mapping and Web Feature Services provided by a GeoServer (GeoServer, 2008). The entire database deployment and data management solution is based on Open Source software. A detailed description of the data management process and the technical integration of the models is discussed in Athanasiadis and Janssen (2008).

5.4. Results

5.4.1. Database on European agricultural systems

5.4.1.1. Full ontology

Figure 5.3 provides a partial view of the ontology developed for the database on European agricultural systems as developed in the SEAMLESS project. Figure 5.3 illustrates the part related to soil, farm and climate data. It includes typology concepts, such as Farm Specialization and Farm Size, concepts that facilitate spatial links, such as NUTS regions and Climate Zones, and concepts that hold the actual data, such as Representative Farm, Soil Characteristics and Daily Climate entities. The current version (October 2008) of the database consists of 379 tables including 2’379 fields and with 487 relations between the tables. The database exceeds 12 million records.

5.4.1.2. Representative farms

A central concept of the ontology is the concept of Representative Farm, which defines a Farm Type in an FADN region in Europe for a specific year. A Farm Type is specified according to the dimensions of farm size, farm intensity and farm specialization (Andersen et al., 2007a) (Fig. 5.4). As an example of a classifying concept, Farm Intensity classifies farms according to their total monetary output of agricultural produce per hectare (Andersen et al., 2007a). If the total output is below

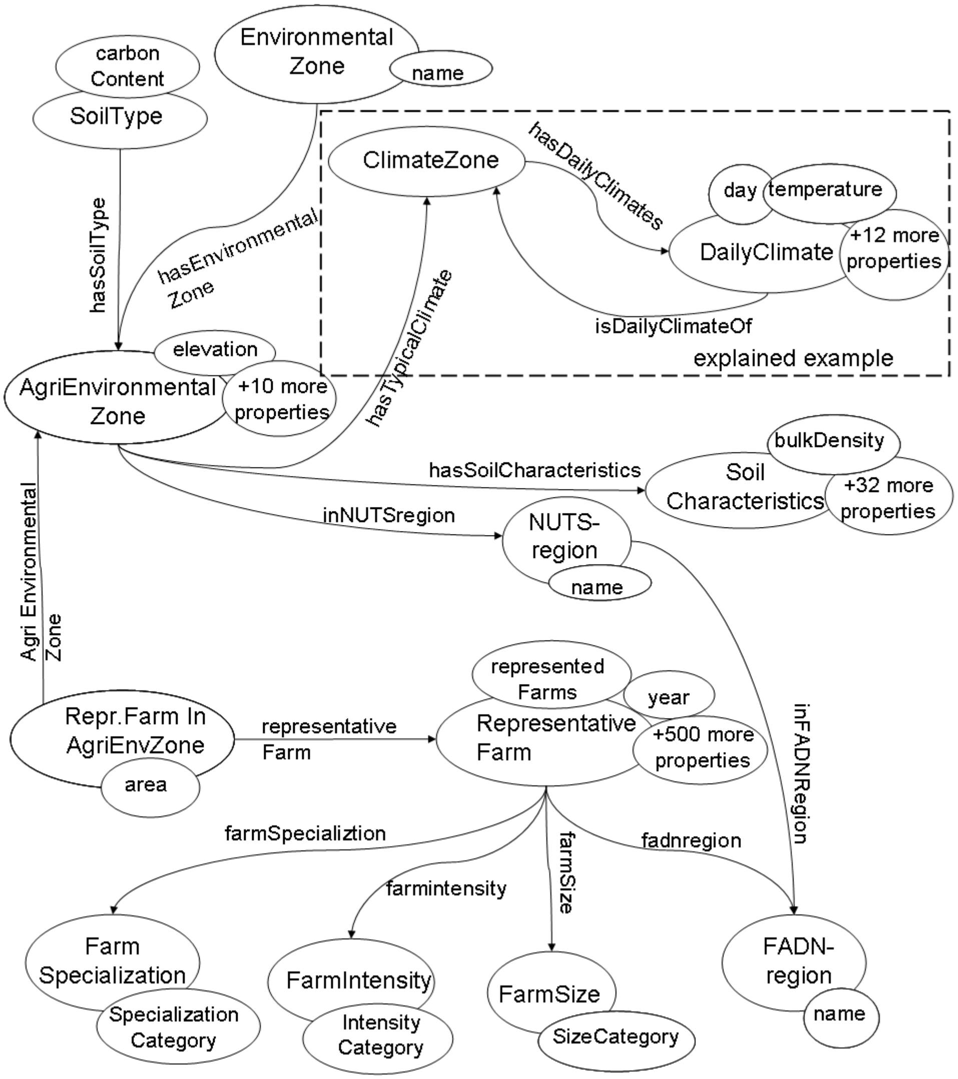


Figure 5.3. An ontology-schema of the database on European agricultural systems showing the parts on farms, soils, climate and their links. Two concepts (ClimateZone and DailyClimate: large ellipses), their relationships (hasDailyClimates and isDailyClimateOf: uni-directional arrows) and the properties of the concepts (name, temperature, and 12 more properties: small ellipses) can be found in the explained example (dashed box). The figure can be read by following the direction of the arrows, for example ClimateZone.hasDailyClimates (DailyClimate). A daily climate is characterised by a day, a temperature of that day and twelve more properties.

500 euros per hectare, then the farm falls in the class of low intensity, if it is between 500 and 3000 euros, then it is medium intensity and if is more than 3000 euros, then it is high intensity. The threshold values are adjusted with yearly producer price indices. The values presented above refer to 2003. While a Farm Type is not linked to a specific region or year, a Representative Farm is associated to a region and a year (Fig. 5.4).

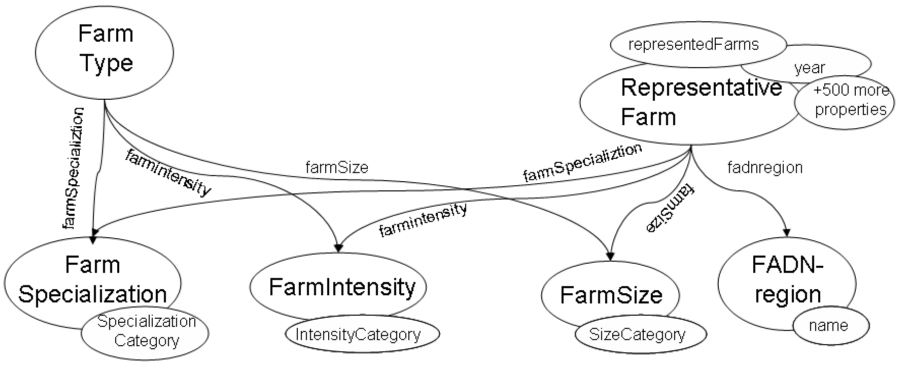


Figure 5.4. The concepts Farm Type and Representative Farm and the relationships to their classifying concepts (For an explanation of how to read this figure, see figure 5.3).

5.4.1.3. Climate and soil data

Another central concept is that of the agri-environmental zone, that links soil and climate data. An AgriEnvironmentalZone is a unique combination of an EnvironmentalZone, the SoilType and NUTS-region. An AgriEnvironmentalZone is the smallest homogenous area in terms of climate and soil data. Environmental Zones are used to stratify the diverse European Union in zones with a similar climate (Metzger et al., 2005). Environmental Zones cover more than one administrative region. A Climate Zone is a unique combination of a NUTS-2 region and Environmental Zone and for each Climate Zone, a set of climate data are available. A Climate Zone is associated with the daily climate data for a 25-years time period. Examples of daily climate data attributes are rainfall, minimum and maximum daily temperature and wind speed at 10m.

Each AgriEnvironmental Zone is linked to a set of soil data, which are classified in Soil Types. Six different Soil Types were defined according to topsoil organic carbon levels (Hazeu et al., 2009). For each unique combination of a Soil Type and a Climate Zone a set of soil data is available as stored in the concept of Soil Characteristics. Examples of properties of the soil characteristics are thickness of soil layers, textural class and maximum usable moisture reserve.

The link between AgriEnvironmental Zones and Representative Farms is made through statistically allocating an area of an AgriEnvironmental Zone to each Representative Farm (Elbersen et al., 2006). This implies that the farmed area within each AgriEnvironmental Zone is allocated to one or more Representative Farms and each Representative Farm manages farmed areas in one or more AgriEnvironmental Zones. As can be seen from Figure 5.3, Representative Farms and AgriEnvironmental Zones are based on different administrative regions. AgriEnvironmental Zones refer to NUTS-regions and Representative Farms refer to FADN-regions. Usually the borders of the NUTS and FADN-regions coincide, however some FADN-regions consist of several NUTS-regions. Through allocating the area of Agri-Environmental Zones to Representative Farms, this mismatch between the borders of FADN-regions and NUTS-regions has been resolved.

5.4.1.4. Agricultural management

As the agricultural management differs between and within regions, Regional Agricultural Management Zones were created. A Regional Agricultural Management Zone can be linked to distinct sets of agricultural management data and each Regional Agricultural Management Zone refers to one or more AgriEnvironmental Zones (Fig. 5.5). The central concept in Figure 5.5 is the RotationElement, which signifies one year of crop rotation as found in a region. A rotation is defined as a sequence of crops in time and space, where the last crop is the predecessor of the first crop (creating a loop) and rotations are widely practiced in agriculture for pest control, soil fertility management and risk diversification. The RotationElement links to one or more ManagementInZones, which means that crop management applied to the crops in a rotation is different across Regional Agricultural Management Zones. For example, in a Regional Agricultural Management Zone with high yield potential (favourable climate and soils) more nitrogen may be applied to the crop than in a Regional Agricultural Management Zone with lower yield potential. This assumes that rotations are the same throughout the NUTS-region, and only the management of the crops in the region can differ. Note that for explanatory purposes focus was given only on arable crop management data. The database holds data on livestock and perennial systems.

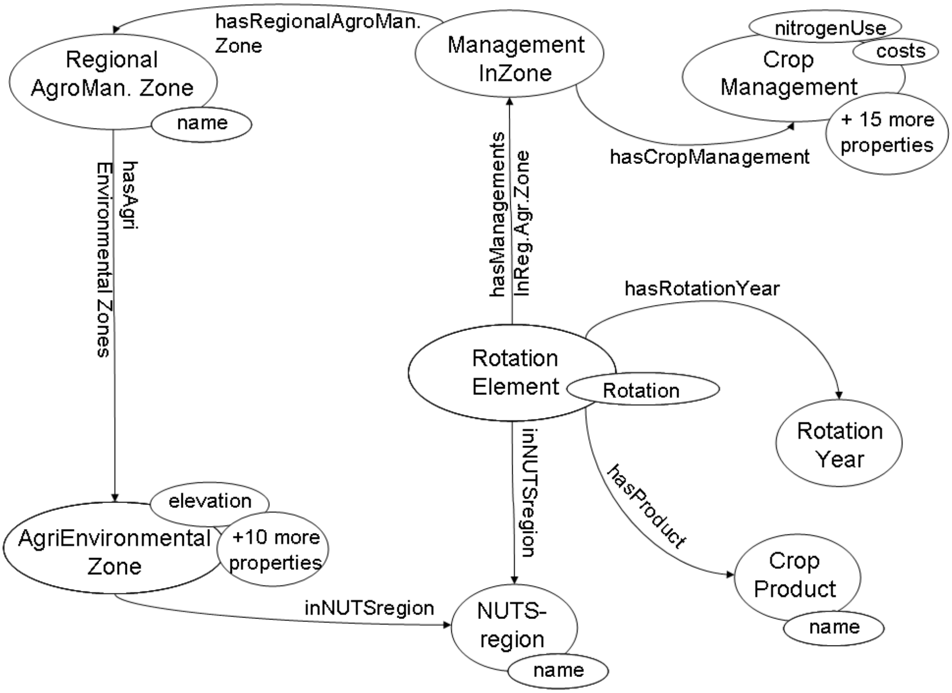


Figure 5.5. Agricultural management data and their links to NUTS-regions and AgriEnvironmental Zones (For an explanation of how to read this figure, see figure 5.3).

5.4.1.5. Policy data

Finally, data on agricultural and environmental policies are linked to Member States and NUTS-regions. Each Member State consists of one or more NUTS-regions. Figure 5.6 shows part of the database schema for policy data, which are related to the premiums the European Union pays to farmers and the decoupling of these premiums as part of the Common Agricultural Policy 2003 reform (EC, 2003). Decoupling means that financial support is not related to production anymore, but farmers receive income support instead. The European Union has established premium amounts per premium groups (e.g. a group of crops or animals for which the same premium is provided), the Basic Premiums. Individual Member States can, with some restrictions, decide on the percentage of decoupling of these Basic Premiums for different Premium Groups. These relationships are shown in Fig. 5.6, as the Basic Premiums are not linked to a NUTS Region and as the relationship 'hasMemberstate' between Coupling Degree and Premium Group in Fig. 5.6. The database on European agricultural systems contains policy data for 14 more policy measures.

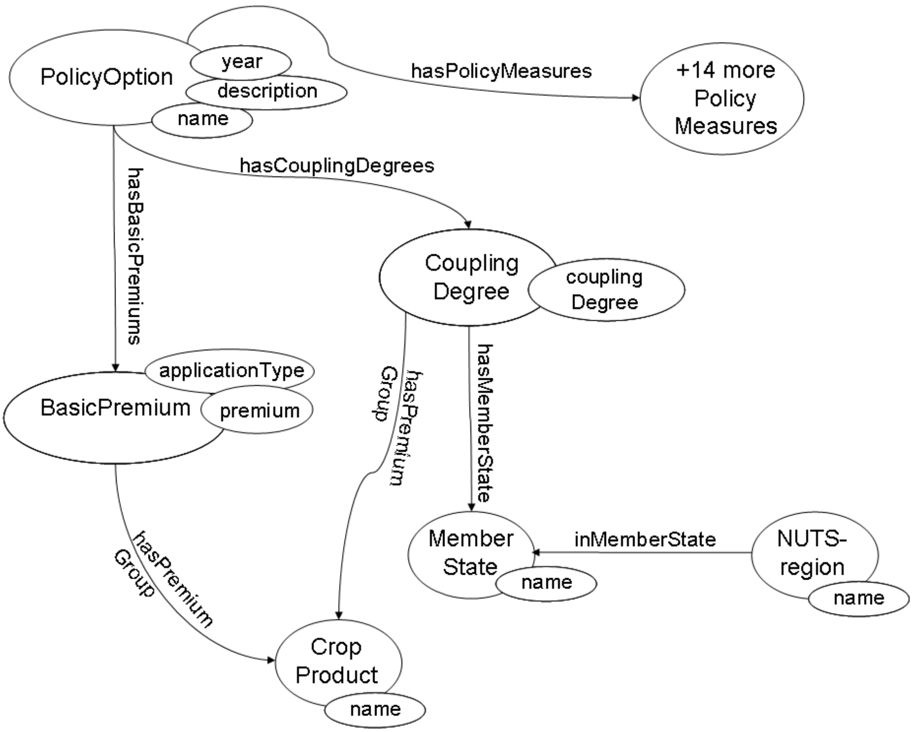


Figure 5.6. Part of the policy data related to premiums for farmers and coupling degree of these premiums to production (For an explanation of how to read this figure, see figure 5.3).

5.4.2. Method to develop the integrated database

To develop a shared ontology for all data-sources, three scientists (a computer scientist, an expert on agri-environmental policies, and a systems analyst) engaged in an integration process. These three scientists involved other domain experts in the integration process, when additional knowledge was required.

The integration process was an iterative procedure, with four milestones: three intermediate "prototypes", each concluded with a stable version of the database schema used for running the models, and one final version. Every prototype began with a phase of ontology building and review in several iterations. The ontology was developed using Protégé (Knublauch, 2005), an ontology editor. Once the shared ontology was fixed, it was exported to a relational database schema using the SeRiDA-framework (Section 5.3.2). Subsequently, the data from the original sources were entered into the database, which led to the identification of obstacles and further issues. These issues were discussed again in a new iteration among the domain scientists involved, and the resolutions were reflected in the ontology, resulting in an updated stable version of the database schema, which was then released as a version and linked to the models. As a final step in each prototype, the relational database schema and shared ontology were reviewed by the three scientists involved and lists of improvements were made. During the review of the database schema of the three prototypes, scientists tried to simplify the shared ontology and relational database schema as much as possible, as the shared ontology had the tendency to grow in detail and complexity.

As part of the fourth and final version of the database schema, metadata have been included as part of the ontology in accordance with ISO (ISO, 2008) and the INSPIRE (INSPIRE, 2008) standard. The metadata document the original data sources, textual descriptions, units and contact persons for the original data source.

5.4.3. Examples of the use of data from the database

5.4.3.1. Examples of data extractions (use of typologies and indicators)

Table 5.1. The share of agricultural area in EU15 managed by different farm types according to intensity in the period from 1995 to 2004[1]

	Share of agricultural area managed by		
	Low intensity farms	Medium intensity farms	High intensity farms
1995	29.5	62.6	7.9
1996	28.6	63.0	8.3
1997	27.9	63.3	8.8
1998	27.3	63.5	9.2
1999	25.8	63.6	10.6
2000	24.4	64.2	11.4
2001	26.3	62.2	11.5
2002	23.8	64.6	11.7
2003	23.3	64.7	12.0
2004	21.0	66.0	13.0

[1] Source: EU FADN-DG AGRI/G-3; SEAMLESS adaptation. The Farm Intensity is defined by the total monetary output per ha. The threshold values vary over time according to the price index on agricultural products. In 2003 the threshold values were: Low intensity < 500 €, medium intensity => 500 and < 3000 € and high intensity => 3000 € (Andersen et al., 2007a).

The following section gives some examples of data that can be extracted directly from the SEAMLESS database providing novel ways of aggregating or combining the original data. The first example provides an overview of the trends in intensity of farming in EU-15 since 1994 (Table 5.1). The example is based on FADN data aggregated using the typology of farm types according to intensity. Clear trends can be identified in Table 5.1: the share of the agricultural area managed by low intensity has declined continuously over the period from almost 30% to close to 20%. At the same time the agricultural area managed by high intensity farms has increased with 5.1% from 7.9% of the area to 13% of the area.

Table 5.2. Number of NUTS2 regions in EU15 according to livestock density (Livestock units per ha) in 2003 and share of area designated as Nitrate Vulnerable Zones (NVZ).

NVZ area in % of total area	Livestock density LU/ha				
	< 0.5	>= 0.5 and < 1	>= 1 and < 2	>= 2	Total
0	17	18	11	2	48
>0 and <33	7	21	12	10	50
>=33 and <66	2	6	5	1	14
>=66 and <100	3	7	3	1	14
100	8	17	28	15	68
Total	37	69	59	29	194

[1] **Source: EU FADN-DG AGRI/G-3; SEAMLESS adaptation.**

The second example shows the relationship between livestock density and the Nitrate Vulnerable Zones designated according to the Nitrates Directive (Table 5.2). The Nitrate Vulnerable zones include catchments that drain to surface and groundwater where the NO3 content exceeds 50 mg/l. As can be seen in the table, there is a tendency towards higher livestock densities in the regions with the largest share of the area as Nitrate Vulnerable Zones. In the 68 regions where the entire area is designated as NVZ, 63% of the regions have an average livestock density above 1, and 22% have a livestock density above 2. Compared to this, livestock density only exceeds 1 in 27%, and 2 in 4% of the regions, where no areas are designated as Nitrate Vulnerable Zones. The analysis can be used to identify the hot spots, where a high livestock density is found in regions with a large share of the area designated.

Table 5.3. Family farm income and livestock density for seven soil types in the EU.

Soil type	Carbon content in topsoil	Family farm income (€/ha)	Livestock density LU per ha
1	< 1.23	866	0.5
2	>= 1.23 and < 2.46	594	0.8
3	>= 2.46 and < 3.94	416	0.9
4	>= 3.94 and < 5.66	671	1.0
5	>= 5.66.and < 8.86	338	1.1
6	> 8.86	435	1.2
7	No soil information	463	1.2

1 Source: EU FADN-DG AGRI/G-3; SEAMLESS adaptation. In SEAMLESS the soil types are defined by the carbon content in the topsoil (Hazeu et al., 2009).

The last example explores whether there is a relationship between soil types, farm family income and livestock densities (Table 5.3). Generally it is assumed that the agronomic potential of the soils increases with increased carbon content in topsoil except for the soils with a very high content, which normally are related to other restrictions on agricultural production. This high agronomic potential might lead to a higher family farm income. However, looking at the family farm income in Table 5.3, no correlation exists: the lowest income per hectare is found on soil type 5 and the highest on soil type 1 with no trend in between. Taking under account the livestock density, on the one hand we observe that the livestock density seems to be correlated with the soil types with an increasing density following increasing carbon content. On the other hand, this cannot be used to explain the variation in family farm income. The SEAMLESS database provides data that can be used to explore these relations also at regional and local levels or to seek for other variables to examine the relationships.

5.4.3.2. Data as model inputs for FSSIM simulations

FSSIM, a bio-economic farm model (Section 5.2.2.2), has been applied to Flevoland, a NUTS2-region in the Netherlands. Flevoland has been reclaimed from the sea in the 1960's and its young and fertile soils are very productive. It is very homogenous in terms of soil and climate as can be seen in Table 5.4. For FSSIM, the data of relevance from the database are the Representative Farms found in Flevoland, as for each of these Representative Farms, FSSIM executes to obtain a simulated cropping pattern. FSSIM execution requires a set of possible farming rotations to choose from, and data on the associated crop management for each rotation. This is illustrated for one sample rotation in Table 5.4. The database provides all these data, and by using hibernate querying facilities, the model is able to retrieve them easily. The results of the model are verified by comparing the simulated cropping pattern with the observed cropping pattern as found in 2003 for each of the representative farms in Flevoland. In Table 5.4 only a limited subset of

Table 5.4. A sample of data for Flevoland region in the Netherlands for 2003 (Source among others. EU-FADN-DG AGRI-G3, Meteorological data Source JRC/AGRIFISH Data Base – EC – JRC)

Representative farms

Farm Specialization / Land Use	Farm Intensity	Farm Size	Usable Farm Area (ha)	Area in potatoes (ha)	Area in sugarbeet (ha)	Area in wheat (ha)	Percentage area per AgriEnvironmental Zone 2993	1317
Arable/Specialised crops	High intensity	Medium scale	16.8	4.8	3.1	2.7	100%	
Arable/Specialised crops	Medium intensity	Large scale	67.2	17.9	11.2	10.4	100%	
Arable/Specialised crops	High intensity	Large scale	69.0	24.8	9.1	11.5	90.3%	9.7%
Arable/Others	High intensity	Large scale	35.2	3.5	1.3	2.0	100%	

Climate-Soil data per AgriEnvironmental Zone

AgriEnvironmental Zone	Environmental Zone	Average Rainfall (mm/day)	Av. Minimum temper-ature (°C)	Average Maximum Temperature (°C)	SoilType: soil carbon content (%)	Clay content (%)	Sand content (%)	Silt content (%)
2993	Atlantic Central	1.69	6.24	14.65	3.94-5.66%	20	45	35
1317	Atlantic North	1.68	6.22	14.49	3.94-5.66%	20	45	35

Table 5.4. A sample of data for Flevoland region in the Netherlands for 2003 (Cont.)

Sample Rotation for Flevoland						
Crop	Year	Costs fertilizer (€/ha)	Sowing date (week of year)	Labour (hours/ha)	Nitrogen use (kg N/ha)	Yield (tonnes/ha)
Soft wheat	1	113	42	10.7	205	8.2
Potato	2	269	11	27.5	255	53.4
Soft wheat	3	113	42	10.7	205	8.2
Sugar beet	4	147	13	19.6	150	65.5

Policy data for direct payments in Flevoland in 2013			
Premium Group	Premium (€/tonnes)	Decoupling degree (%)	Regional reference yield (tonnes/ha)
Energy crops	45	100	4.9
Cereals, oilseeds, pulses	63	0	6.6
Obligatory setaside	63	0	4.9

the input data for FSSIM have been provided, e.g. the database contains more rotations, representative farms for more years and many more properties to describe the soil, climate and representative farms found in Flevoland.

5.5. Discussion

5.5.1. Use of the database

The database on European agricultural systems holds data on different aspects of the agricultural systems, e.g. cropping patterns, production, farm structural data, soil and climate conditions, current agricultural management and policy information. As demonstrated in Section 5.4.3, the database can be used to directly compute indicators related to agricultural and environmental policies in Europe or for policy assessments through the use of one or a set of models. The database in its current form is used by the models APES, FSSIM, SEAMCAP and EXPAMOD (Section 5.2.2). New models and indicators with similar data needs can easily be linked to the database, for example, the database could be useful for computing indicators on soil erosion (Gobin et al., 2004), energy use (Pervanchon et al., 2002), crop diversity (Dramstad and Sogge, 2003), pesticide usage and leaching (Reus et al., 2002) and marginalization based on farm income and employment (EEA, 2005).

The data in the database are organized according to typologies (Section 5.2.1.6), which implies that it is based on aggregated data (e.g. farm typology (Andersen et al., 2007a)), interpolated data (e.g. EICD (JRC, 2008)) or categorized data (e.g. ESDB (ESBN, 2008)). The database does not contain the original data on which these averages, interpolations and categorizations are based, which is required to respect disclosure rules and to avoid data pre-processing for each model and indicator computation.

The database aims to achieve a full coverage of the European Union, but this is not feasible for all data sources. For example, the FADN (EC, 2008a) contains data about the 10 new Member States only for 2004 onwards, while for 12 ‘old’ Member States data are available from 1990 and onwards and for three Austria, Finland and Sweden data are available from 1995 and onwards, as these joined the EU in 1995. There was no European-wide data source available on agricultural management, so a first effort was made in the SEAMLESS project to collect this type of data for a sample (Section 5.2.1.4) out of the 270 NUTS-regions in the EU25. Obviously more work is required to add more regions and to obtain time series in order to increase the representativity of the agricultural management data. Still the database holds the most complete set of data available on agricultural systems in Europe, and data gaps are due to the original data sources on which the SEAMLESS database depends.

5.5.2. Availability, extension, support and maintenance

The database will be made available for non-commercial use in other projects requiring data on agricultural systems in the European Union (Information on access can be found on www.seamlessassociation.org and additional documentation can be found in Andersen et al. (2007c)). Using the SEAMLESS integrated database instead of using original data sources has the advantages that (a) several data sources are available on one server instead of on several locations in different formats, (b) difficult questions of data integration and consistency have been solved by specialists familiar with the original data sources and (c) the pre-processing of the original data sources is already done.

A plan for the maintenance of the database beyond the lifetime of the SEAMLESS project is available that ensures the database will be available in an updated version for at least 3 years and hopefully longer. The maintenance plan provides full documentation of how to update the database with data from the different data sources and ensures that new versions of relevant data sources in the database will be included as they come available, for example a dataset for 2005 for FADN data (EC, 2008a). Not all the original data sources are frequently updated in their structure and content, although for some data sources (e.g. FADN (EC, 2008a) and EICD (JRC, 2008)) new data become available annually. The introduction of new versions of the data sources can be automated, although this is dependent on the stability of the original data sources in their variables and structure. New models and indicators might require new data, that is not currently in the database, for which the database needs to be extended. Extension of the database with new data sources is encouraged and the methods described in Sections 5.3.1 and 5.4.2 for conceptual and technical integration are recommended. For this, the shared ontology would need to be extended for the new dataset and links to the concepts already in the ontology need to be made. Second, the database schema can be made and finally, this can be filled with the new data.

5.5.3. Reflection on development and technical implementation

The integration of multiple data sources into one shared ontology following an iterative process was successful, as it led to one database schema in which all the data from different sources could be stored. The iterative process with different versions was required to step-wise improve the shared ontology. During the review of the first and second version of the shared ontology it was concluded that the shared ontology was too complex and that some relationships between concepts were ambiguous and therefore difficult to understand. The use of shared ontologies can highlight such complexities and ambiguities as scientists are forced to clearly define the concepts part of the ontology and as the concepts have to be consistently and coherently related to other concepts in the ontology. An important test for any shared ontology is whether the data from the data sources can be inserted in the relational database schema based on the shared ontology. Critical success factors in our approach of ontology development are the commitment of participants to the process and the presence of one or more knowledge engineers. Knowledge engineers are impartial scientists who can pro-actively identify and discuss open issues to find agreement, and who do not push their own opinion on the content of the shared ontology.

The database holds data that are spatially and temporally consistent and this difficult task of integration of different data sources has been done by specialists instead of scientists working on indicators or models with poor knowledge on the different data sources, which is an important advantage of the integrated database. Also, users of the data only have to retrieve data from one source instead of different sources. A disadvantage of having one integrated database from the data provider point-of-view is that the data provider has to maintain and oversee a large database with data from different domains instead of a small database requiring knowledge from one domain. This implies that data management needs to be done by more than one person and different data-providers need to interact closely for maintenance, support and extension.

The use of Semantic-Rich Development Architecture (SeRiDA; Section 5.3.2) for traversing across programming paradigms (relational databases, object-oriented programming and ontologies) allows the programmers to benefit from the strengths of each of programming paradigms, and not having to maintain the same conceptual schema in at least two places (the database schema and the data accessing codes). In SEAMLESS, we adopted an explicit process to specify an upper data structure (as an ontology), that was translated through SeRiDA into a database schema and the appropriate source code for retrieving and storing data. This allowed the domain scientists to focus on the actual challenge of domain modelling, instead of details of technical implementation in different programming paradigms. Finally, the database is running as a central repository that supports access rights, ensuring safety and consistency.

5.6. Conclusions

The integrated database on European agricultural systems can support policy evaluation and assessment through providing indicators and model inputs for integrated assessment. The integrated database contains data on cropping patterns, production, farm structural data, soil and climate conditions, current agricultural management and policy information and can be extended with more datasets. The database has been used by the models available in the SEAMLESS project, i.e., a dynamic cropping system model, a bio-economic farm model, an econometric model and an agricultural sector model and can be linked to other models or indicators as required. Data on European agricultural management are absent, but essential for the database and exploiting the modeling capabilities of SEAMLESS. The data on current agricultural management is only available for 16 regions in Europe due to time and budget constraints in the collection of data. A systematic and institutional arrangement at European level is needed to complete and to regularly maintain this data set.

The database has the advantages that (i) several data sources are available on one server; (ii) that difficult questions of data integration and consistency have been solved by specialists familiar with the original data sources and (iii) the pre-processing of the original data sources is already done. We aim to make the database available for non-commercial use.

The integration of different data sources into one database is a difficult and time consuming task (Gruber, 1993; Holsapple and Joshi, 2002), as we experienced in our collaborative process to derive one shared ontology. Such a collaborative and time-consuming process of ontology development is required to derive a schema that integrates a range of data sources from different domains specified at different spatial and temporal scales and to avoid inconsistencies and ambiguities in the meaning and definition of concepts across data sources. The explicit and iterative process of ontology development forced us to focus on the domain knowledge and the consistent and coherent linkage of the different data sources. This process could be potentially useful for extending the database on European agricultural systems with more data sources or to integrate other data sources.

We anticipate the database to be of interest for information specialists and systems analysts in the agri-environmental domain. They can derive or calculate policy relevant information. The chapter also described the method to arrive at an integrated database, which we think can be transferred to attempts in other projects and domains.

Chapter 6. Linking models for assessing agricultural land use change

Abstract

The ex-ante assessment of the likely impacts of policy changes and technological innovations on agriculture can provide insight in policy effects on land use and other resources and inform discussion on the desirability of such changes. Integrated Assessment and Modeling (IAM) is an approach that can be used for ex-ante assessment. It may combine several quantitative models representing different processes and scales into a framework for integrated assessment to allow for multi scale analysis of environmental, economic and social issues. IAM is a challenging task as models from different disciplines have a different representation of data, space and time. The aim of this chapter is to describe our strategy to methodologically, semantically and technically integrate a chain of models from different domains to asses land use changes. The models that were linked are based on different modelling techniques (e.g. optimization, simulation, estimation) and operate on different time and spatial scales. The methodological integration to ensure consistent linkage of simulated processes and scales required modellers representing the different models to clarify the data exchanged and interlinking of modeling methodologies across scales. For semantic integration, ontologies provided a way to rigorously define conceptualizations that can be easily shared between various disciplines. Finally, for technical integration, OpenMI was used and supplemented with the information from ontologies. In our case, explicitly tackling the challenge of semantic, methodological and technical integration of models forced researchers to clarify the assumptions of their model interfaces, helped to document the model linkage and to efficiently run models together. The linked models can now easily be used for integrated assessments of policy changes, technological innovations and societal and biophysical changes.

Sander Janssen, Ioannis N. Athanasiadis, Irina Bezlepkina, Rob Knapen, Hongtao Li, Ignacio Pérez Domínguez, Andrea Emilio Rizzoli and Martin K. van Ittersum, Linking models for assessing agricultural land use change, submitted to Computers and Electronics in Agriculture.

6.1. Introduction

Agriculture uses more than 40% of the European land. Changes in agriculture due to policies or technological innovations are likely to have a big impact on European land use and other natural resources. Increasingly agricultural and environmental policies aim at promoting natural resource quality in addition to traditional aims such as economic viability of farms. Ex-ante assessment of the likely impacts of policy changes and technological innovations on agriculture can provide insight in policy effects on land use and natural resources and inform discussion on the desirability of such changes.

Integrated Asessment (IA) is a method proposed by research for ex-ante analysis of the impacts of policy changes and technological innovations on agriculture. IA is defined by Rotmans and Van Asselt (1996) as an interdisciplinary and participatory process of combining, interpreting and communicating knowledge from diverse scientific disciplines to allow a better understanding of complex phenomena. Integrated Assessment and Modelling (IAM) is based on quantitative analysis involving the use of different modelling tools (Harris, 2002; Parker et al., 2002; Letcher et al., 2007). One particular challenge for IAM is to effectively transfer multi-disciplinary scientific and socio-cultural knowledge to an increasingly participatory policy domain (Oxley and ApSimon, 2007). Different types of IAM tools exist, e.g. meta modelling, Bayesian networks, agent-based systems and linking of comprehensive models into model chains. This chapter focuses on this latter IAM approach, as frequently employed for assessing land use changes (Verburg et al., 2006), e.g. ATEAM (Rounsevell et al., 2005), EURURALIS (Van Meijl et al., 2006) and SENSOR (Helming et al., 2008).

The land use modeling community has been one of the early adopters of IAM, recognizing that a single disciplinary modelling approach falls short of capturing the growing complexity in sustainable land use. Mono-disciplinary models cover only a few processes from a single domain, be it economic, agricultural or environmental and lack descriptions of some relevant processes. These models generally do not cover the relevant multiple scales to handle all assessment questions. Mono-disciplinary models can complement each other and thereby provide comprehensive and balanced assessments across scales (Van Tongeren et al., 2001; Pérez Domínguez et al., 2008). Land use models are 'highly evolved, readily available and easy to use (Kok et al., 2007)' and are therefore suitable to be linked in model chains.

In order to arrive at an operational model chain for applications in integrated assessment procedures, semantic, methodological and technical integration of models is required. To show why different types of integration are required in IA studies, we present here the model linking of a set of (agricultural) models from different domains to arrive at a model chain that can be re-used for a range of IA questions. First, we address the meaning and content of the methodological, semantic and technical integration by providing an overview of relevant literature. Second, we show how a model chain can be described comprehensively with these concepts thereby becoming re-usable for a range of IA questions. In this chapter, we do not describe an application of the models, the integration of data-sources for such a model chain (Chapter 5), or the definition of scenarios for such a model chain (Chapter 7).

We present the model linking as achieved in the integrated project System for Environmental and Agricultural Modelling; Linking European Science and Society (SEAMLESS) (Van Ittersum et al., 2008) for an agronomic model, an agronomic-economic model and economic models. Ultimately these linked models provide a means to achieve up-scaling and the interdisciplinary assessment of agricultural and agri-environmental policies, technological innovations and societal and biophysical trends, that would not be possible with the individual models.

Section 6.2 defines in more detail semantic, methodological and technical integration for this chapter by providing an overview of relevant literature. Section 6.3 introduces the SEAMLESS-Integrated Framework (IF) and the choices for theories and technologies made in integration for SEAMLESS-IF. Subsequently, the semantic, methodological and technical integration as achieved in SEAMLESS-IF is presented in Section 6.4. Section 6.5 discusses the lessons learned with respect to and during the integration and the further use of the model chain. Finally, some conclusions are provided.

6.2. Model linking and integration

6.2.1. Semantic integration

Ambiguous terminology and a lack of shared understanding between disciplines have often been mentioned as important obstacles in integrated assessments (Jakobsen and McLaughlin, 2004; Scholten et al., 2007). Semantic integration means speaking a common language and achieving a shared understanding between all models and modellers working together. This is a crucial challenge for any integrated modeling project (Jakobsen and McLaughlin, 2004; Tress et al., 2007; Hinkel, 2008; Scholten, 2008), as it is provides the building blocks for the technical and methodological integration and as it ensures the consistency and transparency in definitions and terms required for the methodological and technical integration.
Very few practical applications of possible methods for semantic integration of models could be found in literature, with the exceptions of Hinkel (2008) and Scholten (2008). Possible methods are variable mapping, mathematical formalism, concept maps and ontologies. Variable mapping is an ad hoc process of investigating which variables could be exchanged between models and then mapping them to each other. As variable mapping is not formalized and ad hoc, it remains a black box. Hinkel (2008) uses mathematical formalism as a methodology to firstly align terminology between models and secondly the model equations across models and uses this to undertake a semantic integration for model linking in a number of modelling projects. One disadvantage of using mathematical formalism is as Hinkel (2008) mentioned, that non-modellers need explanation and training in order to be involved. Concept maps (Novak and Cañas, 2006) are graphs representing knowledge, in which concepts are expressed in circles and relationships are shown by lines connecting two concepts. Finally, like concept maps ontologies

consist of concepts and relationships between concepts (Antoniou and Van Harmelen, 2004), but these have the advantage that these are expressed in man readable and machine understandable format.

6.2.2. Methodological integration

The methodological integration focuses on aligning different scientific methodologies and identifying required model improvements necessary for meaningful linkage. Methodological modelling is a vital first step to facilitate communication between modellers, non-modelling researchers and stakeholders (Liu et al., 2008). Good practice guidelines (Refsgaard and Henriksen, 2004; Jakeman et al., 2006; Scholten, 2008) exist for methodological development of a model in all steps of model building for a mono-disciplinary model. Methodological challenges for land use models are to model appropriate '(1) level of analysis; (2) cross scale dynamics; (3) driving forces; (4) spatial interaction and neighbourhood effects; (5) temporal dynamics; and (6) level of integration (Verburg et al., 2004).'

Methodological integration in a model chain requires that the data produced by one model are a meaningful input to another model, usually operating at a different temporal and spatial scale. Different process descriptions represented by different modelling techniques (e.g. optimization, estimation, simulation) affect the interpretation of outputs, as processes can be modelled in many different ways, modelling techniques model different outputs and models have different levels of uncertainty.

Methodological integration deals with calculations of a concept out of other concepts or converting one concept into another concept. Spatial and time scales are crossed through these calculations and conversions, e.g. moving from daily estimates to an estimate for one or several years or from the representative farms to regions or provinces. These calculations describe the behavior of the system (e.g. linked models) in mathematical terms and often include strong assumptions. In methodological integration all the calculations have to become explicit, preferably in mathematical terms. An example can be found in Hinkel (2009) based on mathematical formalism.

6.2.3. Technical integration

6.2.3.1. Types of coupling

Technical integration with respect to linking models in model chains is classified by Brandmeyer and Karimi (2000). They distinguish five hierarchical types of model coupling for IAM tools: i. one way data transfer, in which output files of one model are used as input files to the next model, ii. loose coupling, in which two models automatically send each other data, iii. shared coupling, in which two models are executed through a common graphical user interface and common data storage, iv. joined coupling, in which a shared graphical user interface and data storage is used and in which one model embeds other model(s) and v. tool coupling, in which models are linked together in a modelling framework with a common graphical user interface and data storage.

Tool coupling requires the closest integration of models and it facilitates re-use of models, making it the most flexible type of technical integration when dealing with model chains. A tool coupling requires first a central repository for data storage for all models, scenarios and data sources (Chapter 5). Second, tool coupling requires one graphical user interface from which all models can be parameterized and executed (Wien et al., 2009). Third, tool coupling requires a modelling framework, that supports the execution of models in a model chain (Hillyer et al., 2003; Rahman et al., 2003; Moore and Tindall, 2005). This chapter further only discusses the importance of modelling frameworks, since modelling frameworks are most relevant to the actual model linking and can have effects on the set-up of the models.

6.2.3.2. Modelling frameworks

The linking of models assumes the exchange of data between models. Model linking is especially challenging when modelers from different domains use different programming languages, tend to stick to their own pre-cooked solutions and when the best type of model linking they can achieve is only through the exchange of data files. As long as model linking is a one time exercise, it is still possible to use an ad hoc file-based exchange, but when the linked models must be used to analyze a large number of scenarios, then the file-based exchange becomes excessively laborious, error-prone and non-repeatable. Automated, documented and standardized model linking in a modelling framework is definitely preferred and recommended. Some available modelling frameworks exist. Open Modelling Interface and Environment (OpenMI - Moore and Tindall, 2005) is a software standard for dynamically linking models at runtime, which can potentially be used in many domains, but is currently mainly applied in the water domain. TIME (Rahman et al., 2003) is, like OpenMI, a generic computational framework for building and executing models that may be applicable across domains. ModCom (Hillyer et al., 2003) is used for linking biophysical process-based models in crop growth simulation. Moore et al (2007) propose the Common Modeling Protocol which nests dynamic models in a hierarchy with a common interface on top and also focuses on dynamic and biophysical models. Computational frameworks, with a stronger technical instead of methodological focus are lacking for the land use and socio-economic models, although frameworks like OpenMI, TIME or ModCom might be useful.

6.3. Integration methods of choice for SEAMLESS-IF

6.3.1. Purpose of model linking

In an IAM research project codenamed SEAMLESS (Van Ittersum et al., 2008) the causal chain of impacts of farmers' actions is modelled by linking and combining field, farm, regional and market models. When farmers' decisions on land use allocation are aggregated to a higher scale through the level of production, this may have profound market impacts and, hence, in turn influence agricultural commodity prices. Moreover, farmers' decisions in land allocation directly impact the environment through their crop choices (e.g. maize instead of wheat) and through their use of inputs (e.g. nitrogen fertilizer causing nitrogen leaching). Therefore it would not be adequate to study the land allocation patterns at the farm scale (e.g. through bio-economic farm models) without taking into account also the market (e.g. trade agreements and policy changes by the European Union through partial equilibrium market models) and field scale (e.g. technological innovation and integrated production by farmers through cropping system models).

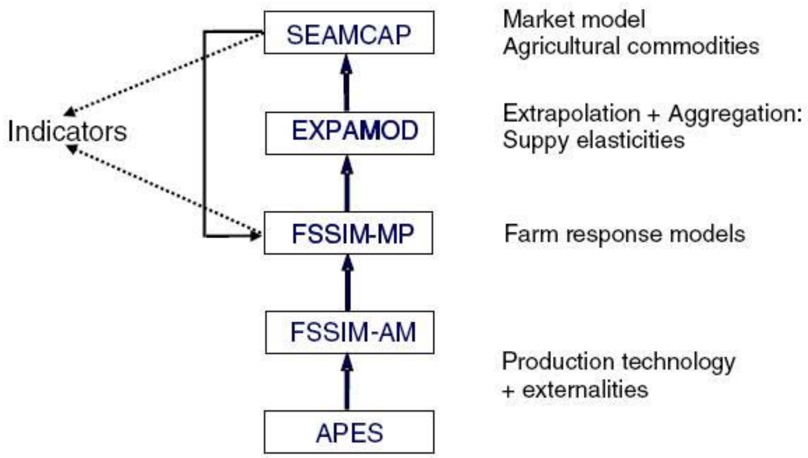


Figure 6.1. The models in SEAMLESS (after Van Ittersum, et al. (2008)). APES: Agricultural Production and Externalities Simulator; FSSIM-AM: Farm System SIMulator-Agricultural Management; FSSIM-MP : FSSIM-Mathematical Programming; EXPAMOD: EXtraPolation and Aggregation MODel and SEAMCAP: adapted version of the Common Agricultural Policy Regional Impact (CAPRI) model.

The SEAMLESS-Integrated Framework (IF) (Van Ittersum et al., 2008) has been developed to assess the sustainability of agricultural systems in the European Union at multiple scales. In SEAMLESS-IF methods to methodologically and technically link different models (Fig. 6.1) are used to facilitate the re-usability of models for different purposes (Van Ittersum et al., 2008). By linking field-farm-market models in SEAMLESS-IF, the land use changes can be analysed at multiple levels through a selected number of economic, environmental and social indicators, accounting for the impacts of farm responses that could not be analysed by using only the individual models as stand-alone tools. An example of such a question is ‘what are the impacts of implementations of the Nitrate Directive on farm income, on-farm labour, non-point source pollution and resource use in the European Union and in the regions Poitou Charentes (France) and Flevoland (Netherlands)?' With respect to this question, a bio-economic farm model can provide an estimate of the impacts on farm income in either Poitou Charentes or Flevoland, while a cropping system model can estimate the impacts on non-point source solution and resource use. A market model can estimate the impacts on farm income, trade and markets in the entire European Union. When these models are linked, the impacts can be calculated at all scales and for all indicators in a consistent manner.

6.3.2. Semantic integration

SEAMLESS applied ontologies for semantic integration, since i. ontologies are in machine readable format, e.g. the Web Ontology Language (McGuinness and Van Harmelen, 2004), ii. ontologies are based on first order logic upon which a computer can reason, iii. the developed ontologies are a separate product independent of the models to which they are applied and iv. both modellers and non-modellers can contribute to the ontology development. Recently, a branch of computer science focusing on knowledge representation and engineering has introduced ontologies as a means to provide a specification of a conceptualization (Gruber, 1993). In the context of integrated modeling, ontologies can be useful for defining data structures describing model inputs and outputs (Athanasiadis et al., 2006; Rizzoli et al., 2008; Scholten, 2008).

Only the specification of the interfaces between the models has to adhere to the shared ontology, while the internal specification of the knowledge in the model does not have to adhere to the shared ontology (Gruber, 1993). An ontology separates knowledge captured in the model from the actual implementation in a programming language e.g. JavaTM, FORTRAN, Matlab, or STATA (Gruber, 1993) and thus ensure that knowledge is not hidden in programming languages (Athanasiadis et al., 2006). Ontologies help to formalize the knowledge exchanged between models, thus facilitating re-usability and exchangeability of model knowledge (Rizzoli et al., 2008), supporting portability (Gruber, 1993) and working in a multi-disciplinary environment.

6.3.3. Methodological integration

The models in SEAMLESS are a cropping system model APES, a bio-economic farm model FSSIM, an econometric estimation model EXPAMOD, and a partial equilibrium optimization model SEAMCAP. Each of these operate on different spatial and temporal scale and is based on different modelling techniques. The cropping system model Agricultural Production and Externalities Simulator (APES) operates at the field systems level, and represents one hectare (or a point) (Donatelli et al., 2009). On the basis of agricultural activities, soil and climate data, APES simulates the yield and environmental effects resulting from those activities. It presently includes components for simulation of crops, grassland, vineyards and agroforestry. Examples of other components are those that simulate water balances in the soil, carbon-nitrogen dynamics in the soil, the fate of pesticides and agricultural management. It is a dynamic simulation and it usually simulates a period of 10 to 25 years with a daily time step.

The bio-economic farm model and partial equilibrium optimization model are both optimization models based on mathematical programming techniques. These models are built based on assumptions with respect to the functioning of economic agents, i.e. farms or market parties at continental scale. These models are comparative static, i.e. they have no interdependence of outcomes across years, and model results represent the equilibrium situation for a year. The Farm System Simulator (FSSIM) is a bio-economic farm model developed to assess the economic and ecological impacts of agricultural and environmental policies and technological innovations (Louhichi et al., 2009). A bio-economic farm model links decisions on management of farm's resources to current and alternative production possibilities describing input-output relationships and associated externalities (Chapter2).

SEAMCAP is a variant of the Common Agricultural Policy Regionalised Impact (CAPRI) model adapted for inclusion in SEAMLESS-IF (Britz et al., 2009). CAPRI is a spatial economic model that makes use of non linear mathematical programming tools to maximise regional agricultural income. It explicit considers Common Agricultural Policy instruments in an open economy and price interactions with other regions of the world are taken into account (Heckelei and Britz, 2001). Major outputs of the market module include bilateral trade flows, market balances and producer and consumer prices for the products and world country aggregates.

Finally, the econometric estimation model EXtraPolation and Aggregation MODel (EXPAMOD) is an econometric meta-model describing price-production responses of farms given specific farm resources and biophysical characteristics (Pérez Domínguez et al., 2009). EXPAMOD accounts for land use changes via production volume. After the calculations done in EXPAMOD, the regional supply modules of the market model SEAMCAP can behave like a representative aggregate of the FSSIM models of the same region. The extrapolation routine operates with prices, farm characteristics and regional biophysical characteristics obtained from other models or European databases. The output of EXPAMOD are price-supply elasticities on which the regional supply functions in the market model SEAMCAP are calibrated.

6.3.4. Technical integration

In the development of SEAMLESS-IF The Open Modelling Interface and Environment (OpenMI; Moore and Tindall, 2005) was applied to link the models at run time into a model chain. OpenMI was chosen as it can in principal be applied to models from all domains and as it is a standard instead of an implemented modelling framework in a specific programming language. OpenMI represents a standard for the definition of the interface of a software component (Gregersen et al., 2007). In principle, a model that complies to the OpenMI standard and is designed as a software component can, without any programming, be configured to exchange data at run-time with other OpenMI-compliant models (Gregersen et al., 2007). The OpenMI standard aims at an easy migration of existing models to comply with the standard, without the need for re-implementing the whole models. To achieve such

an easy migration, wrappers are proposed that comply with the OpenMI-standard and that leave the model internally unchanged with respect to specification and programming paradigm (Gregersen et al., 2007).
The OpenMI standard version 1.4 is based on a pull-approach in which the last model in the chain pulls its outputs from other models in the chain by calling "getValues()"-methods, which means requesting outputs from a model or data source (Moore et al., 2007). Before "getValues()"-calls can be successfully enacted at run time, the links between the two OpenMI-compliant models need to be defined by the modeller by specifying so-called "Links." These links define the output item of a model that is linked to an input item of another model.

6.4. Integration in SEAMLESS

This section describes the results of the integration efforts to link the models (i.e. APES, FSSIM, EXPAMOD and SEAMCAP). It starts with the semantic integration by describing three ontologies (e.g. crop-product, elasticity and activity) crucial to understand the model linkage. The concepts in these ontologies provide the building blocks that are subsequently used to describe the methodological and technical integration. The calculations described in the methodological integration are the responsibility of the models or scaling procedures in between models. For methodological integration, the calculations to link on the one hand FSSIM, EXPAMOD and SEAMCAP and on the other hand FSSIM and APES are described, without describing all calculations of the models in detail. The technical integration describes the use of the ontologies derived in the semantic integration and the impact of OpenMI on the models.

6.4.1. Semantic integration

6.4.1.1. Crop-product ontology (APES-FSSIM-EXPAMOD - SEAMCAP)

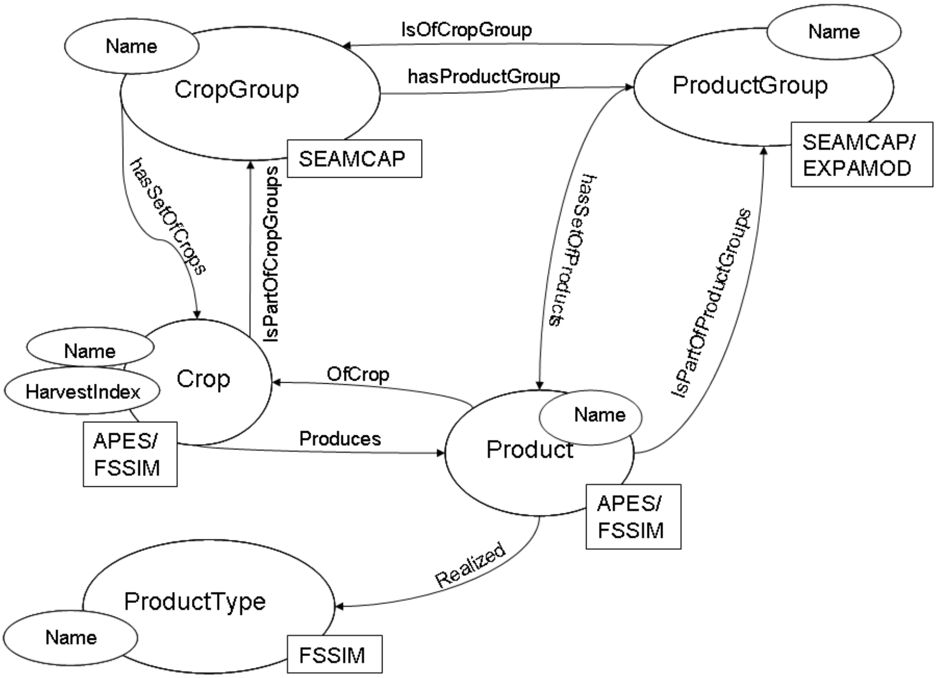


Figure 6.2. The Crop-Product ontology showing the relationships (arrows) between the concepts Crop, Product, ProductType, CropGroup and ProductGroup (ellipses) and their properties (small ellipses). The models using the concept are indicated in the boxes. The figure can be read along the lines.

In the initial discussion, it appeared that all models dealt with cropped areas and used Crops and Products produced by these crops as concepts. Each of the models referred to these concepts, although sometimes with different names (e.g. crop in APES, crop in FSSIM and activity group in SEAMCAP). It seemed that the ontology could thus be simple, only referring to Crop and Product-concepts and relationships between them. This simple structure proved to be invalid, when confronted with the list of Crops and Products used by each of the models. The reason for models to use different groupings of Crops is that they have originally

Table 6.1. Examples of crops, products, product types, crop groups and product groups.

a. Example of crop groups with associated crops

CropGroup		Crop
Wheat	hasSetOfCrops	WinterSoftWheat SpringSoftWheat WinterDurumWheat SpringDurumWheat
Potatoes	hasSetOfCrops	Potatoes
Textiles	hasSetOfCrops	Flax Hemp

b. Example of products, product types and crops.

ProductType		Product IsProduced By		Crop
Straw	Realises	WinterSoftWheatStraw SpringSoftWheatStraw WinterDurumWheat Straw SpringDurumWheat Straw	Is produced by	WinterSoftWheat SpringSoftWheat WinterDurum Wheat SpringDurum Wheat
Grain	Realises	WinterSoftWheatGrain SpringSoftWheatGrain WinterDurumWheat Grain SpringDurumWheat Grain	Is produced by	WinterSoftWheat SpringSoftWheat WinterDurum Wheat SpringDurum Wheat
Ware	Realises	PotatoesWare FlaxWare HempWare	Is produced by	Potatoes Flax Hemp

c. Example of products and product groups.

Product		ProductGroup
WinterSoftWheatStraw SpringSoftWheatStraw WinterDurumWheatStraw SpringDurumWheatStraw	isPartofGroup	Straw
WinterSoftWheatGrain	isPartofGroup	SoftWheat
SpringSoftWheatGrain		
WinterDurumWheatGrain	isPartofGroup	DurumWheat
SpringDurumWheatGrain		
PotatoesWare	isPartofGroup	Potatoes
FlaxWare HempWare	isPartofGroup	Textiles

been developed for different purposes and scales. For example APES models crop growth for a field whereas SEAMCAP models markets of crop commodities.

Consequently, these lists of Crops and Products were further investigated, and a suitable structure was found for the ontology as shown in Fig. 6.2. In this ontology, each Crop produces one or more Products, which are realized by a ProductType. Products and Crops can be grouped together in ProductGroups and CropGroups. These ProductGroups and CropGroups are an input to the higher scale models SEAMCAP and EXPAMOD, that operate on the region and market scale, while the Crops, Products and ProductTypes are used by the lower scale models APES and FSSIM, that operate on the field and farm scale.

An example of the data associated with the Crop-Product Ontology is given in Table 6.1a, b and c. From Table 6.1a, it can be read that the wheat CropGroup has a set of crops WinterSoftWheat, SpringSoftWheat, WinterDurumWheat and SpringWinterWheat, while the potato CropGroup has only one crop, which is Potatoes. Similarly, Table 6.1b displays that the Straw ProductType realizes the products WinterSoftWheatStraw, SpringSoftWheatStraw, WinterDurumWheatStraw and SpringDurumWheatStraw, while the crops Potatoes, Flax and Hemp produce the products WarePotatoes, WareHemp and WareFlax.

6.4.1.2. Price-elasticity ontology (FSSIM-EXPAMOD-SEAMCAP)

The unambiguous definition of crops and products as presented in the previous section 6.4.1.1 is used to define other relevant concepts for the links between the models. Crucial concepts for the linking between FSSIM, EXPAMOD and SEAMCAP are price elasticity and supply response (Fig. 6.3). The concept price elasticity is the output of EXPAMOD and the input to SEAMCAP, whereas supply response is the output of FSSIM and the input to EXPAMOD.

A price elasticity is the percentage change in supply as a result of one percent change in price. Price elasticity in the ontology (Fig. 6.3) has three dimensions, as it refers to two ProductGroups through ‘to’- and ‘from’-relationships and a NUTS2-

region. NUTS stands for Nomenclature of Territorial Units for Statistics (EC, 2008b) and the NUTS2 level corresponds to provinces in most countries.

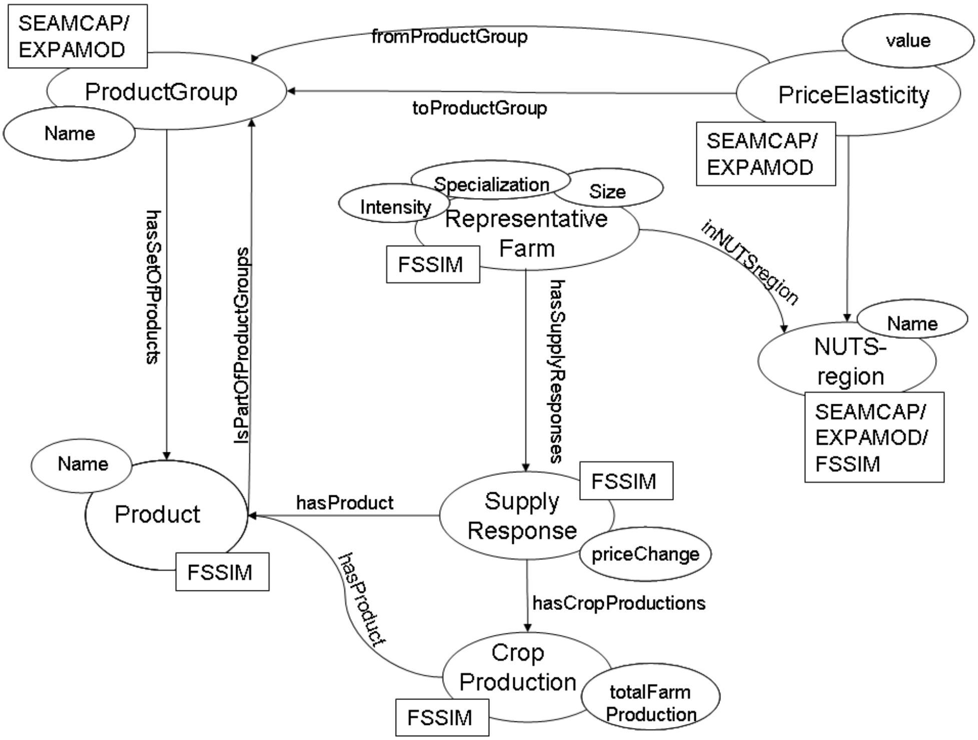


Figure 6.3. Price elasticity ontology. The large ellipses show concepts, relationships can be read along the arrows, the small ellipses are data-properties of the concepts, and the boxes indicate the models using the concept.

Supply responses describe the responses of representative farms to changes in prices (Fig. 6.3). Each representative farm (Chapter 5) refers to sets of supply responses. Each supply response captures the price change for a product and multiple CropProductions in response to the price change. One CropProduction is the total farm production of a product for the representative farm.

6.4.1.3. Activity ontology (APES-FSSIM)

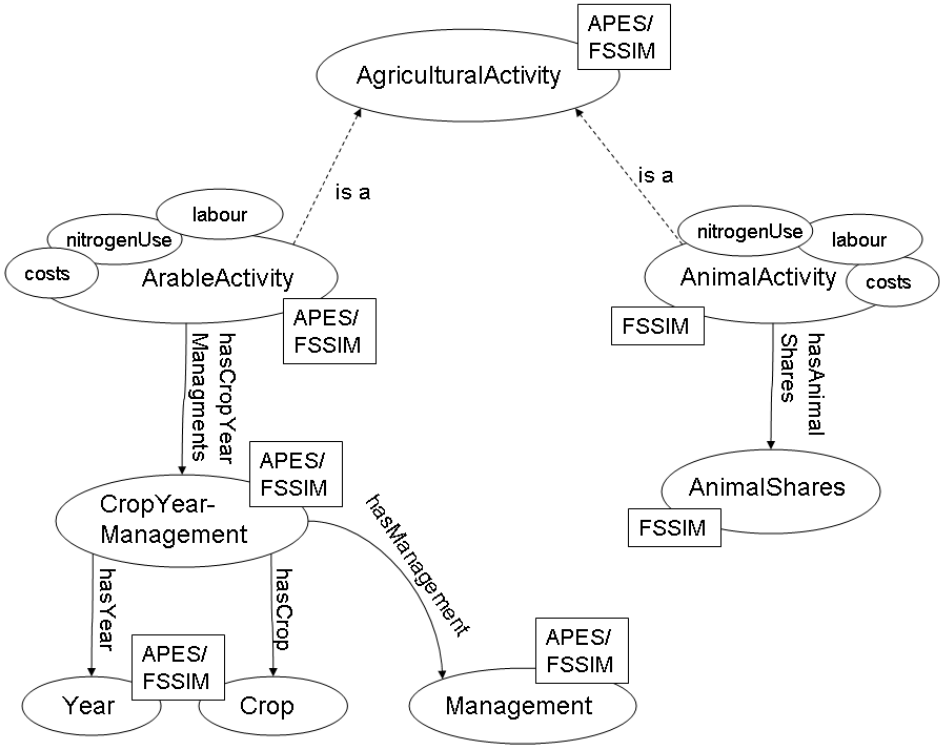


Figure 6.4. Part of the activity ontology. The large ellipses show concepts, relationships can be read along the arrows, the small ellipses are data-properties of the concepts, the dotted arrows indicate an 'is a'-relationship and the boxes indicate the models using the concept.

Farmers have many different production possibilities on their farm. They might decide to grow crops, plant trees, or have livestock. Within these three basic choices, many more choices exist between different crops, different trees or different types of animals. Also the intensity and type of management of a crop, animal or tree might change. To capture the broad range of options available to the farmer and make the linking between the models APES and FSSIM explicit, the activity ontology was created. Figure 6.4 shows a small part of this activity ontology related to arable and animal activities and some illustrative relationships. According to this ontology, farmers can have on their farms arable activities and/or animal activities. An arable activity entails several CropYearManagements that capture the unique combination of a crop (Fig. 6.4), a year and management. For example, in year 1 the farmer grows potatoes with an intensive management, while in year 2 he grows barley with an extensive management. Together potato and barley form a two year rotation; the management within this rotation differs between the crops from intensive to extensive.

Both the construction and selection of agricultural activities for a specific farm type is done by FSSIM (Section 6.3.3). APES (Section 6.3.3) operates on the arable activity by simulating for each activity the succession of CropYearManagements over time and providing the yields and environmental effects as an output. The arable activity is thus a shared concept between FSSIM and APES. Different models use different properties of a concept, as is shown in Table 6.2. Whereas the variable costs of an arable activity are of relevance to FSSIM, the sowing date and nitrogen use are of relevance to APES. The activity ontology captures the shared concepts used by the models and allows them to work on different parts of this shared concept (Table 6.2).

Table 6.2. Example of arable activities specified according to the activity ontology for the Auvergne Region in France

Model		APES	APES	FSSIM	APES/ FSSIM	APES
		Nitrogen use	Water use	Variable costs	Grain yield	Sowing week
Year	Crop	kg N/ha	m3/ha	euro/ha	Tonnes /ha	Week number
Activity Identifier = 1364						
1	Maize	100	0	350	6.0	14
2	Maize	200	1000	696	10.0	14
3	Sunflower	40	0	288	2.0	15
Activity Identifier = 1196						
1	Maize	100	0	350	6.0	14
2	Softwheat	90	0	318	4.0	43
3	Winter Barley	100	0	300	4.0	42

6.4.2. Methodological integration

For methodological integration two calculations representing scaling procedures are crucial. First, a calculation is required to aggregate supply responses at the farm scale in FSSIM to price elasticities of product groups at market scale for SEAMCAP. Second, the field scale APES model and the farm scale FSSIM model are interlinked through agricultural activities and upscaling procedures are required to move from field and annual simulations to averages across years and activities.

FSSIM provides supply responses (Fig. 6.3) at farm scale. These supply responses are the results of multiple runs of FSSIM with changed product prices (Fig. 6.3) (Pérez Domínguez et al., 2009). Each change in product price leads to another optimal solution in FSSIM, and thus to a changed supply of products. Through the multiple runs, FSSIM generates one price supply response for each product on each representative farm. EXPAMOD uses the supply responses as observations in its estimation procedures per product.

In the estimation, the supply responses are regressed on properties of the representative farm (e.g. machinery, buildings, size, climate and soil conditions) (Pérez Domínguez et al., 2009). The function obtained through this regression can subsequently be used to predict the supply responses of representative farms in regions, for which FSSIM has not been run. In both the regression and extrapolation the properties of the representative farm are multiplied by the weighing factor. This weighing factor is calculated as the area of the farm divided by the area of all representative farms in the NUTS2-region, under the assumption that the representative farms cover 100% of the region. Through the regression and extrapolation price elasticities per product in a region are derived. To derive the price elasticitices per product group as needed by SEAMCAP (Section 6.3.3), the price elasticities per product are averaged with the quantity shares of each product in total production.

APES receives as input data from FSSIM the specification of an arable activity and the specification of an agri-environmental zone. The arable activity has a limited rotation length from 1 to 8 years. The agri-environmental zone is associated with soil data, which is constant over time, and climate data for a period of 10 to 25 years. APES starts a simulation on the first day and ends with the last day of the climate data. The arable activity is iteratively run over the simulation period. This implies that year 1 of the arable activity coincides with year 1 of the climate data and year 1 of the arable activity is run again after the last year in the arable activity, till the final year in the climate data has been reached. If an arable activity has a rotation length of 3 years (e.g. soft wheat, potato, sugar beet) and 25 years of climate data are available, then the arable activity is run 8 times with one additional run for the first crop during the simulation. Implicitly this assumes that years in the activity are associated with years in the climate data, e.g. activity year 1 with year 1 in the climate data and activity year 2 with year 2 in the climate data. Changes of values of soil variables are simulated by APES.

During its simulation, APES produces multiple estimates of yield and externalities for an arable activity, depending on the number of runs of the activity. FSSIM requires only one single estimate and a standard deviation of the yield for each product of the activity and one estimate and a standard deviation of each externality (e.g. nitrate leaching or soil erosion) for the activity. To obtain these single estimates, the yearly simulated estimates are averaged over the number of years. For example, with a 3 year activity and a 25 years of climate data, 9 estimates of the yield of product grain of soft wheat grown in the year 1 of the activity and 25 estimates of the nitrate leaching for the activity are obtained. These 9 estimates and 25 estimates are then averaged to produce one estimate for the yield of grain of soft wheat and one estimate of nitrate leaching.

6.4.3. Technical integration

With the shared concepts clarified in the semantic integration and the conversion of these concepts clarified in the methodological integration, a modelling framework (Fig. 6.5) was designed that supports the execution of the models APES, FSSIM, EXPAMOD and SEAMCAP in different model chains (Ewert et al., 2009). In the modelling framework the shared ontology achieved in the semantic integration is used to provide a common access to the data layer and to define the links between models as OpenMI components (Fig. 6.5). The ontologies in Web Ontology Language (OWL) were automatically translated into a relational database schema according to the specifications of the Semantic-Rich Development Architecture (SeRiDA) (Athanasiadis et al., 2007b; Athanasiadis et al., 2007c). The SeRiDA acts as a bridge between different programming languages, e.g. object-oriented programming, relational databases and ontologies (Athanasiadis et al., 2007b). The relational database schema was made accessible to the modelling framework and models through Entity Java Beans (DeMichiel and Keith, 2006), which can be used to develop the wrappers for the models as OpenMI components. The models are linked to the database through Hibernate (JBOSS, 2008). The integrated database is running on a PostgreSQL database server (PostgreSQL, 2008).

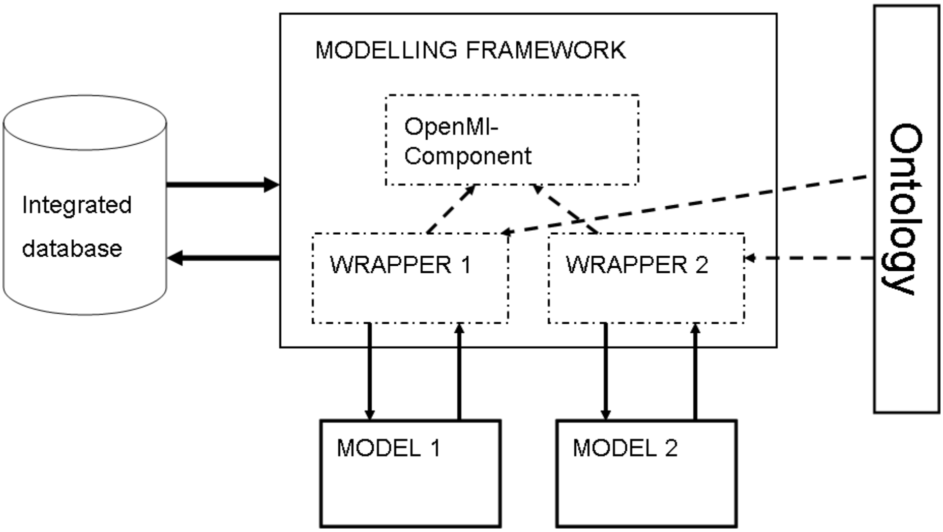


Figure 6.5. Architecture of the modelling framework in SEAMLESS (source: Wien et al., 2009)

The models all remain programmed in their native programming language (e.g. GAMS for EXPAMOD, SEAMCAP and part of FSSIM; C# for APES and JavaTM for the other part of FSSIM). Each model is wrapped through a wrapper that translates the data into an appropriate format for the model, executes the model, and translates the model output data into a suitable format for the framework. The wrappers of the models have been developed as OpenMI-components, which implies that the models themselves are not aware of or affected by the OpenMI-standard. An extension of the OpenMI standard was required to make it usable for the SEAMLESS model chain (Knapen et al., 2009). This extension to capture the complex data types of the models implied that data exchanged between models are objects or complex data structures and not primitive data types (e.g. float, integer, string, character) like in the current OpenMI standard. For models with complex data types from different disciplines such an extension of the OpenMI standard is required for OpenMI to be relevant.

6.5. Discussion

6.5.1. Semantic Integration

Through semantic, methodological and technical integration, we achieved a chain of agricultural models to assess the impacts of policy and technology changes on European agricultural systems. We experienced benefits of our integration approach, which will be described in this section. Through the use of shared ontologies, we managed to explicitly establish a shared understanding between the modelers and their models. In our case, the use of ontologies forced researchers to clarify the assumptions of their model interfaces and to set forth parts of their modeling knowledge, typically kept within their models. An important benefit of this approach is that knowledge on model linking is not solely contained in the model source code or in the modelers mind, but is documented as part of the framework and can help to explain model linkages to non-modelers. This opens up the model linking to scrutiny from a wider community than just the modelers involved in the linking.

In our case study, a shared ontology has been developed for model linking of four models, as demonstrated through the examples in the previous section. The ontology is re-usable independently of the models and documents the concepts used and agreed upon for model linking. Ontologies from SEAMLESS are available on http://delivered.seamless-ip.org:8060/browser/zul/main.zhtml or they can be referred to through the Uniform Resource Identifier (URI) each concept is linked to, for example concept Crop can be found on http://ontologies.seamless-ip.org/crop.owl#Crop. Other modelers can build upon, extend and improve the ontologies. The ontologies are supplied with metadata (Brilhante et al., 2006) and browsable through a simple search tool, in order to facilitate their re-use.

While on the internet there is a growing number of ready ontologies available, for example the core software ontology (Gangemi et al., 2008), unfortunately many could not be re-used for our model linking tasks, as these were not yet specific to the agricultural domain and not concrete enough. Similarly, it might seem that the ontologies developed for SEAMLESS-IF are specific to the linking of the models APES, FSSIM, EXPAMOD and SEAMCAP. Although the ontologies have been made with the aim of linking these specific models, they exist independently of the models and there exists no concept like "SEAMCAP" or "FSSIM" in any of the ontologies. As a true test of the genericity of these ontologies one could re-use them for the linking or developing of models simulating cropping systems, farm responses and market behavior.

6.5.2. Methodological Integration

A first methodological benefit is that we identified calculations to link cropping system models to bio-economic farm models, and bio-economic farm models to partial equilibrium market models in a sensible and consistent manner. These calculations are based on jointly setting parameters of activities and aggregating supply responses to price elasticities through an estimation model. These calculations and links between cropping system models, bio-economic farm models and partial equilibrium models may be re-usable in future research linking these model types, because these links help to cross temporal and spatial scales of the different models and are based on standard outputs of these types of models.

A methodological benefit is that the explicit model linking helped to efficiently (re)run models in model chains. Examples are an application of FSSIM to a large number of regions to assess supply-responses for EXPAMOD (Pérez Domínguez et al., 2009) and an application of APES for a large number of activities to supply yields and environmental effects for FSSIM (Belhouchette and Wery, 2009). Such applications can now easily be repeated for different samples of regions or activities and are easily reproducible in the modelling framework, thereby ensuring scientific transparency and rigor.

For the methodological integration no generic method was used to link the different models. Although many different loosely or tightly linked models (e.g. Rounsevell et al., 2005; NMP, 2006; Jansson et al., 2008; Verburg et al., 2008) are available, there is no generic method to achieve the methodological integration of a set of models for land use modelling. Relevant aspects (e.g. time and spatial scales, process definitions, modelling techniques) of methodological integration can easily be identified and have been discussed also for our methodological integration. A generic method may facilitate the model linking by providing guidelines and a conceptual framework for scientists to achieve a model linking. An example of such a generic method is found in Letcher et al. (2007) for IAM of water allocation problems, which is based on the nature of interactions between decisions and the hydrological cycle and the assumptions with respect to perfect knowledge or uncertainty. Established scientific theories like hierarchical systems theory (Smith and Sage, 1973) may supplement a generic method, but such theories always represent a perspective of the model linking considered. A thorough review of the available linked models in the land use domain is a useful first step in the development of a more generic method for methodological integration.

6.5.3. Technical Integration

The use of OpenMI and the development of a modelling framework helped to execute the model chain on a computer. Our use of OpenMI demonstrates that OpenMI can be applied to models outside the water domain, as OpenMI facilitated the link between agronomic and economic models. The use of OpenMI had two benefits. First, the definition of data exchanged in Links and getValues() (e.g. outputs) forced models to be specific about their inputs and outputs. Second, wrapping the model as an (OpenMI) component facilitated the definition of models independently of each other, of data sources and of the graphical user interface. The extension of the OpenMI standard version 1.4 to work with complex data types may be of interest to incorporate in future updates of the OpenMI standard (OpenMI, 2009), if the OpenMI standard targets applicability in more different domains and models based on different modelling techniques. OpenMI is based on the use of wrappers that allow to keep the model in its original programming language. Disadvantages are that the wrappers require maintenance and updating with changes in the model and that the model itself is quite distant from OpenMI. This distance may lead to problems in developing the wrapper, if the wrapper-developer and modeler are not the same person.

The ontology achieved in the semantic integration was intensively used in the technical integration by translating it to source code through the SeRiDA-framework. A benefit of a tight link between semantic and technical integration is that modelers are forced to focus on content of their model and not on the implementation of a model into programming language. A second benefit is the explicit separation of data from model specification as is advocated in good modeling practices (Jakeman et al., 2006), allowing to easily validate a model against other data sources. This separation is facilitated through the database schemas which are built on the basis of ontologies (Athanasiadis et al., 2007b) and provide a natural container for data persistence. A disadvantage from the modelling perspective is that the models cannot easily change their input and output data specification, as this first has to be aligned with the ontology in the semantic integration.

6.6. Conclusion

The models APES, FSSIM, EXPAMOD and SEAMCAP are now linked in the modelling framework SEAMLESS-IF. These models allow assessment of the socio-economic, biophysical and environmental impacts of changes in agricultural and environmental policies and technological innovations across spatial and temporal scales. Examples of possible applications at EU, individual region or farm scale are the assessment of the impacts of a trade liberalization as discussed in the frame of the World Trade Organization (Adenäuer and Kuiper, 2009), the introduction of the EU Nitrate Directive (Belhouchette and Wery, 2009), the EU Water Directive, the consequences of increases in bio-fuel production, the changes in production due to high commodity prices and of the introduction of agricultural technologies (e.g. zero-tillage, improved irrigation implements). Our integration effort led to a credible and transparent model linking with an explicit consideration of the concepts (e.g. activities, crops, products, product type, crop group, product group, price elasticity, supply response) and calculations (e.g. parameter calculation of activities and aggregation of supply responses to price elasticity) of relevance implemented in an advanced modelling framework based on OpenMI and semantic modelling.

The subdivision of the integration effort in methodological, semantic and technical aspects was useful to comprehensively consider all aspects of integration and to avoid a bias to one of them. In future research projects that link models, it is advised to first define the semantic and methodological integration, if models are linked that have yet to be developed. If existing models are linked, semantic integration of concepts across models is the most suitable starting point.

Chapter 7. Defining assessment projects and scenarios for policy support: use of ontology in Integrated Assessment and Modelling

Abstract

Integrated Assessment and Modelling (IAM) provides an inter-disciplinary approach to support ex-ante decision-making by combining quantitative models representing different systems and scales into a framework for Integrated Assessment. Scenarios in IAM are developed in the interaction between scientists and stakeholders to explore possible pathways of future development. As IAM typically combines models from different disciplines, there is a clear need for a consistent definition and implementation of scenarios across models, policy problems and scales. This chapter presents such a unified conceptualization for scenario and assessment projects. We demonstrate the use of common ontologies in building this unified conceptualization, e.g. a common ontology on assessment projects and scenarios. The common ontology and the process of ontology engineering are used in a case study, which refers to the development of SEAMLESS-IF, an integrated modelling framework to assess agricultural and environmental policy options as to their contribution to sustainable development. The presented common ontology on assessment projects and scenarios can be re-used by IAM consortia and if required, adapted by using the process of ontology engineering as proposed in this chapter.

Sander Janssen, Frank Ewert, Hongtao Li, Ioannis. N. Athanasiadis, Jan-Erik J.F. Wien, Olivier. Thérond, Rob M.J. Knapen, Irina Bezlepkina, Johanna Alkan-Olsson, Andrea E. Rizzoli, Hatem Belhouchette, Mats Svensson and Martin K. van Ittersum, in press. Defining assessment projects and scenarios for policy support: use of ontology in Integrated Assessment and Modelling, Environmental Modelling & Software, in press.

7.1. Introduction

Integrated Assessment and Modelling (IAM) is increasingly used to assess the impacts of policies, technologies or societal trends on the environmental, economic and social sustainability of systems (Harris, 2002; Parker et al., 2002; Oxley and ApSimon, 2007; Hinkel, 2009). Prominent examples are the assessment of climate change impacts (Weyant et al., 1996; Cohen, 1997; Warren et al., 2008) and the assessment of quality and allocation effects in water resource management (Turner et al., 2001; Letcher et al., 2007; Ticehurst et al., 2007). Integrated assessment is defined by Rotmans and Van Asselt (1996) as an interdisciplinary and participatory process of combining, interpreting and communicating knowledge from diverse scientific disciplines to allow a better understanding of complex phenomena. IAM is then a methodology to combine several quantitative models representing different systems and scales into a framework for Integrated Assessment (Parker et al., 2002). Consequently IAM can cover several organisational and spatio-temporal scales to provide quantitative assessment of impacts on sustainable development.

Core features of any IA are the integration among disciplines and between scientists and stakeholders (Rotmans, 1998; Parker et al., 2002). Scenario analysis is an important technique in integrated assessment (Rotmans, 1998), where scenarios are developed and used in the interaction between scientists and stakeholders to anticipate and to explore possible futures and to assess potential consequences of different strategies into the future. The literature provides many different definitions of the concept scenario. 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 a scenario as ‘a coherent, internally consistent and plausible description of a possible future state of the world.’ In strategic business planning, where scenarios are often used as planning tool, scenarios are defined (according to Schoemaker, 1993) as ‘focused descriptions of fundamentally different futures presented in a coherent script or narrative.' Peterson et al. (2003) provides a definition of scenario which is closer to modelling, i.e. 'as variation in the assumptions used to create models.'

Next to a wide range of definitions for scenarios, also different classifications and typologies of scenarios exist (Rotmans, 1998; Greeuw et al., 2000; Alcamo, 2001; Van Notten et al., 2003; Borjeson et al., 2006): forecasting vs. backcasting scenarios, descriptive vs. normative scenarios, quantitative vs. qualitative scenarios, trend vs. peripheral scenarios, baseline vs. policy vs. business-as-usual scenarios and exploratory vs. anticipatory scenarios. A wide diversity of terms are associated with scenarios, such as indicators, driving forces, time horizon, time steps, storyline or narrative, processes, states, events, consequences and actions. It is not clear how these classifications and terms relate to each other and how they are used in constructing scenarios for IA.

Confusion and misunderstanding is particularly high when it comes to the implementation of scenarios. A researcher who is working in an IAM team, will be confronted with different types of stakeholders and scientists, with the latter covering a wide variety of disciplines and experiences. Each scientist will have a specific understanding of the concept scenario which is not consistent across disciplines and models. Discussions among scientists from different disciplinary domains and stakeholders are likely to result either i) in developing a 'container' term for scenario which serves as the magical solution whenever researchers are unclear about the way forward, or ii) in lengthy discussions on the meaning of scenario without arriving at any conclusion acceptable to the whole group. Again, the critical issue is that different models and policy problems have a specific implementation of scenarios targeted at that specific model or policy problem. There is a need for a clear set of rules and protocols with respect to scenarios in IA, as concluded by Rotmans and Van Asselt (1996), to avoid the danger of in-transparent, inconsistent, narrowly-defined and ad hoc setting of parameters (Rotmans, 1998; Van Asselt and Rotmans, 2002).

Chapter 7

This chapter considers a case study of achieving consensus on scenario definition in an an IAM consortium, System for Environmental and Agricultural Modelling; Linking European Science and Society (SEAMLESS) (Van Ittersum et al., 2008). It provides an computerized framework (SEAMLESS-IF) to assess the sustainability of agricultural systems in the European Union at multiple scales. The SEAMLESS consortium includes 30 institutions and more than 100 researchers from agronomy, economics, landscape ecology, social science, environmental science and computer science with dissimilar research background, leading to many different views on the meaning of scenario and its implications for the computerized integrated framework (SEAMLESS-IF). For example, biophysical simulation models (Van Ittersum and Donatelli, 2003) used for climate change impact assessment often apply the SRES scenarios framework (IPCC, 2000). In contrast, in a market model (Britz et al., 2007) a scenario typically refers to a policy that might be implemented in the future and that affects the market.

This chapter proposes a unified structured view for model-based scenario and assessment projects and a process of arriving at this result within a large community of researchers in a consortium. We demonstrate the use of common ontologies (see next Section for explanation) in building this shared conceptualization through a case study. This chapter describes our experiences in the challenging task of arriving at a shared conceptualization among researchers from different disciplines with dissimilar education and research experiences. We suggest that the process and the methods used are reusable for different integrated assessment tools or consortia developing such tools.

In the next Section, the theory behind common ontologies and the process of ontology engineering will be explained. Also, our case study based on the SEAMLESS consortium is introduced. In Section 7.3, the developed common concept on scenario and assessment projects is presented, including one fictitious example of the use of the common concept in an integrated assessment project at the regional scale. The common concept is discussed in Section 7.4. In the final Section we address our main findings as to the unified structured view on scenarios and assessment projects that we propose in this chapter. Throughout the chapter, we list some of the lessons we learned in our exercise to achieve this common understanding.

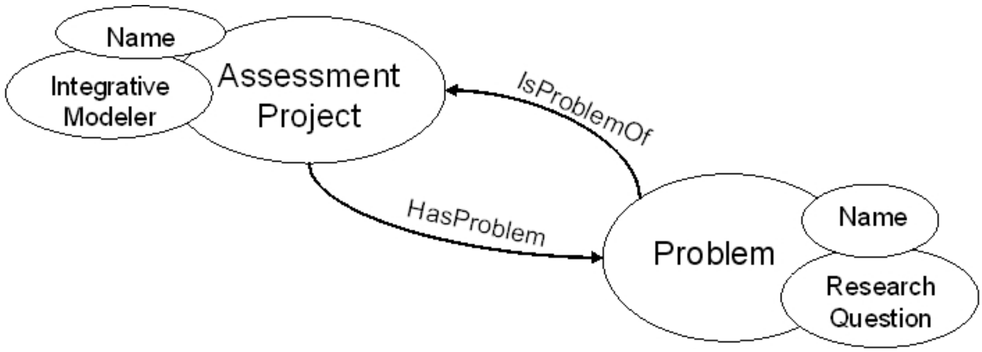


Figure 7.1. A part of an ontology showing two concepts (in ovals; Assessment Project and Problem), their relationships (uni-directional arrows; relationship as Assessment Project has Problem and relationship as Problem is Problem of Assessment Project) and their data-properties (Name for Concept Asessment Project and Problem, Integrative Modeller only for Concept Asessment Project and Research Question only for concept Problem).

7.2. Material and methods

7.2.1. Ontologies

In the context of integrated modelling, ontologies are useful to define the shared conceptualization of a problem. Ontologies consist of a finite list of concepts and the relationships among these concepts (Antoniou and Van Harmelen, 2004) (Fig. 7.1) and are written in a language, e.g. Web Ontology Language (McGuinness and Van Harmelen, 2004), that is understandable by computers. The term ontology originates from philosophy and was coined by classical philosophers Plato and Aristotle 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 that we wish to represent for some purpose' (Gruber, 1993). A computer can understand an ontology, because it is structured according to concepts and relationships on which it can reason, as opposed to unstructured files like documents or html (Antoniou and Van Harmelen, 2004) (Fig. 7.2). This difference is illustrated in Figures 7.1 and 7.2. Figure 7.1 can be understood by a human, while Figure 7.2 can be understood by computers. Applications of ontologies are known in the field of medical research (e.g. Musen, 1992; Flanagan et al., 2005) for lexicon or taxonomy-like descriptions of diseases or the genome, and computer science (e.g. Antoniou and Van Harmelen, 2004) for information and document management.

```
<owl:Class rdf:ID="Problem"/>
<owl:Class rdf:ID="AssessmentProject"/>
<owl:FunctionalProperty rdf:ID="ResearchQueston">
  <rdfs:domain rdf:resource="#Problem"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="IntegrativeModeler">
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
  <rdfs:domain rdf:resource="#AssessmentProject"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="isProblemOf">
  <owl:inverseOf>
    <owl:InverseFunctionalProperty rdf:ID="hasProblem"/>
  </owl:inverseOf>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty"/>
  <rdfs:domain rdf:resource="#Problem"/>
  <rdfs:range rdf:resource="#AssessmentProject"/>
</owl:FunctionalProperty>
```

Figure 7.2. A snippet of an OWL-file, describing the concepts Problem and AssessmentProject and relationships ResearchQuestion, IntegrativeModeller and isProblemOf from Figure 7.1. In an OWL-file, the ontology is stored in computer understandable format.

Scientists from various disciplines can define a common conceptual schema that their domains share as a basis for the integration of their models. A common assessment project ontology, i.e. an ontology which is shared by all domains considered for integration, serves as a knowledge-level specification of the joint conceptualization, in our case of the project and scenario definition. Each scientist can refer to and should adhere to the semantics of the concepts in the assessment project ontology, including restrictions on the concepts and relationships between the concepts.

7.2.2. Process of ontology engineering

The process of ontology engineering consists of set-up, design, approval and dissemination phases. In the set-up phase, the need for a common ontology is identified in the research consortium. In the design phase, agreement on the content of the common ontology is reached through a collaborative process. The common ontology is confirmed by the responsible researchers in the research consortium during the approval phase, while the communication of the common ontology to the whole research consortium occurs in the dissemination phase. In the remainder of this Section, we focus on the design phase, because this is the most complex and challenging phase in building the common ontology.

In the design phase, the following steps should be undertaken: (i) iterative discussion with relevant researchers to define the content of the common ontology; (ii) edit the common ontology in a dedicated ontology editor and (iii) use the common ontology for software development of model, database and graphical user interface. The first step in developing a common ontology, is that a group of scientists must agree and adopt one tight, well-reasoned and shared conceptualization. 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 has been developed (e.g. Protégé OWL; 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 proposed to be used. A collaborative approach is based on ‘development as a joint effort reflecting experiences and viewpoints of persons who intentionally cooperate to produce it' and it thus requires a consensus-building mechanism (Holsapple and Joshi, 2002). A collaborative approach has two advantages. First, researchers from different disciplines are diverse in their contributions, which reduces the chance of blind spots and which has more chances of getting a wide acceptance (Holsapple and Joshi, 2002). Second, it can incorporate approaches other than the collaborative approach (e.g. inductive, inspirational, deductive approaches) as required for development of parts of the ontology. For example, we built parts of the assessment project ontology through the inductive approach, e.g. by observing and examining cases from the literature on scenarios in integrated assessments.

The second step in the design phase is annotating the ontology in a computer understandable language by entering the ontology in a dedicated ontology editor (Knublauch, 2005). The third step is using the ontology for the development of databases, models and graphical user interfaces. The common ontology, which provides a conceptual layer independent of different programming paradigms, can be translated in source code for different programming paradigms (e.g. relational database, object-oriented programming). The Semantic-Rich Development Architecture (SeRiDA) (Athanasiadis et al., 2007b) can derive from this common ontology an object model and relational database schema. An object model is a schema of objects, properties and methods used in object-oriented programming. The SeRiDa facilitates the usage of appropriate tools for the tasks: (i) ontologies are used for storing semantics and supporting logical operations by reasoners, (ii) the object model is used for programming applications, graphical user interfaces, models and structuring the input to the models and (iii) the relational database schema is used for the persistent storage of data on assessment projects, scenarios, model inputs and results (Athanasiadis et al., 2007b).

7.2.3. Case study: policy assessment for sustainable development

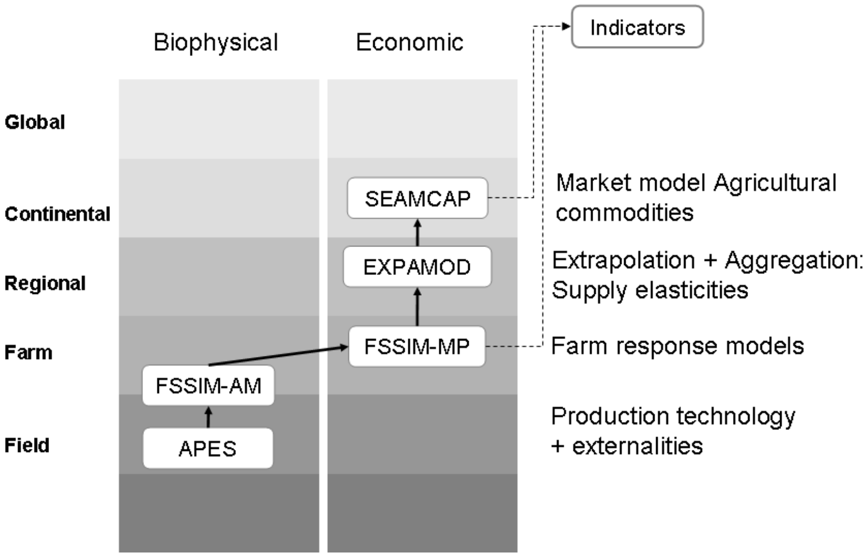


Figure 7.3. Backbone model chain of SEAMLESS-IF for field, farm and market level analysis, from the bottom to the top of the figure, respectively.

The SEAMLESS consortium develops a computerized and integrated framework (SEAMLESS-IF) to assess the impacts on environmental, social and economic sustainability of a wide range of policies and technological improvements across a number of scales (Van Ittersum et al., 2008). In SEAMLESS-IF different types of models and indicators are linked into model chains, where each model uses the outputs of another model as its inputs and ultimately indicators are calculated. With respect to the models (Fig. 7.3), macro-level economic partial or general equilibrium models (Britz et al., 2007) are linked to micro-level farm optimization models (Louhichi et al., 2009) and field crop growth models (Donatelli et al., 2009), using micro-macro upscaling methods (Pérez Domínguez et al., 2009). These models provide, through their outputs, the basis for the calculation of indicators of interest to the user. Each of these models are derived from different disciplines, operate on different time and spatial scales, are programmed in different programming languages and have a different implementation of scenarios.

Within SEAMLESS, modelling and stakeholder involvement are considered equally important in the assessment procedure proposed by SEAMLESS-IF. For applying SEAMLESS-IF, we foresee an integrative modeller working together with a policy expert (Van Ittersum et al., 2008). Accordingly SEAMLESS-IF must be designed to facilitate such a participatory approach (Ewert et al., 2009; Thérond et al., 2009). Potential users have a different understanding of scenarios than the modellers and they should not be confronted with the different implementations of scenarios in the models.

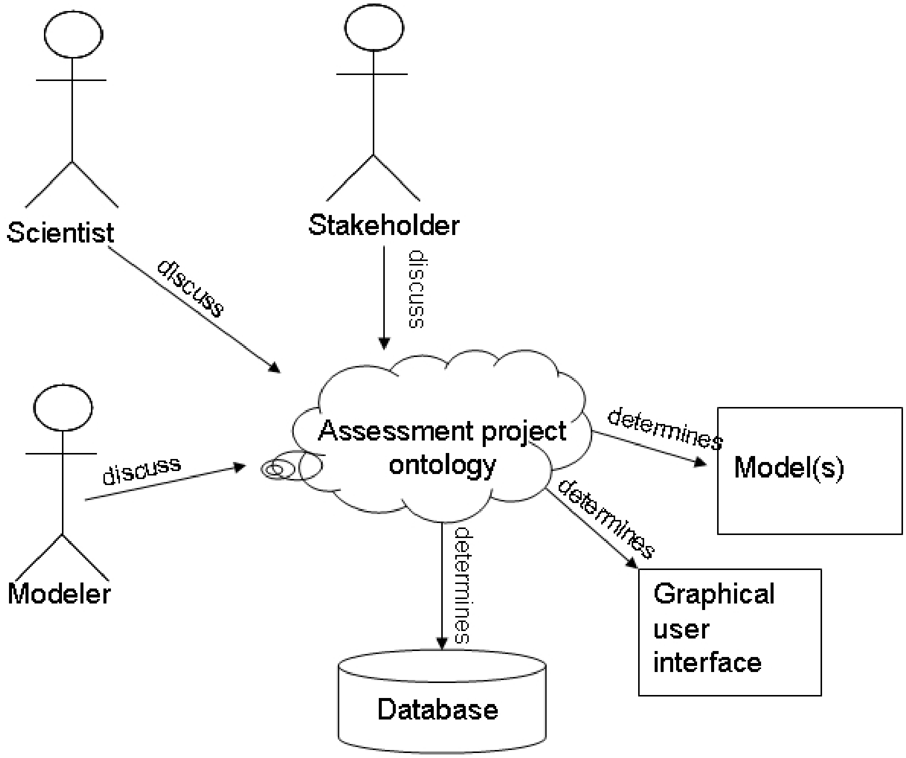


Figure 7.4. Role of an assessment project ontology in an integrated assessment modelling project

An assessment project ontology is thus required within SEAMLESS to unify the different implementations of scenarios in the different models across the different scales, indicators, programming languages and assessment problems. The assessment project ontology is a common ontology for definition of assessment projects and scenarios and it acts on the interfaces between modellers and other scientists and between scientist and users after the development of the SEAMLESS-IF (Fig. 7.4).

In our case study of the SEAMLESS consortium, one example of an application of the common assessment project ontology is presented. The example refers to an integrated assessment project for the region, Midi-Pyrénées in the South of France, concerning the impacts of the 2003 reform of the Common Agricultural Policy (CAP) (EC, 2003) as requested by two members of a regional government agency. The CAP2003 reform involves major changes in the subsidies that farmers receive for crops and animals (EC, 2003). The assessment must also incorporate the impact of CAP2003 reform on conservation agriculture in the Midi-Pyrénées region.

7.3. Results

7.3.1. Collaborative approach

The collaborative approach consisted of set-up, design, approval and dissemination phases. In the set-up phase, the need for a project ontology was identified by scientists responsible for integration in the research consortium. The method to make the project ontology was proposed and agreed, after which the design phase started. The method is to develop one shared document in Microsoft Word on the meaning of scenario and assessment projects between a group of seventeen researchers from different disciplines working in different parts of the SEAMLESS consortium.

In the design phase, ten iterations of the document were used and after each iteration an ontology constructed in Protégé OWL (Knublauch, 2005) was adjusted to the outcomes of each iteration. Two knowledge engineers acted as impartial facilitators, who pro-actively identified and discussed open issues to find agreement, without imposing own opinions about the content of the common ontology. They also edited the common ontology in an ontology editor. With each iteration, more scientists were involved starting from four for this first iteration up to seventeen for the tenth iteration (Fig. 7.5). Most scientists offered voluntarily to contribute to the document, as they realised the need for the document and were committed to the research consortium. Three scientists were included through invitations to contribute to the document, because of their crucial role in the research consortium and a balanced representation of the different research domains and roles in the consortium.

At the start of the document a clear and precise description of the aim and requested actions of the participants were provided, which was needed to avoid confusion. Due to the choice for a document, the descriptions of concepts and relationships in the document had to be such that the descriptions are not open to multiple interpretations. Formulations like 'concept has one and only instance of another concept' and 'concept has one or more instances of other concepts' were used for relationships and 'concept is …' or 'concept is defined as…' for definitions. In case of conflicts on the meaning of concepts or relationships, the two impartial knowledge engineers could mediate to build consensus. The consensus building usually occurred through asking questions to the domain scientists to further explain their ideas on the meaning of concepts and relationships. By asking questions new insights were obtained and the project ontology developed into a more advanced state. In some cases, meetings were organised, in which the domain scientists discussed unclear parts of the project ontology. During these discussions, the knowledge engineers made proposals on possible ontology structures until an ontology was accepted by all present.

In the approval phase after the tenth iteration, both the document and the ontology were 'closed' after the approval by the core group of researchers. At the tenth iteration, a set of actions was formulated to elaborate specific parts of the project and scenario definition. An example of an action was to investigate the relationship between scale and scenarios. Also, a set of four fictitious sample assessment projects was formulated during the iterations as a testing exercise of the ontology developed so far. One of these examples is presented below (Section 7.3.2.7.).

In the subsequent dissemination phase, a group of seventeen scientists with high commitment to the assessment project ontology were available that consequently helped to further explain and establish the ontology with the scientists in the consortium. Interestingly, the scientists not involved in the process did not indicate any need to re-discuss the project ontology. These scientists were mainly interested in how their own research fitted to the developed ontology. Eventually, the ontology has been evaluated and accepted within the consortium. The wider evaluation of the common ontology is facilitated by making it open to scientists outside the consortium (see Section 7.4.2).

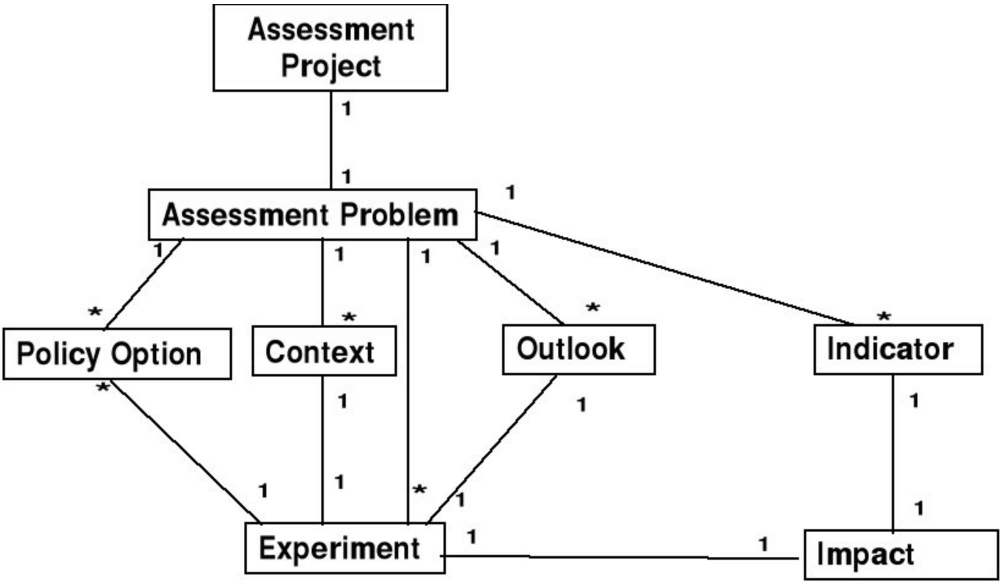


Figure 7.5. A simplified data model of the project ontology with annotations between the concepts, indicating whether it is a 'One-to-One' relationship (1__1; One project is only related to one assessment problem and vice versa) or a 'One-to-Many' relationship (1__*; one Experiment can have one or more policy options).

These four phases of set-up, design, approval and dissemination required about one and half year. The set-up and approval phase were both relatively short, e.g. a month. The design phase required about six months, with the two knowledge engineers working for 50% of their time on the assessment project ontology and domain scientists spending about one day at each iteration. The total time investment in the design phase is estimated at one and half man-years for the seventeen scientists involved. The dissemination through presentations or meetings to the rest of the consortium took about a year till all researchers were accustomed with the assessment project ontology. The set-up phase was initiated at the end of the second year on a total of four years of the research consortium. Advantages of initiating it at that time in the research consortium were that scientists were familiar with each other and each other's work and that a group of committed scientists interested in such an exercise could easily be identified.

7.3.2. Assessment project ontology

The content of the assessment project ontology is further verbally described based on the document developed. An assessment project in SEAMLESS refers to the assessment of changes in policies or technological innovations on the sustainability of agricultural systems. An assessment project consists of one or several experiments that capture a specific perspective on the assessment problem. A project has one and only one assessment problem. One problem has the following properties: (i) one spatial and temporal scale, (ii) one or more contexts, (iii) one or more policy options, (iv) one or more outlooks, (v) one or more experiments and finally, (vi) one or more indicators (Fig. 7.5 and 7.6).

Table 7.1. Feasible scales of the assessment problem and models that can address a problem at that specific scale.

Extent	Resolution	Models
Continental	Agri-environmental zone	APES
Continental	Farm type	CAPRI- FSSIM-AM/MP
Continental	Region	CAPRI
Region	Agri-environmental zone	FSSIM – APES
Farm type	Agri-environmental zone	FSSIM – APES

7.3.2.1. Scale

Scale refers to the physical dimensions (most commonly space and time) of observed entities and phenomena (meaning that dimensions and units of measurement can be assigned). Each scale has two relevant attributes: the extent and the resolution. The extent defines the boundaries, the area or the magnitudes, for example from year for a temporal scale or continent for a spatial scale. Resolution refers to the finest detail that is distinguishable, for example a day for a temporal scale or member state for a spatial scale. Based on the models available in an integrated assessment project, a limited set of assessment scales is feasible. An example of possible assessment scales based on the SEAMLESS project is given in Table 7.1.

7.3.2.2. Context

Each experiment within a problem will be based on one context that can be different from those of other experiment(s). The context describes the delineation of the object of interest. The delineation determines what is inside and what is outside to the system modeled and define the range of options or possibilities within which changes due to policy options and outlooks can occur. The properties of the context describe the input parameters of the simulation and combinatorial models. These models require assumptions to define and simulate options or possibilities used by other models to assess the consequences of a policy change or innovation. The context must contain assumptions on what is technologically possible in the future, for example will genetically modified cultivars become available at a large scale? Also, the context makes the abstract temporal and spatial scales (Section 7.3.2.1) concrete by specifying the temporal and spatial delineation. For example, for an assessment problem on the continental scale, the context specifies that the member states of the European Union in 2008 are of interest.

7.3.2.3. Policy Option

Each experiment within a project assesses the effects of one or a combination of several policy options. One policy option refers to one or more policy measures as part of it. Each policy option has a set of policy parameters within a given timeframe or for a given time series, that are not modified by any of the models in the assessment while running. An example of a policy option is the introduction of decoupled payments in the EU as part of the Common Agricultural Policy (EC, 2000). This policy option consists of two policy measures, which are the introduction of direct income-support and cut of area- or head-based premiums. These policy measures are quantified by the reference yield for a region to calculate the income support level and the premium levels, which are cut.

7.3.2.4. Outlook on the future

An assessment problem can have one or more Outlooks on the future. 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, but which are not modeled endogenously. Examples of outlook parameters of relevance to SEAMLESS are atmospheric CO2-concentration, shifts in demands for agricultural products and energy prices. Outlooks are usually highly contestable images of what might happen in the future, and therefore it is recommendable to assess a problem under contrasting alternative outlooks, e.g. an economically-oriented versus an environmentally-oriented outlook, a globalization versus a regionalization outlook, a high-economic growth versus a low-economic growth outlook. Sometimes these outlooks are based on discussions between a large group of researchers and stakeholders, for instance the IPCC Special Report on Emission Scenarios (SRES) (IPCC, 2000).

7.3.2.5. Experiments

One assessment problem has at least two or more experiments. One experiment represents the assessment of one or a combination of several policy options in a given context and outlook on the future, which translates into one run of the models within SEAMLESS-IF and calculates values for a set of indicators. One experiment describes the reference situation, i.e. the baseline experiment (Alcamo, 2001). This baseline experiment consists of a policy option describing the policy instruments that are already phased in, or have been agreed upon, an outlook describing the projection of current trends and a context describing the current situation. The definition of one or more experiments assures that a with/without or before/after analysis of changes can be made. The experiments define the changes as compared to the baseline experiment, 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). The maximum number of experiments is the full factorial combination of contexts, outlooks and policy options, although some combinations of contexts, outlooks and policy options may not be sensible and useful to assess.

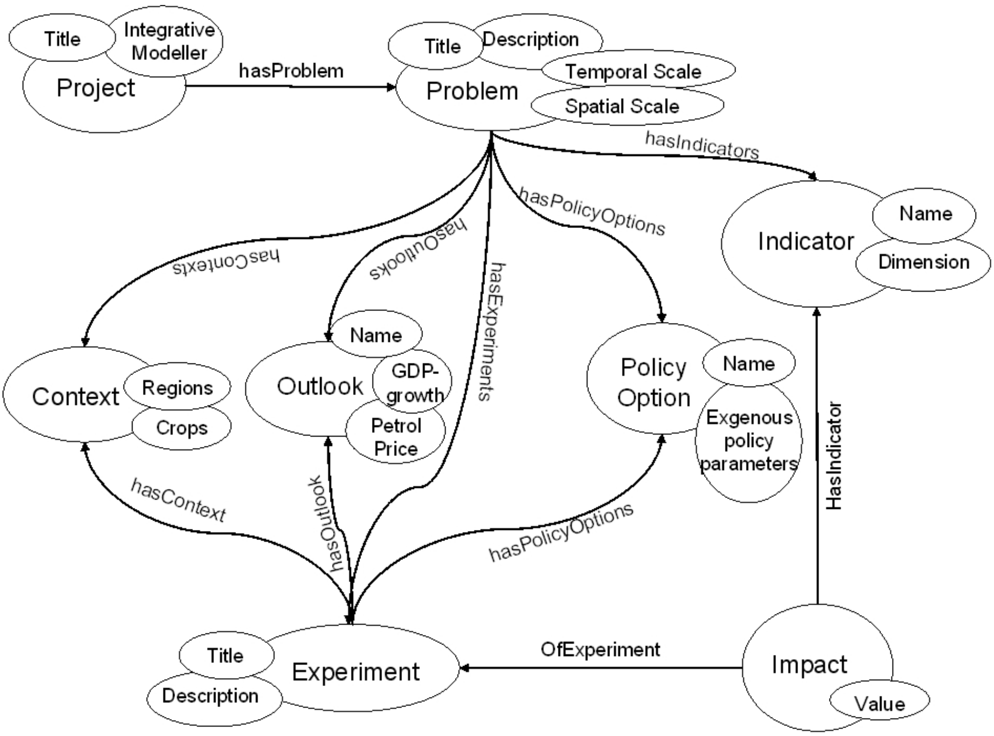


Figure 7.6. Schematic overview of the assessment project ontology.

7.3.2.6. Indicators

Each assessment problem is associated with a set of indicators that are of interest for the policy expert. Indicators synthesize relevant data and model outputs and indicate the change or define the status of something (Gallopin, 1997). A value for each indicator is calculated with a model run for an experiment. The indicators must be the same among experiments in one assessment problem allowing comparison of indicator values among experiments. Impacts are the changes in indicator-value for one experiment due to changes in policy options, context and outlook as compared to the baseline experiment.

Table 7.2. Experimental set up for an assessment problem on impacts of CAP2003 reform and conservation agriculture in Midi-Pyrénées region in France

Experiments	Policy option	Outlook	Context
1. Baseline	Only current policies apart from CAP 2003 reform	Business as Usual	No conservation agriculture
2. CAP 2003 reform	CAP 2003 Reform	Economically oriented	No conservation agriculture
3. No support	CAP 2003 Reform	Environmentally oriented	Conservation agriculture
4. Conservation oriented in regional world	CAP 2003 Reform and subsidies for conservation agriculture	Environmentally oriented	Conservation agriculture
5 Conservation oriented in a global world	CAP 2003 Reform and subsidies for conservation agriculture	Economically oriented	Conservation agriculture

7.3.2.7. Example of regional assessment project in Midi-Pyrénées

The example introduced in Section 7.2.3 refers to an integrated assessment project for Midi-Pyrénées of the 2003 reform of the Common Agricultural Policy (CAP) and impacts on conservation agriculture. The spatial scale for this example has the extent of a region and the resolution of a farm type, as the example focuses on one region and on the impacts on specific groups of farms. The temporal scale has an extent of the period from 2003 to 2013 with a resolution of a year. The year 2003 is used for calibrating the models. The experimental set up of the assessment problem with descriptions of experiments, outlooks, policy options and context can be found in Table 7.2. Relevant indicators for this assessment problem are the regional cropping pattern, the farmers' income, the amounts of subsidies, the % of no-ploughing tillage, the area for the intercrops mustard and clover and the level of erosion.

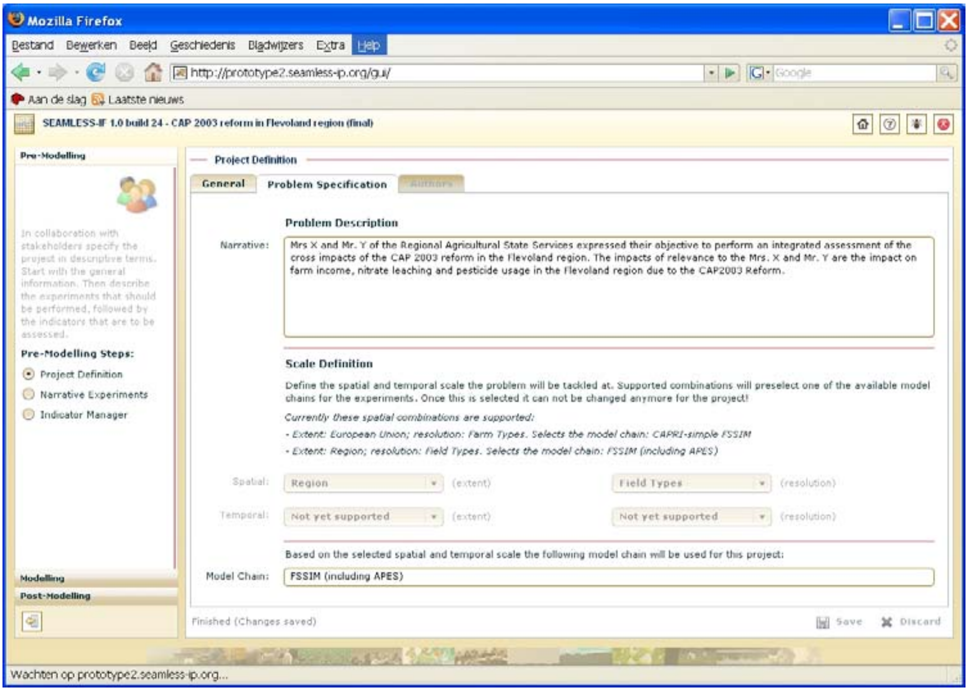


Figure 7.7. Screenshot of the GUI displaying an assessment problem.

7.3.3. Ontology use for software development

The assessment project ontology is shown in diagrams, i.e. one datamodel (Fig. 7.5) and one ontology-schema (Fig. 7.6). The datamodel can only be translated into a database schema, while the ontology-schema can be translated both into a database schema and a set of classes for object-oriented programming through SeRiDA (Section 7.2.2). The assessment project ontology was used to generate a set of tables to store the project information in a relational database and a set of JavaBeans for communication between graphical user interface, models and database (Fig. 7.5 and 7.6). The JavaBeans are used to deliver parameters described in the ontology as inputs to the models. The assessment project ontology has impacted the design and set up of the Graphical User Interface (GUI) of SEAMLESS-IF, as can be seen in Fig. 7.7. In Fig. 7.7 the part of the GUI is shown, where the problem is defined, by providing a description and selecting the temporal and spatial scale of the assessment problem. Through the specification of the scales, the model chain of relevance is selected by the GUI and displayed. The GUI through the assessment project ontology enforces the explicit definition of the link between an assessment problem, a model chain and a spatial scale. Thereby the assumptions required to link an assessment problem, a model chain and a spatial scale become transparent.

7.4. Discussion

7.4.1. Scenario and its meaning

In our assessment project ontology as presented in Section 7.3, we have no explicit concept scenario. In the iterative process of building the common ontology, we experienced that scenario had different meanings for different scientists. During the process, some scientists thought of scenarios as experiments, so a perspective of future changes in parameters of policy options, outlooks and context, and thereby determining the input parameters for the models. Other scientists thought of scenarios as a set of impacts in the sense of indicator values that change depending on 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. In the approval phase, the multiple meanings of scenario were demonstrated to all participants involved in the collaborative approach. The core group of scientists approving the proposed project ontology decided on a suitable definition of the word scenario for the research consortium, i.e. a scenario represents the changes or driving forces in policy options, outlooks and contexts in an experiment compared to the baseline experiment (Thérond et al., 2009). Through the collaborative approach the multiple meanings of scenarios became managed and explicit decisions were taken, which increased transparency and clarity for scientist participating in the research consortium.

The concept scenario is further detailed through the assessment project ontology to cover a range of models and disciplinary understanding of what a scenario is. In the proposed assessment project ontology other concepts instead of scenario were chosen that could be defined unambiguously without multiple historical connotations, and agreed upon to avoid risk of confusion. Through the flexibility offered by concepts like context, policy option, outlook, experiment and assessment problem, the project ontology is able to cover all the different meanings which the concept scenario can have, and offers an opportunity to comprehensively describe an integrated assessment problem. Scenario definition as held by other stakeholders outside the science-community (e.g. policy makers) is not included yet in our assessment project ontology.

The different definitions and classifications of scenarios from literature as described in Section 7.1 were not readily usable as content in the assessment project ontology. We consider the assessment project ontology as a definition of scenario for multi-disciplinary and multi-scale research consortia in Integrated Assessment. Subsequent research should investigate, if it can become a standard for definition of scenarios and assessment projects across research consortia. The assessment project ontology in Section 7.3.2 presents a first simple formulation, that can be extended and detailed in further research. The simple formulation in Section 7.3.2 indicates that advanced and complex definitions and classifications from the literature are obsolete and not targeted.

7.4.2. Project ontology and models

The selection and configuration of models is not explicitly mentioned in the project ontology and the fictitious sample project as presented in Section 7.3.2, although a link exists between the properties of the context, outlook and policy option and input parameters for the models. As mentioned by Parker et al. (2002), scale is recognised as an important concept in integrated assessments and in our project ontology it is as a central node that determines (i) the models/model chains that should be run, (ii) the parameters or properties that should get a value with respect to outlook, context and policy option (Table 7.3), (iii) the indicators which can be selected and (iv) the results to be presented. In an integrated assessment we must make a distinction between the scales of the assessment problem and the scale(s) of the models. The scale of the assessment problem refer to the research question, properties of policy option, properties of context, properties of outlook and indicators. The scale(s) of a model is defined by the modeler and refers to the scale(s) at which relationships are modeled and outputs are simulated that are used at the scale of assessment.

Table 7.3. The relevance of properties of policy option, context and outlook for different types of models.

Models	Policy option	Outlook	Context
Crop growth simulation model	--	++	+++
Farm model	+++	++	+++
Market model	+++	++	+
General Equilibrium model	+++	++	+++

-- = no properties for this model

+ = limited number of properties

++ = average number of properties

+++ = many number of properties

Each assessment problem is linked to one spatial and one temporal scale, although this does not mean that multi-scale assessments are not possible. A multi-scale sustainability theme such as climate change or CAP2003 reform has to be subdivided in several assessment problems, each on their own scale with relevant assessment question, indicators, model chain and properties of outlook, policy option and context. For example, in assessing the impact of climate change on agriculture, one feasible assessment problem is to study the impact of climate change on farmer income and environmental farm performance in a region, while another feasible assessment problem is the impact of climate change on farm production and trade in agricultural commodities in the European Union. Both assessment problems require different models at several spatial and temporal scales (Table 7.1), leading to two multi-scale assessments in terms of models and indicator values. Indicator values can be calculated at the scale of the assessment problem and finer scales, at which indicators can reasonably be calculated from available model outputs.

The properties of context, policy option and outlook are the input parameters to the models. One property can be an input parameter to more than one model. For example, a quota policy is defined by a value of the quota and a product to which the quota is applied for each farm type. This quota policy can be used both by a market-scale model and a farm-scale model. By specifying properties of policy option, context and outlook a library of possible model input parameters is created that can be used by different models. Hereby we decouple the description of an assessment project through relevant parameters from the use and implementation of these parameters by the models. This decoupling shifts the focus from the technical capabilities of the models to the assumptions made while defining values for the different model input parameters and defining the experiments (Rotmans, 1998; Greeuw et al., 2000). The use of experiments in defining projects also helps to make assumptions explicit, because these experiments capture the changes between a baseline experiment and the other experiments. By considering explicitly the differences between experiments, the changes in indicator values can be analysed. If many differences between two experiments occur, then it is more difficult to interpret the changes in indicator values. Designing sensible and useful experiments is therefore a challenging task.

By decoupling our understanding of scenarios and projects as captured by the assessment project ontology from the model input parameters, the assessment project ontology can be reused for other integrated assessment modelling research that deals with policy assessment and sustainable development and thus is a separate part of knowledge produced by a group of scientist as foreseen in the vision of the semantic web (Berners Lee et al., 2006). The project ontology is available on http://delivered.seamless-ip.org:8060/browser/zul/main.zhtml.

7.4.3. Use of ontologies and ontology engineering

To build the assessment project ontology a collaborative approach was used that involved scientists with different disciplinary backgrounds. 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. As an ontology can only support concepts, relationships between concepts and restrictions on relationships or concepts, scientist could only discuss in these terms. In other words, three conditions have to be met for a concept to be included in a common ontology: (i) the concept has to be clearly defined; (ii) the concept has to be consistent and coherent with other concepts in the ontology, (iii) one or more scientists have to provide the 'burden of proof' to fulfil the previous conditions.

With ten iterations and seventeen participating scientists, the collaborative approach required a clear objective, two persons managing the process (by setting deadlines, determining the type of contributions and the required participants) and a set of actions for each iteration, which made it a time-consuming task. Up to five participants sent contributions and feedback to each iteration of the document, which then had to be evaluated on their merits and which had to be discussed in case of diverging opinions. Critical success factors in the collaborative approach were the commitment of participants to the process and the presence of one or more knowledge engineers.

Many suitable tools to edit ontologies (see Knublauch (2005) and GO-Consortium (2007)) exist and we used these to edit the project ontology once consensus was reached. In the collaborative approach to reach consensus, we used Microsoft Word-documents. Documents had two advantages compared to dedicated ontology editors. First, all participants in the collaborative process have Microsoft Word installed on their computer and are used to communicate with documents. Second, the agreed ontology in the ontology editor was shielded from participants, as it is only necessary that a knowledge engineer edits the ontology in a dedicated ontology editor. Through track-changes and comments in the document, multiple participants were able to simultaneously edit the common ontology and their individual contributions could be followed and synthesized to a joint understanding of the problem at hand. We did not invest in the development of a tool for collaborative ontology editing, as initially we did not know the requirements for such a tool and the way the participants would work in this process. Through our experience of building the project ontology in a shared document, we learned that a website for ontology editing as proposed by Farquhar et al. (1995) could be helpful. However, such a website for ontology editing, in which all participants can edit the ontology, is only useful if it registers the users and their activities, if it allows a knowledge engineer to finalise parts of the ontology and make them non-editable, if it has a very simple and intuitive user interface to propose concepts and the relationships to other concepts and if it forces users to use specific formulations to define concepts and their relationships. Wiki-technology could provide a useful starting point for the development of such a website.

7.5. Conclusions

Although literature provides many advanced and complex definitions and classifications of scenarios, these definitions and classifications cannot be made operational for research consortia in IAM. Our common ontology on assessment projects and scenarios provides an operational and simple definition of scenarios and assessment projects. It improves the consistency, transparency and applicability range across disciplines of scenarios, as (i) a set of concepts is provided to describe different types of model input parameters, (ii) the focus is on assumptions made in defining these input parameters instead of on the models, GUI's or databases themselves and (iii) experiments are explicitly constructed capturing the different perspectives and assumptions on the future. The assessment project ontology can be reused by other Integrated Assessment and Modelling consortia that deal with policy assessment and sustainable development and could become a standard for the definition of scenarios and assessment projects in the future.

We recommend for any Integrated Assessment consortium to clarify with its participants the meaning of scenario, associated concepts or other concepts with vague and ambiguous meaning (e.g. driving forces, indicators). We achieved such a clarification 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. The common ontology can be directly used for development of databases, models and graphical user interfaces. A collaborative approach for clarifying concepts in a multi-scale multi-disciplinary research consortium was developed, while building our common ontology. This collaborative approach can be re-used to extend the assessment project ontology or to build a shared understanding in other IAM research consortia.

Chapter 8. Discussion, conclusions and broader perspective

8.1. Introduction and Reading Guide

Different aspects of methodological, semantic and technical integration were described in Chapters 2 to 7 organized along two case studies. i.e. bio-economic integration and multi-issue integration. Bio-economic integration considered an integration covering two scales (e.g. field to farm) within a bio-economic farm model involving a relatively small team of ca. ten agronomic and farm economic scientists (Chapters 2, 3, 4). Multi-issue integration described the integration in modelling European agricultural systems, which covers five scales (e.g. field, farm, region, country and continental) and links different types of models for the calculation of impacts through indicators. A large team of about hundred scientists from agronomy, economics, landscape ecology, information technology and environmental sciences (Chapters 5, 6, 7) was involved.

Table 8.1 gives an overview of the relative importance of methodological, semantic and technical integration in each of the chapters. In the case study of bio-economic integration, this thesis mainly considered methodological integration relevant to bio-economic farm models (Table 8.1). In this case study, semantic and technical integration are described to show how these facilitated the methodological integration. The case study of multi-issue integration described methodological, semantic and technical integration in detail, with some emphasis on semantic integration. This chapter starts with a discussion of integration in the two case studies. This discussion devotes a section to methodological, semantic and technical integration. Subsequently, this chapter presents main conclusions from the previous chapters and the discussion. Finally, the link between social and institutional barriers to integration and methodological, semantic and technical integration is investigated based on experiences from the SEAMLESS project.

Table 8.1. Importance of methodological, semantic and technical integration in the different Chapters (+++ = very important; ++ = important; + = present; -- = not present)

Case study	Chapter	Methodological Integration	Semantic integration	Technical integration
Bio-economic integration	Assessing farm innovations and responses to policies: A review of bio-economic farm models (Chapter 2)	+++	--	--
	A flexible and comprehensive conceptualization of alternative agricultural activities for bio economic farm models (Chapter 3)	+++	--	+
	A generic bio economic farm model for environmental and economic assessment of European agricultural systems (Chapter 4)	+++	++	++
Multi-issue integration	A database for integrated assessment of European agricultural systems (Chapter 5)	+	+++	++
	Linking Models for Assessing Agricultural Land Use Change (Chapter 6)	+++	+++	+++
	Defining assessment projects and scenarios for policy support: use of ontology in Integrated Assessment and Modelling. (Chapter 7)	++	+++	+

8.2. Discussion

8.2.1. Methodological integration

8.2.1.1. Bio-economic farm models

To date many different (applications of) BEFMs have been published, which cover a wealth of different problems, data sources and modelling techniques (Chapter 2). The comparison of these BEFMs and modelling techniques is difficult due to the lack of provision of model source code with publications, the limited descriptions of data sources, the lack of model evaluation and a strong emphasis in their description on an application of a model to some policy change or farm innovation. Yet, methodological challenges for bio-economic farm models have been identified (Chapter 2). First, the type (e.g. normative vs. positive; policy assessment, technological innovation or both; empirical vs. deterministic) and purpose of the model must be explicitly mentioned in any modelling study. Second, model evaluation through sensitivity analysis, correspondence of model results to reality or validation must be explicitly and comprehensively addressed. Third, sources and descriptions of agricultural activities used as model inputs must be explicitly considered and presented. Fourth, strategic decision making and social environment have to be more explicitly incorporated in bio-economic farm models. Finally, an easily transferable BEFM with a generic and modular structure is to be developed, which would enable a group of researchers to jointly work on one BEFM and benefiting from the synergies in model development.

These shortcomings of model description, evaluation and application-bias lead to a lack of credibility of results and transparency of the methods and data. This lack of credibility and transparency indicates that BEFM as a scientific method has not yet matured to be used as an ex-ante assessment tool. Transparency can be increased through standardization, which makes review of models and model results by scientists and other stakeholders (e.g. policy makers, farmers and environmental organizations) easier and more thorough. Standardization facilitates reproducibility and re-use of models and model applications, but is not a prescription of the type of model to make or the scientific content of the model. Standardization means setting up generally accepted, implemented and adopted procedures to develop, document, program, use and evaluate a model. Examples of standardization are a standard source-code implementation of an objective function based on profit maximization and a procedure to report a BEFM in a scientific publication or report. A procedure for publishing BEFMs in scientific journals provides guidelines for the topics in the article (e.g. appendix with model equations, section on model evaluation) and the delivery method of the model and data (e.g. downloadable from a public website). Such procedures can become a type of standard, e.g. the ISO norms (ISO, 2009) or a quality label as used in food production (cf. organic or Marine Stewardship Council (MSC, 2009)). Available standards and guidelines can be adapted and furthers specified for BEFMs. For example, Jakeman, et al. (2006) provide ten steps in development of a model for natural resource management, that can be adapted to BEFMs. Advantages of standardization are, that i. the reviewer is more likely to be familiar with modelling solutions and advances proposed in a particular BEFM; ii. modellers cannot easily hide ‘quick fixes’ and are forced to provide a comprehensive insight into their BEFM and iii. novices to BEFMs can more easily learn to develop a BEFM, as documentation and expertise are easily accessible and available.

Standardization is targeted by explicitly considering the type of BEFM, which would make it easier to classify and compare different BEFMs. The classification presented in Chapter 2 can be used. The description and explicit consideration of alternative agricultural activities described in Chapter 3 also contributes to this standardization, as a procedure is proposed to help avoid blindspots in the definition of alternative agricultural activities.

A robust and standardized procedure for model evaluation is crucial to build credibility with researchers and stakeholders. The development of a standard procedure has to be combined with an assessment of the strengths and weaknesses of existing methods in back-casting experiments, see for example Kanellopoulos et al. (2009). Such an assessment provides the scientific grounding of proposed modelling techniques.

Standardization is targeted by the development of an easily transferable BEFM with a generic and modular structure. A first step in this direction has been made with the development of FSSIM (Chapter 4) by a group of agro-economic researchers. Still, the development and use of FSSIM only represents a first step, as extensions (e.g. perennials, multi-functionality, different types of decision making and alternative livestock activities) and a critical evaluation of model and model results through extended peer review (Funtowicz and Ravetz, 1990) with stakeholders and users are required. In the future, it might be useful to move away from FSSIM as a model, to FSSIM as a library of adaptable model functions relevant to BEFMs. For example, if many different objective functions (e.g. normative profit maximization, positive profit maximization using Positive Mathematical Programming, expected utility, long term profit maximization) are available in FSSIM, then a modeller can compile his "FSSIM" by selecting the objective function that suits his purpose.

8.2.1.2 Integrated assessment and modelling

Linking models into model chains is a common practice in IAM and it requires harmonization between different modelling techniques (e.g. simulation, optimization and estimation), different time (e.g. day, decades and years) and spatial (e.g. region, ecosystem, landscape and continent) scales, and different economic, social and biophysical processes descriptions (e.g. farmer decision making, plant water uptake, market equilibrium). Time in these process descriptions can be incorporated in a static or dynamic way, while space is either referred to as a point or a spatial entity (e.g. a farm, a region). Although there are some methods to methodologically align modelling techniques in one discipline (e.g. systems dynamics, state-rate analysis, mathematical formalism (Hinkel, 2008)), there is a lack of a generic method across disciplines as was concluded in Chapter 5. The lack of a generic method implies that methodological integration across models remains an ad hoc activity. An advantage of methodological integration as an ad hoc activity is that it allows for flexible solutions depending on the models and purpose of the model linking. Disadvantages are that it makes it difficult to compare different chains of linked models and to provide a transparent and general acceptable format of describing and reviewing a linked model chain. As a first step to develop a generic method, it might be insightful to compare and contrast a number of existing model chains (e.g. the SEAMLESS model chain in Chapter 6, IMAGE model (NMP, 2006) and SENSOR model chain (Jansson et al., 2008)) with respect to their specific and general aspects in methodological integration. More steps are needed to develop and establish such a generic method for methodological integration.

Crucial concepts to link the models in SEAMLESS were agricultural activity between FSSIM and APES and price elasticity between FSSIM, EXPAMOD and CAPRI. Agricultural activity describes the inputs and outputs involved in agricultural production on a field, and FSSIM and APES both set different parameters of this agricultural activity. Price elasticity describes the percentage change in supply with one percent change in price, and EXPAMOD aggregates and extrapolates these out of FSSIM supply responses on farm level for CAPRI. Probably in linking other models that are of the same type (e.g. cropping system model, bio-economic farm model, econometric estimation model and partial equilibrium market model) and on the same scale (e.g. field, farm, region and continental) similar concepts and links can be established.

In this thesis a strict separation was maintained between the models, data and scenarios for linked model chains. For data sources (Chapter 5) and scenarios (Chapter 7) an integration was achieved, which was used by the models (Chapter 6) and which can be used independently of the models. Advantages of such a strict separation are that non-modelling experts can work on the integration of the data sources and the definition of scenarios, that other model chains or integrated assessment projects can also make use of the integrated database or scenario definition and that maintenance of the independent models, data and scenarios is easier. Disadvantages are that the models might need some adaptation to such an integrated database or scenario definition and that stronger coordination is needed between data, model and scenario specialists to arrive at shared conceptualization of time and spatial scales.

8.2.2. Semantic integration

8.2.2.1 Bio-economic farm models

Semantic integration in or for BEFMs is described in almost all chapters of this thesis with the exception of Chapters 2 and 3. The ontology created during the semantic integration helps to define and develop components of a BEFM (Chapter 3), to link components within the BEFM (Chapter 4), to integrate the data-sources for a BEFM (Chapter 5), to link a BEFM to other models (Chapter 6) and to define scenarios for a BEFM (Chapter 7). The ontology has been set up as a granular ontology, meaning that it exists of different sub-ontologies which are hierarchically linked (Athanasiadis et al., 2009). This implies that some of the sub-ontologies are specific for BEFMs. Examples of these are a sub-ontology on farm optimization as done in FSSIM-MP (Chapter 4) and a sub-ontology on definition of rotations as conceptually defined in the Production Enterprise Generator (Chapter 3). These ontologies capture part of the knowledge of the model by describing the data structures, e.g. activities, representative farms, quotas, prices. Although other BEFMs may use different methods to work with the data structures, the data structures may be similar independent of their original models. The ontologies offer standardized data structures for bio-economic farm models that can be formally extended and adapted. By adhering to the data structures in the ontologies, different bio-economic farm models refer to the same set of meta-data and definitions. Also, bio-economic farm models can easily switch between data-sources, if these are compatible with the ontology. The current ontologies present a first prototype, that require an extensive review and use before these can be adopted as formal standard for bio-economic farm models.

8.2.2.2 Integrated assessment and modelling

A shared definition of terms and concepts through a semantic integration is required in IAM to ensure scientific rigor, transparency and trust in databases (Chapter 5), models (Chapter 6) and scenarios (Chapter 7). This thesis achieved this shared definition of terms and concepts by developing an ontology in an iterative collaborative communication process. Descriptions of the technical usage and deployment of ontologies are available (Musen, 1992; Farquhar et al., 1995; Rizzoli et al., 2008; Scholten, 2008) and this thesis supplements this with a collaborative approach to build the ontology.

For the development of each sub-ontology of the overall ontology, the collaborative approach was adapted to the domain members involved. For the definition of the concept “scenario” (Chapter 7), a large group of researchers with dissimilar research backgrounds had to be involved, so the collaborative approach was based on jointly editing a Word document in fast iterations. More scientists were involved at each iteration. For the development of the integrated database (Chapter 5), the collaborative approach was arranged in long iterations resulting in prototypes of the database. The collaborative approach was carried out by one domain scientist and two knowledge engineers. The domain scientist consulted a small group of domain scientists when required. Finally, for the ontologies of the models (Chapter 3, 4 and 6), adaptations were made to the collaborative approach according to the models involved. For the FSSIM models (Chapter 3 and 4), the ontology was largely built through an inductive approach of examining model inputs, outputs and equations. This inductive approach was supplemented with extensive discussions of the resulting ontologies between one domain scientist, one computer scientist and one knowledge engineer. Finally, for the model chain (Chapter 6) workshops were organised with relevant domain members to develop parts of the ontology.

Adapting the collaborative approach to the participants involved and pro-actively engaging domain scientists was helpful to realise the ontologies. Implementing the ontology in an ontology language (e.g. OWL; McGuinness and Van Harmelen, 2004) is not a difficult or time consuming task, but agreeing on the content of the ontology is time-consuming and contentious. Automation of part of the collaborative approach to build ontologies is desirable to more efficiently build ontologies in the future. Such automation can be based on description logics (Villa et al., 2009; Buccella et al., 2009), which are the mathematical basis of ontologies. Simultaneously, further development of the collaborative approaches is required. Automation may build on data-mining techniques to parse data and propose a proto-ontology, which can subsequently be defined through a collaborative approach.

During the collaborative approach described in this thesis, different types of confusion on the meaning of concepts and terms (Wien et al., 2009) have been experienced due to dialects and methaphors. Dialects (Wear, 1999; Bracken and Oughton, 2006) are the specialized languages used by each of the disciplines. One consequence of dialects is that the same concept is used for different meanings. This happened with the definition of crops across the models (Chapter 6). Another consequence of dialects is that different concepts might be used, which have the same meaning. Goble and Stevens (2008) provide a powerful example: WS-1 protein has ten different names. Metaphors are abstract notions used within a context or discipline to illuminate an argument, develop thinking in a new direction or refer to the unknown and these metaphors might become so entrenched that they seem true or real (Wear, 1999; Bracken and Oughton, 2006). An example of a metaphor in this thesis is the concept scenario (Chapter 7). Another example are the concepts exogenous and endogenous in models from economics versus parameter and variable in biophysical models. In the collaborative approach described in this thesis, one other type of confusion was experienced both with dialects and metaphors. The type of confusion concerns relationships between concepts, that are understood in a different way across or even within disciplines. An example is the complex relationships between NUTS regions, agri-environmental zones, farm types and representative farms (Chapter 5).

Dialects are easiest to solve, because a different understanding of concepts is relatively easy to identify. Metaphors require time and effort, as meaning of a concept is vague and abstract or many meanings exist due to the large number of participants involved. Clarifying metaphors entails defining new concepts and relationships, and researchers might not be willing to give up the relative freedom of the vagueness. These metaphors typically occur if researchers are not sure about something or need a container term to hide poorly defined concepts. Confusion on relationships is the most challenging to identify, because the differences in understanding only become apparent through detailed discussion or inspection of data sources when there is already agreement on the meaning of concepts. Further research is required to explicitly study confusion occurring in collaborative approaches and corroborate or adjust the types of confusion suggested here.

8.2.3. Technical integration

8.2.3.1. Modelling and software engineering

Technical integration is about integration of programming paradigms, data and models into modelling frameworks and graphical user interfaces. As the issues with respect to technical integration are similar for bio-economic farm models and for integrated assessment and modelling, these will be jointly discussed. In technical integration of models, methods from computer science and domain modelling have to be combined into a modelling framework. Domain modelling refers to model development, implementation and testing as done in a specific scientific domain (e.g. agronomy or economics). Computer science develops different types of technologies (e.g. ontology, OpenMI, tiered-applications, object- or aspect-oriented programming or design patterns) to facilitate building large complex computer-based programs. In computer science a strong push exists to innovate by adopting or developing the latest technologies. In domain modelling, the focus is on the conceptual model development and it is sufficient to work with one familiar technology, which fulfills the modelling requirements.

To link models in a modeling framework that is easy to maintain, extend and applicable for a range of models from different domains, advanced technologies from software engineering are required. Technologies used in domain modelling are not capable of managing multiple models and providing persistent data storage and advanced user interfaces. Advanced technologies from software engineering may have been applied at a small scale to one or two models, but applying them on a larger scale with many models covering multiple disciplines and scales is challenging and requires high resource investments. During the technical integration in the case studies, it was noted that software engineers had a tendency to overestimate the ease of adoption and experienced benefits of advanced IT technologies. Consequently, modellers had high expectations of what these technologies can deliver. On the other hand, modellers can be highly skeptical and think that the modeller must decide solely on the fly what is required (Hinkel, 2009). Such modellers are reluctant to change or fundamentally adapt their model to such technologies and path dependence exists to stick to the existing source code implementation of a model.

An example from the technical integration in the case studies is that the software engineers were used to work with relational database systems to ensure integrity, accuracy, intellectual property rights and consistency of data, while the modellers preferred binary formats (e.g. GDX for GAMS (GAMS, 2008)) or spreadsheet files. The modellers called these spreadsheets or binary files databases. Adopting relational database systems required investment from the modellers to learn working with a remote server containing their data and to adapt their model to retrieve data from a relational database system. Another example from the case studies, again related to data management, is the use of Hibernate (JBOSS, 2008), a technology to map class definitions from object-oriented programming to relational database entities. Hibernate is interoperable between different database dialects (e.g. PostgreSQL, MySQL or Oracle), which was presented by the software engineers as interoperability between different 'databases.' This led to the fallacy by modellers that with Hibernate they can just replace their spread-sheet database by a PostgreSQL database developed for another purpose. The software engineers forgot to explain that for interoperability purposes these databases must adhere to exactly the same database schema. In conclusion, cooperation between domain modelling and computer science might be problematic due to expectations held in each discipline and the gap in common practices in modelling and software engineering. During the technical integration described in this thesis, obstacles existed in modelling practices that inhibit an easy integration of models into a modelling framework. A first obstacle was a poor separation of data and algorithms in many models, which makes it difficult to achieve a comprehensive definition of all inputs required for the model. Models can have (very) complex data types, which have to be made explicit for the model to be integrated in a model chain. A second obstacle was legacy code (Feathers, 2004), which is working code for a specific purpose under a set of assumptions, but it is used for other purposes with different assumptions. Legacy code typically lacks adequate documentation and tests describing the purpose and assumptions of the source code and is usually understood by few developers. A third obstacle was a modelling practice to calibrate a model in an ad hoc way through its parameter values (Chapter 2). An ad hoc calibration may evolve into manipulating or tweaking models. An ad hoc calibration cannot be done when integrating a model with other models and running it in an automated fashion through a modelling framework. These three obstacles are advantages, when an individual modeller develops its own model, as it provides full flexibility to the creative modeller to investigate and learn from the behavior of his model. It is proposed to first develop and test a model separately from the other models before to link the tested model to a chain of models through an explicit model integration procedure (Chapter 4). During the separate development the modeller must keep in mind that the model will eventually be linked. Legacy code must be documented, model parameters separated from algorithms and procedures to automatically calibrate the model developed to avoid ad hoc calibration procedures in the separate development.

8.2.3.2. Modelling framework

To bridge the gap between modellers and software engineers, first, scientists are needed who are familiar with both modelling techniques and software engineering.

Second, an architecture of the modelling framework is required that allows both groups to do what they do best: modelling or programming. The architecture in SEAMLESS-Integrated Framework (Wien et al., 2009) allowed this by leaving the models in their original programming paradigm and using wrappers to translate between the programming paradigms of the different models, the framework and the database (Fig. 8.1). Although the architecture leaves the models relatively untouched, the models loose their direct link to the database or data-source. The development of the wrappers is a tedious, difficult and time-consuming task. Each wrapper is specific to a model. The wrapper therefore has to be changed, if the model changes, which is difficult for maintenance. The development of a wrapper required good and intensive communication between the modeller and the software engineer.

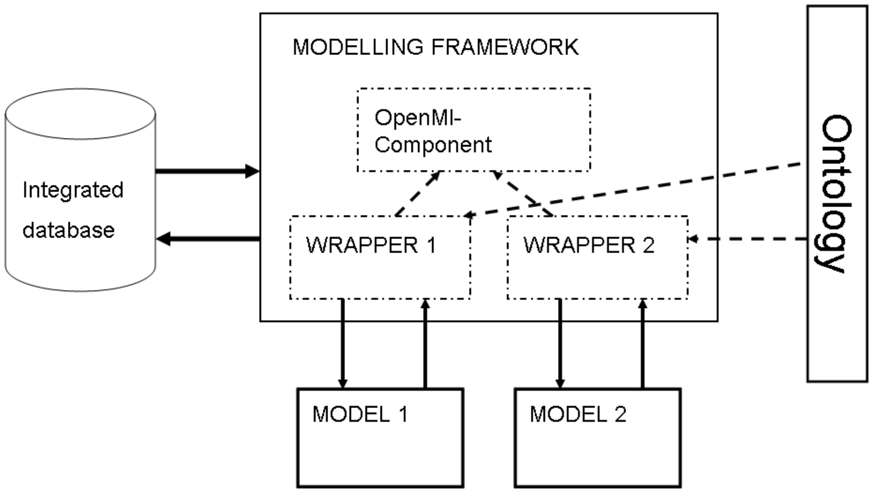


Figure 8.1. Architecture of the modelling framework in SEAMLESS (Source: Wien et al., 2009)

The ontology developed through the semantic integration was extensively used in the technical integration. The ontology covers all data-structures exchanged between models, database, wrappers and modelling framework. It does not cover the methods (i.e. also referred to as algorithms or process descriptions) of the models, wrappers and modelling framework, which would be a useful and challenging extension for the future. Such an extension would be comparable to alignment of models by mathematical formalisms as done by Hinkel (2008).

Through the SeRiDA framework (Athanasiadis et al., 2007b) source code was generated for the database, wrappers, models (if these are in the JavaTM programming language like FSSIM-AM (Chapter 3 and 4)) and modelling framework. By generating source code on the basis of the ontology, modellers and software engineers were less able to hide knowledge or data in the source code of the model, wrapper or database. Specific changes made to the generated source code are deleted and lost, whenever the source code was regenerated. The ontology was only updated a few times for each prototype and provided a stable basis for development of applications by modellers and software engineers. Frequent updates of the ontology by a group of developers might lead to chaotic development, because many different views exist on the optimal data structure in different programming paradigms.

8.3. Conclusions

8.3.1. Bio-economic farm models

This thesis and in particular Chapters 2 to 4 lead to the following conclusions as to BEFMs. BEFMs have potential to be useful in integrated assessment of policy changes and technological innovation. Current shortcomings hindering assessment of policy changes and technological innovations with BEFMs are in model description, availability of source code or program, and evaluation. Also, there is a lack of incorporation of strategic decision making and social environment and of an easily usable generic BEFM. Using BEFMs in integrated assessment requires an explicit procedure to define alternative activities that farms might take up in the future. Such a procedure must be based on generating alternative enterprises and management of those enterprises on the basis of knowledge on production ecology and economics. Farm System Simulator (FSSIM) is a generic BEFM, as it can be applied to a range of agri-environmental conditions, farm types and purposes at different levels of detail and with links to macro-economic and cropping and livestock system models.

8.3.2. Integration in Integrated Assessment and Modelling

This thesis and in particular Chapters 5 to 7 lead to the following conclusions as to IAM. Integration appears as a multi-headed Hydra snake for IAM-projects, due to the many different types of integration that have to be achieved in parallel and due to the important role of communication in integration. If the IAM project can manage all but one head of the Hydra, it is bound to fail in achieving integration and succeeding as an interdisciplinary research project.

In IAM projects that link models into a model chain, an integration of models, data-sources and diverse scenario definitions is required. In this thesis an integrated database on European agricultural systems was described. Such an integrated database has the advantages that: a range of data sources are available on one location; difficult questions of data integration and consistency have been solved by specialists familiar with the original data sources and the preprocessing of the original data sources is already done. An integrated database can be used for all tools in an IAM, if these tools are aligned with the content of the database. Linking models in a model chain for IAM requires the explicit and deliberate methodological, semantic and technical integration, e.g. aligning different models in their process definitions space, time and modelling techniques, developing a shared conceptual language across models and ensuring execution of models on a computer. Although literature provides many advanced and complex definitions and classifications of scenarios, these definitions and classifications cannot be made operational in development of tools in IAM. The common ontology on assessment projects and scenarios presented in Chapter 7 provides an operational and simple definition of scenarios and assessment projects.

8.4. Communication and organization for integration

'In days of scarce resources we need to keep a very clear eye …on priorities. These are often best understood by the science community but trust has been so eroded that the purchasers will not listen to the science community who are being perceived to be 'barrow pushing'. After all, … the scientists always have a vested interest and are often cast in the guise of mendicant priests.' (Harris, 2002)

8.4.1. Introduction

The Chapters 2 to 7 from this thesis are developed in the context of the SEAMLESS integrated framework project. Methodological, semantic and technical integration require an organizational structure to facilitate it and communication to exchange information and to achieve consensus, especially in large IAM projects like SEAMLESS. This Section will consider the methodological, semantic and technical integration as described in Chapters 2 to 7 in the broader context of an interdisciplinary IAM research project, like SEAMLESS. The broader context concerns social and institutional barriers to integration and more specifically, the communication and the organization of an IAM research project. First, recommendations for communication and organizational structures based on Jakobsen and McLaughlin (2004) will be presented. Second, these recommendations are related to methodological, semantic and technical integration in SEAMLESS. Third, a reflection is offered on the topic of integration for interdisciplinary research, considering methodological, semantic and technical integration, communication strategies and organizational structure.

A successful integration depends on "good" communication (Harris, 2002; Bruce et al., 2004; Jakobsen and McLaughlin, 2004), which sounds obvious. It is not obvious provided that only Jakobsen and McLaughlin (2004) explicitly discuss communication and organization in an interdisciplinary project, in this case on integrated ecosystem management. They found that an organized formal communication process between project members across personal and disciplinary boundaries is required to achieve integration. They recommend that such an organized communication process includes i. training on crucial methods and concepts, ii. a recruitment process to select group members, iii. a strategy on types, timing and forms of communication, and iv. interactive activities that facilitate integration (Jakobsen and McLaughlin, 2004). These four recommendations are used to evaluate the organization and communication processes in SEAMLESS.

In our integration effort as part of the SEAMLESS project, different communication strategies were used with respect to methodological, semantic and technical integration. The SEAMLESS project was organized in work packages and tasks within these work packages. The participants in a task usually had a similar disciplinary or research background, while work packages were interdisciplinary. For instance, in the modelling work package, the modellers were divided in one task for biophysical modelling, one task for farm-economic modelling, one task for upscaling of supply responses and one task for market level economic modelling.

8.4.2. Methodological integration

Methodological integration was achieved through an organized communication process as suggested by Jakobsen and McLaughlin (2004) by organizing a group of senior researchers[1] from different disciplines and institutes in one integrative work package. Each senior researcher represented one of the other work packages. The work package targeted to achieve the methodological integration and to manage the progress in the SEAMLESS project. Although the tasks of methodological integration and management can potentially complement each other, they can also be conflicting for time and resources. For example, in discussing a link between two models, the senior researchers had to discuss the conceptual link, to decide on the relevant junior researchers[2] to work on the link and to set the time-planning to finish

[1] Senior researcher is a researcher in a managing and coordinating role that requires leadership. Senior does not refer to the age of the researcher.

[2] Junior researcher is a researcher in an executive role that work on project tasks on a day-to-day basis and hands-on fashion. Junior does not refer to the age of the researcher.

the work. In discussing the conceptual link, attention could too rapidly turn to who to work on it or when to finish it. Also, senior researchers could try to keep the integration work away from their own work package by claiming that their work package had done its work. This resulted in finger-pointing at the other work packages and the integrative work package.

The senior researchers had many other tasks both within and outside the project, and had limited time to go into detail in many problems of methodological integration. Junior researchers worked on the methodological integration on a day-to-day basis, but these were initially excluded from the integrative work package. Information on methodological integration relevant for the work package had to be passed through the senior researchers, resulting in long communication lines. The junior researchers were sometimes missing information on decisions taken by the work package or the opportunity to provide ideas or explanations of their work. During the project, some junior researchers were added to this group, which led to an improved sharing of relevant information and decisions, and it enhanced progress.

8.4.3. Semantic integration

'Economists, as well as those in other fields, communicate mainly with powerful figures of speech—in particular metaphors and appeals to authority—that offer up a compact and rich way of communicating within a peer group (even when these figures are enacted without a full understanding of their content). They also have the effect of excluding others from the conversation.' (Wear, 1999)

An organized communication process was not initially planned to realize the semantic integration, although a knowledge base that integrated the knowledge of the project was a planned product. The development of the scientific content of the knowledge base required the involvement of all disciplines, e.g. computer scientists, economists, agronomists and landscape ecologists. In the proposal, the responsibility for the semantic integration effort was held by none of the disciplines. In year 2 of the project, a task force was started to establish such an organized communication process, as by that time the lack of responsibility and progress was identified. This task force consisted of a small group of knowledge engineers and of domain experts, which were involved upon request, from different parts of the project. The task force started with the drafting of a work plan, in which the role and the mission of the task force were formalized. This work plan was crucial to build a shared vision between participants and clearly define the desired products of the task force. Actions of the task force included the development of an integrated database (Chapter 5), definition of exchanged data types between the models (Chapter 6), joint conceptualization of scenarios and assessment projects (Chapter 7), definition of indicators and concepts related to indicators (Thérond et al., 2009), specifications to bridge programming paradigms (Athanasiadis et al., 2007c) and a browser of the knowledge base to disseminate the content.

The membership of the task force was on a voluntary basis, which had as advantage that participants were motivated to participate and that more participants could easily join the task force, after initial promising actions of the task force generated interest. Disadvantages were that there was considerable time needed to involve participants and that there was a high turn-over rate of participants as many participants left the group due to other priorities on top of participants changing jobs. The task force worked in an iterative process of developing prototypes and a final version. At the end of each iteration, a version of the knowledge base was delivered. The setting up of each prototype started with planning of the activities for that prototype and the interactions required.

8.4.4. Technical integration

From an organizational point of view, the technical integration had to be jointly achieved between the work packages of the modellers and computer scientists. There was no organized communication process established at the start of or during the project for modellers and computer scientists to form a group. There was considerable confusion between the two groups, because the modellers could not understand what the computer scientists wanted to achieve with their technologies and the computer scientists did not always understand what the models made by the modellers were doing. This led to wrong expectations of what models or IT technologies could do, and diverted attention from the task at hand, which was a rather laborious one of converting data from one programming language into another programming language, e.g. from JavaTM to C# through XML. One part of the technical integration was resolved in the task force for semantic integration, which was the integration of the ontology with the database (Chapter 4 and Athanasiadis et al., 2007c).

Ultimately, the technical integration was solved on an ad hoc basis involving both modellers and computer scientists. This required many interventions from senior researchers managing the project. Some modellers with adequate IT skills and one computer scientist developed the wrappers for the models and played a bridging role between the modellers and computer scientists. Towards the end of the project, an intensive self-organizing communication process was established to facilitate close cooperation between modellers and computer scientists. This communication process consisted of daily conference calls, in which each participant could update on his or her progress made and plans for that day, a continuous chat conversation and meetings to jointly develop source code. The intensive self-organising communication process was very successful to achieve rapidly the technical integration.

8.4.5. Lessons learned

Table 8.2 summarizes the organized communication processes used for methodological, semantic and technical integration and links these with the recommendations provided by Jakobsen and McLaughlin (2004). Technical integration shows the poorest performance as it only weakly fulfills one of the four recommendations and as organized communication processes were not deliberately and formally established. Semantic and methodological integration are strong on different recommendations, but do not fulfill the recommendations to the same extent. The main weakness of the task force used for semantic integration was the voluntary participation of researchers, which caused problems as people left or felt little responsibility to continue their work. For the methodological integration, participation of members was ensured through formal roles in the project and the project progress being discussed at the same time. Semantic integration was strong in interactive activities, which took the form of workshops in which domain content and the work of the task force were presented and discussed. Overall, adhering to the recommendations by Jakobsen and McLaughlin (2004) led to more scientific output (e.g. tools, publications) and a lower managerial load to solve problems of unclear tasks or roles of researchers.

Table 8.2. Using four recommendations from Jakobsen and McLaughlin (2004) to evaluate the communication strategies in the SEAMLESS project for methodological, semantic and technical integration 0 = not present; + = present with major shortcomings; ++ = present, although not explicitly planned; +++ = present and planned.

Recommendations Jakobsen and McLaughlin (2004)	Methodological integration	Semantic integration	Technical integration
Training on crucial methods and concepts	++	++	0
Strategy to recruit participants	+++	+	0
Strategy of communication	++	++	++
Interactive activities	0	+++	0

For future interdisciplinary IAM projects, it is recommended that organized communication processes are ensured for methodological, semantic and technical integration. The success of the methodological, semantic and technical integration was influenced by the initial organization of the project in work packages and fairly disciplinary tasks. A suitable organizational structure cannot be derived from this research, but a requirement is that it somehow ensures that an integration group is formally organized. The integration group i. uses a recruitment process to involve both senior and junior participants, ii. writes a communication and interaction plan,

iii. starts with training on the integration and scientific methods, and iv. jointly drafts an adaptable work plan with a clear mission and vision statement and clearly defined integrated products. The rest of the project can be organized according to more disciplinary lines. One option is to set up an integration group as part of the research project, which has a mix of senior and junior researchers. A second option is to organize the whole project along the lines of methodological, semantic and technical integration.

8.4.6. Interdisciplinary research and integration

'The division-of-labor model of separate departments is obsolete and must be replaced with a curriculum structured like a web or complex adaptive network. Responsible teaching and scholarship must become cross-disciplinary and cross-cultural. … There can be no adequate understanding of the most important issues we face when disciplines are cloistered from one another and operate on their own premises.' (Taylor, 2009)

If it is desired to combine the reductionist disciplinary solutions to an interdisciplinary answer to large societal problems (e.g. poverty, hunger, climate change, biodiversity loss, soil salinity), methods for integration in interdisciplinary large research projects are required. Such large interdisciplinary research projects are funded at different levels, e.g. the Interdisciplinary Research and Education Fund (INREF) programme of Wageningen University (WU, 2009), Besluit Subsidies Investeringen KennisInfrastructuur (BSIK) programme in the Netherlands (SenterNovem, 2009), the Seventh Framework programme for Research and Technological Innovation of the European Union (EC, 2009) and Challenge programmes of CGIAR for the world (CGIAR, 2009). This thesis offers insights into the methodological, semantic and technical integration for large research projects with more than two disciplines and ten researchers. This thesis only provides a start as the methods proposed need to be evaluated, tested and developed in other settings and as many questions are still unanswered. For example, if the communication processes are formally planned and organized as recommended by Jakobsen and McLaughlin (2004) and this thesis, are these successful? What type of communication processes and integration methods fit with the nature of the researcher as a "lone boffin" (Harris, 2002)? What are suitable methods to achieve simultaneously methodological, semantic and technical integration? What are organizational structures that facilitate and nurture integration in interdisciplinary research?

Literature (Bruce et al., 2004; Jakobsen and McLaughlin, 2004; Tress et al., 2007; Hinkel, 2008) is scant and anecdotal on methods for integration in large research projects. Pleas (Norgaard, 1992; Metzger and Zare, 1999; Harris, 2002; Parker et al., 2002) are available calling for more integration and emphasis on interdisciplinary research, but these do not provide a unified view. With institutional, social, methodological, semantic and technical barriers to integration, interdisciplinary research in large projects is more likely to fail than to succeed. With such a high chance of failure, development of methods for integration cannot be expected to come about by itself, even if funding agencies encourage interdisciplinary research (Metzger and Zare, 1999; Bruce et al., 2004). A shift in research needs to occur from the current accidentally achieved good practices in integration (Jakobsen and McLaughlin, 2004; Hinkel, 2008; this thesis) to a deliberate formal development of integration methods for research.

References

Abadi Ghadim, A. K., 2000. Water repellency: a whole-farm bio-economic perspective. Journal of Hydrology 231-232, 396-405.

AbouZahr, C., Adjei, S., Kanchanachitra, C., 2007. From data to policy: good practices and cautionary tales. The Lancet 369 (9566), 1039-1046.

Acs, S., Berentsen, P. B. M., Huirne, R. B. M., 2007. Conversion to organic arable farming in The Netherlands: A dynamic linear programming analysis. Agricultural Systems 94 (2), 405-415.

Adenäuer, M., Kuiper, M., 2009. A typical application of SEAMLESS-IF at macro level: a trade-liberalization scenario applied to the EU. PD6.3.5.1, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org, Wageningen, pp. 51.

Alcamo, J., 2001. Scenarios as tools for international environmental assessments. Environmental issue report No. 24 Experts' corner report Prospects and Scenarios No 5, European Environment Agency, Copenhagen, http://reports.eea.europa.eu/environmental_issue_report_2001_24/en, pp. 31.

Ali, R., Wheitner, D., Talbott, E., Zborowski, J., 2007. Connecting environmental health data to people and policy: integrating information and mobilizing communities for environmental public health tracking. Journal of Community Health 32 (5), 357-374.

Alkan Olsson, J., Bockstaller, C., Stapleton, L. M., Correia, T. P., Geniaux, G., Bellon, S., Turpin, N., 2009. Indicator frameworks supporting impact assessment of new policies for rural systems – a review of a goal oriented frameworks and its indicators. Environmental Science & Policy In press.

Alterra, INRA, 2005. New soil information for CGMS (Crop Growth Monitoring System) (SINFO). Alterra - INRA, Wageningen, pp. 219.

Andersen, E., Elbersen, B., Godeschalk, F., Verhoog, D., 2007a. Farm management indicators and farm typologies as a basis for assessments in a changing policy environment. Journal of Environmental Management 82 (3), 353.

Andersen, E., Elbersen, B., Godeschalk, F., Verhoog, D., 2007b. New policy objectives – new farm typologies. In: Poppe, K., Boone, K., Teeuwen-Vogelaar, C. (Eds.), Changes in farming and the effects on FADNs. Agricultural Economic Institute (LEI) The Hague, pp. 111-121.

Andersen, E., Elbersen, B. S., Hazeu, G., Van Diepen, C., Baruth, B., Verhoog, D., Terluin, I., Borkowski, N., Janssen, S., 2007c. The environmental component, the farming systems component and the socio-economic component of the SEAMLESS database for the Prototype 2. D4.3.5-D4.4-D4.5.4., SEAMLESS integrated project, EU 6th Framework Programme, (contract no. 010036-2), www.SEAMLESS-IP.org, pp. 195.

Anderson, J., Dillon, J., Hardaker, J., 1985. Socio-economic modelling of farming systems. In: Hardaker, J. B. (Ed.) Agricultural systems research for developing countries: proceedings of an international workshop, Hawkesbury Agricultural College, Richmond, N.S.W., Australia, 12-15 May 1985, pp. 77-88.

Annetts, J. E., Audsley, E., 2002. Multiple objective linear programming for environmental farm planning. Journal of the Operational Research Society 53, 933-943.
Antle, J. M., Capalbo, S. M., 2001. Econometric-Process Models for Integrated Assessment of Agricultural Production Systems. American Journal of Agricultural Economics 83 (2), 389-401.
Antoniou, G., Van Harmelen, F., 2004. A semantic web primer. The MIT Press, Cambridge, Massachusetts; London, England, pp. 238.
Apland, J., 1993. The use of field days in economic models of crop farms. Journal of Production Agriculture 6 (3), 437-444.
Athanasiadis, I. N., Janssen, S., 2008. Semantic mediation for environmental model components integration. Information Technologies in Environmental Engineering 1, 3-11.
Athanasiadis, I. N., Janssen, S., Andersen, E., Rizzoli, A. E., Wien, J. J. F., 2009. Granular ontologies for integrated assessment of agricultural systems. In: Proceedings of Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Policy and Society (AgSap), Egmond aan Zee.
Athanasiadis, I. N., Rizzoli, A. E., Janssen, S., Van Ittersum, M. K., 2007a. Generation of artificial crop rotation schemes as cyclic permutations. In: Donatelli, M., Hatfield, J., Rizzoli, A. E. (Eds.), Farming Systems Design 2007: An International Symposium on Methodologies on Integrated Analysis on Farm Production Systems, Catania, Italy.
Athanasiadis, I. N., Villa, F., Rizzoli, A. E., 2007b. Enabling knowledge-based software engineering through semantic-object-relational mappings. In: Stojanovic, L., Sabbouh, M. (Eds.), 3rd International Workshop on Semantic Web Enabled Software Engineering, 4th European Semantic Web Conference, Innsbruck, Austria, pp. 15.
Athanasiadis, I. N., Villa, F., Rizzoli, A. E., 2007c. Ontologies, JavaBeans and relational databases for enabling semantic programming. In: 31th IEEE Annual International Computer Software and Applications Conference (COMPSAC), Beijing, China pp. 341-346.
Athanasiadis, I., Rizzoli, A. E., Donatelli, M., Carlini, L., 2006. Enriching software model interfaces using ontology-based tools. In: Voinov, A., Jakeman, A., Rizzoli, A. (Eds.), the iEMSs Third Biannual Meeting "Summit on Environmental Modelling and Software Third Biennial Meeting", Burlington, USA.
Austin, E. J., Willock, J., Deary, I. J., Gibson, G. J., Dent, J. B., Edwards-Jones, G., Morgan, O., Grieve, R., Sutherland, A., 1998. Empirical models of farmer behaviour using psychological, social and economic variables. Part I: linear modelling. Agricultural Systems 58 (2), 203.
Bachinger, J., Zander, P., 2007. ROTOR, a tool for generating and evaluating crop rotations for organic farming systems. European Journal of Agronomy 26 (2), 130-143.

Barbier, B., Bergeron, G., 1999. Impact of policy interventions on land management in Honduras: results of a bioeconomic model. Agricultural Systems 60 (1), 1-16.
Bartolini, F., Bazzani, G. M., Gallerani, V., Raggi, M., Viaggi, D., 2007. The impact of water and agriculture policy scenarios on irrigated farming systems in Italy: An analysis based on farm level multi-attribute linear programming models. Agricultural Systems 93 (1-3), 90-114.
Baruth, B., Genovese, G., Montanarella, L., 2006. New soil information for the MARS Crop Yield Forecasting System. Joint Research Centre of European Commission, Ispra, http://eusoils.jrc.it/projects/sinfo/pdf/mars_crop_yield_forecasting_system.pdf, pp. 95.
Beaumont, R., Dulphy, J. P., Sauvant, D., Tran, G., Mesch, F., Aufrère, J., Peyraud, J. L., Champciaux, P., 2007. Les tables de valeurs des aliments, Alimentation des Bovins, Ovins et Caprins. In: Quae (Ed.) Besoins des animaux valeurs des aliments. pp. 476.
Belhouchette, H., Wery, J., 2009. A typical application of SEAMLESS-IF at meso level:The Nitrate Directive and Water Framework Directive scenarios in Midi-Pyrenees. PD6.3.5.2, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org, Wageningen, pp. 36.
Bellman, R., 1957. Dynamic Programming. Princeton University Press, Princeton, New Jersey, USA.
Benoit, M., Veysset, P., 2003. Conversion of cattle and sheep suckler farming to organic farming: adaptation of the farming system and its economic consequences. Livestock Production Science 80 (1-2), 141-152.
Berentsen, P. B. M., 2003. Effects of animal productivity on the costs of complying with environmental legislation in Dutch dairy farming. Livestock Production Science 84 (2), 183-194.
Berentsen, P. B. M., Giesen, G. W. J., Schneiders, M. M. F. H., 1998. Conversion from conventional to biological dairy farming: Economic and environmental consequences at farm level. Biological Agriculture and Horticulture. 16 (3), 311-328.
Berentsen, P., Giesen, G., 1994. Economic and environmental consequences of different governmental policies to reduce N losses on dairy farms. Netherlands Journal of Agricultural Science 42 (1), 11-19.
Berger, T., 2001. Agent-based spatial models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis. Agricultural Economics 25 (2-3), 245-260.
Bergevoet, R. H. M., Ondersteijn, C. J. M., Saatkamp, H. W., Van Woerkum, C. M. J., Huirne, R. B. M., 2004. Entrepreneurial behaviour of dutch dairy farmers under a milk quota system: goals, objectives and attitudes. Agricultural Systems 80 (1), 1-21.
Berners Lee, T., Hall, W., Hendler, J., Shadbolt, N., Weitzner, D. J., 2006. Creating a science of the web. Science 313, 769-771.

Berntsen, J., Petersen, B. M., Jacobsen, B. H., Olesen, J. E., Hutchings, N. J., 2003. Evaluating nitrogen taxation scenarios using the dynamic whole farm simulation model FASSET. Agricultural Systems 76 (3), 817-839.
Beukes, P. C., Cowling, R. M., Higgins, S. I., 2002. An ecological economic simulation model of a non-selective grazing system in the Nama Karoo, South Africa. Ecological Economics 42 (1-2), 221-242.
Blanco Fonseca, M., Flichman, G., 2002. Dynamic optimisation problems: different resolution methods regarding agriculture and natural resource economics. Working Paper, Universidad Politécnica de Madrid and CIHEAM-Institut Agronomique Méditerranéen de Montpellier Montpellier, pp. 35.
Bland, W. L., 1999. Toward integrated assessment in agriculture. Agricultural Systems 60 (3), 157-167.
Borjeson, L., Hojer, M., Dreborg, K.-H., Ekvall, T., Finnveden, G., 2006. Scenario types and techniques: Towards a user's guide. Futures 38 (7), 723.
Borkowski, N., Zander, P., Stokstad, G., Uthes, S., Reinhardt, F.-J., Bellocchi, G., Hecker, M., 2007. Conceptual approach to identify and assess current activities. PD3.3.9, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org, Wageningen, pp. 23.
Bos, J. F. F. P., 2000. Agronomic, environmental and socio-economic characteristics of prototype mixed farming systems in Flevoland: a modelling approach. In: Gagnaux, D., Poffet, J. (Eds.), Livestock-farming-systems:-integrating-animal-science-advances-into-the-search-for-sustainability. Proceedings-of-the-fifth-international-symposium-on-livestock-farming-systems, Posieux, Fribourg, Switzerland, 19-20 August,1999; 5 ref., pp. 290-293.
Bouma, J., Stoorvogel, J., Van Alphen, B. J., Booltink, B. W. G., 1999. Pedology, Precision Agriculture, and the Changing Paradigm of Agricultural Research. Soil Science Society of America Journal 63, 1763-1768.
Bracken, L. J., Oughton, E. A., 2006. 'What do you mean?' The importance of language in developing interdisciplinary research. Transactions of the Institute of British Geographers 31 (3), 371-382.
Brandmeyer, J. E., Karimi, H. A., 2000. Coupling methodologies for environmental models. Environmental Modelling and Software 15 (5), 479.
Brilhante, V., Ferreira, A., Marinho, J., Pereira, J. S., 2006. Information Integration through Ontology and Metadata for Sustainability Analysis. In: Voinov, A., Jakeman, A., Rizzoli, A. (Eds.), the iEMSs Third Biannual Meeting "Summit on Environmental Modelling and Software Third Biennial Meeting", Burlington, USA.
Britz, W., Pérez Domínguez, I., Heckelei, T., 2009. A comparison of CAPRI and SEAMLESS-IF as Integrated Modelling Systems. In: Brouwer, F., Van Ittersum, M. K. (Eds.), Environmental and agricultural modelling: integrated approaches for policy impact assessment. Springer.
Britz, W., Pérez, I., Zimmermann, A., Heckelei, T., 2007. Definition of the CAPRI core modelling system and interfaces with other components of SEAMLESS-IF. SEAMLESS Report No.26, SEAMLESS integrated project,

EU 6th Framework Programme, contract no.010036-2, www.SEAMLESS-IP.org, pp. 116.
Bruce, A., Lyall, C., Tait, J., Williams, R., 2004. Interdisciplinary integration in Europe: the case of the Fifth Framework programme. Futures 36 (4), 457-470.
Buccella, A., Cechich, A., Fillottrani, P., 2009. Ontology-driven geographic information integration: A survey of current approaches. Computers & Geosciences 35, 710-723.
Cain, P., Anwar, M., Rowlinson, P., 2007. Assessing the critical factors affecting the viability of small-scale dairy farms in the Punjab region of Pakistan to inform agricultural extension programmes. Agricultural Systems 94 (2), 320-330.
Calker, K. J. v., Berentsen, P. B. M., de Boer, I. M. J., Giesen, G. W. J., Huirne, R. B. M., 2004. An LP model to analyse the economic and ecological sustainability on Dutch dairy farms: model presentation and application for experimental farm. Agricultural Systems 82, 139-160.
Carberry, P. S., Hochman, Z., McCown, R. L., Dalgliesh, N. P., Foale, M. A., Poulton, P. L., Hargreaves, J. N. G., Hargreaves, D. M. G., Cawthray, S., Hillcoat, N., Robertson, M. J., 2002. The FARMSCAPE approach to decision support: farmers', advisers', researchers' monitoring, simulation, communication and performance evaluation. Agricultural Systems 74 (1), 141-177.
Carpenter, S. R., Mooney, H. A., Agard, J., Capistrano, D., DeFries, R. S., DÃ-az, S., Dietz, T., Duraiappah, A. K., Oteng-Yeboah, A., Pereira, H. M., Perrings, C., Reid, W. V., Sarukhan, J., Scholes, R. J., Whyte, A., 2009. Science for managing ecosystem services: Beyond the Millennium Ecosystem Assessment. Proceedings of the National Academy of Sciences 106 (5), 1305-1312.
Castellazzi, M. S., Wood, G. A., Burgess, P. J., Morris, J., Conrad, K. F., Perry, J. N., 2008. A systematic representation of crop rotations. Agricultural Systems 97 (1-2), 26-33.
CGIAR, 2009. CGIAR Challenge Program. http://www.cgiar.org/impact/challenge/index.html.
Cobb, C. W., Douglas, P. H., 1928. A Theory of Production. The American Economic Review 18 (1), 139-165.
Coelli, T., Rao, D. S. P., Battese, G., 1998. An Introduction To Efficiency And Production Analysis. Kluwer Academic Publishers Boston
Cohen, S. J., 1997. Scientist–stakeholder collaboration in integrated assessment of climate change: lessons from a case study of Northwest Canada. Environmental Modelling & Assessment 2 (4), 281-293.
Confalonieri, R., Bocchi, S., 2005. Evaluation of CropSyst for simulating the yield of flooded rice in northern Italy. European Journal of Agronomy 23 (4), 315-326.
Csaki, C., 1977. Dynamic linear programming model for agricultural investment and resources utilization politics. I.I.A.S.A., Laxenburg, pp. 28.

De Buck, A. J., Hendrix, E. M. T., Schoorlemmer, H. B., 1999. Analysing production and environmental risks in arable farming systems: A mathematical approach. European Journal of Operational Research 119 (2), 416-426.
De Koning, G. H. J., Van Keulen, H., Rabbinge, R., Janssen, H., 1995. Determination of input and output coefficients of cropping systems in the European Community. Agricultural Systems 48 (4), 485-502.
De Ridder, N., 1997. Hierarchical level in agro-ecosystems: selective case studies on water and nitrogen. Wageningen University, PhD thesis, Wageningen, pp. 193
De Wit, C. T., 1992. Resource use efficiency in agriculture. Agricultural Systems 40 (1-3), 125-151.
DeMichiel, L., Keith, M., 2006. Enterprise javabeans 3.0 specifications. http://jcp.org/aboutJava/communityprocess/pfd/jsr220/index.html.
Dent, J. B., 1990. Optimising the Mixture of Enterprises in a Farming System. In: Jones, J., Street, P. (Eds.), Systems Theory Applied To the Agriculture and the Food Chain. Elsevier Applied Science, London, pp. 113-130.
Dent, J. B., Edwards-Jones, G., McGregor, M. J., 1995. Simulation of ecological, social and economic factors in agricultural systems. Agricultural Systems 49 (4), 337-351.
Detlefsen, N. K., Jensen, A. L., 2007. Modelling optimal crop sequences using network flows. Agricultural Systems 94 (2), 566-572.
Deybe, D., Flichman, G., 1991. A regional agricultural model using a plant growth simulation program as activities generator-- an application to a region in Argentina. Agricultural Systems 37 (4), 369-385.
Dogliotti, S., Rossing, W. A. H., Van Ittersum, M. K., 2003. ROTAT, a tool for systematically generating crop rotations. European Journal of Agronomy 19 (2), 239-250.
Dogliotti, S., Rossing, W. A. H., Van Ittersum, M. K., 2004. Systematic design and evaluation of crop rotations enhancing soil conservation, soil fertility and farm income: a case study for vegetable farms in South Uruguay. Agricultural Systems 80 (3), 277.
Dogliotti, S., Van Ittersum, M. K., Rossing, W. A. H., 2005. A method for exploring sustainable development options at farm scale: a case for vegetable farms in South Uruguay. Agricultural Systems 86 (1), 29-51.
Donaldson, A. B., Flichman, G., Webster, J. P. G., 1995. Integrating agronomic and economic models for policy analysis at the farm level: The impact of CAP reform in two European regions. Agricultural Systems 48 (2), 163-178.
Donaldson, A. B., Flichman, G., Webster, J. P. G., 1995. Integrating agronomic and economic models for policy analysis at the farm level: The impact of CAP reform in two European regions. Agricultural Systems 48 (2), 163-178.
Donatelli, M., Russell, G., Rizzoli, A. E., Acutis, M., Adam, M., Athanasiadis, I., Balderacchi, M., Bechini, L., Belhouchette, H., Bellocchi, G., Bergez, J. E., Botta, M., Braudeau, E., Bregaglio, S., Carlini, L., Casellas, E., Celette, F., Ceotto, E., Charron-Moirez, M. E., Confalonieri, R., Corbeels, M.,

Criscuolo, L., Cruz, P., Guardo, A. D., Ditto, D., Dupraz, C., Duru, M., Fiorani, D., Gentile, A., Ewert, F., Gary, C., Habyarimana, E., .Jouany, C., Kansou, K., Knapen, M. J. R., Filippi, G. L., Leffelaar, P., Manici, L., Martin, G., Martin, P., Meuter, E. C., Mugueta, N., Mulia, R., Noordwijk, M. v., Oomen, R., Rosenmund, A., Rossi, V., Salinari, F., Serrano, A., Sorce, A., Vincent, G., Theau, J. P., Therond, O., Trevisan, M., Trevisiol, P., Van Evert, F. K., Wallach, D., Wery, J., Zerourou, A., 2009. APES: The Agricultural Production and Externalities Simulator In: Brouwer, F., Van Ittersum, M. K. (Eds.), Environmental and agricultural modelling: integrated approaches for policy impact assessment. Springer Academic Publishing.

Dorfman, R., Samuelson, P. A., Solow, R. M., 1958. Linear Programming and Economic Analysis. McGraw-Hill, New York, USA.

Dorward, A., 1999. Modelling embedded risk in peasant agriculture: methodological insights from northern Malawi. Agricultural Economics 21 (2), 191-203.

Dorward, A., Parton, K., 1997 Quantitative farm models and embedded risk in complex, diverse and risk prone agriculture. Quarterly Journal of International Agriculture 36, 317-330.

Dramstad, W., Sogge, C., 2003. Agricultural impacts on landscapes: developing indicators for policy analysis. In: Proceedings from an OECD Expert Meeting, Oslo, Norway, October 2002. NIJOS/OECD, Oslo, pp. 343.

EC, 1991. Nitrates Directive 91/676/EEC. 91/676/EEC.

EC, 1999. Contribution of the European Community on the Multifunctional Character of Agriculture. European Commission, DG Agriculture, Brussels.

EC, 2000. Directive 2000/60/EC of the European Parliament and of the Council, of 23 October 2000, establishing a framework for Community action in the field of water policy. Official Journal of the European Communities, 22.12.2000.

EC, 2002. Implementation of Council Directive 91/676/EEC concerning the protection of waters against pollution caused by nitrates from agricultural sources -Synthesis from year 2000 Member States reports- European Commission, Luxembourg.

EC, 2003. Council regulation (EC) No 1782/2003 of 29 September 2003. European Commission, No1782/2003, pp. 111.

EC, 2005. Impact Assessment Guidelines. European Commission, Brussels, pp. 99.

EC, 2006. Rural Development 2007-2013, Handbook on common monitoring and evaluation framework, Guidance document. European Commission, Directorate General for Agriculture and Rural Development, Brussels, pp. 15.

EC, 2007. Towards sustainable water management in the European Union - First stage in the implementation of the Water Framework Directive 2000/60/EC –. Communication from the Commission to the European Parliament and the Council, European Commission, Brussels, pp. 13.

EC, 2008a. Farm Accountancy Data Network (FADN) Source: EU-FADN-DG AGRI-G3. http://ec.europa.eu/agriculture/rica/.

EC, 2008b. Nomenclature of territorial units for statistics - NUTS Statistical Regions of Europe. http://ec.europa.eu/comm/eurostat/ramon/nuts/home_regions_en.html.
EC, 2009. Seventh Framework Programme (FP-7): Cooperation. http://cordis.europa.eu/fp7/cooperation/home_en.html.
Edwards-Jones, G., McGregor, M., 1994. The Necessity, Theory and Reality of Developing Models of Farm Households. In: Dent, J., McGregor, M. (Eds.), Rural and Farming Systems Analysis: European Perspectives. CAB International, Wallingford, pp. 338-352.
EEA, 2005. Agriculture and environment in EU-15 - the IRENA indicator report. European Environment Agency, Copenhagen, pp. 122.
El Gohary, N., El Diraby, T., 2005. Achieving ontology interoperability using formal concept analysis: an approach to inter-organizational collaboration. In: Scherer, R. J., Katranuschkov, P., Schapke, S.-E. (Eds.), Proceedings of 22nd CIB-W78 Conference on Information Technology in Construction, Dresden, pp. 1-8.
Elbersen, B., Kempen, M., Van Diepen, K., Hazeu, G., Verhoog, D., Andersen, E., 2006. Protocols for spatial allocation of farm types. SEAMLESS project deliverable (PD 4.7.1). SEAMLESS Integrated Project EU FP6 (contract nr. 010036), Wageningen.
ESBN, 2008. European Soil Database. http://eusoils.jrc.it/esbn/Esbn_overview.html.
Eurostat, 2008. European Dissemination Database (Formerly New Cronos). http://epp.eurostat.ec.europa.eu/.
Eurostat, 2008. European Dissemination Database (Formerly New Cronos). http://epp.eurostat.ec.europa.eu/.
Ewert, F., Van Ittersum, M. K., Bezlepkina, I., Therond, O., Andersen, E., Belhouchette, H., Bockstaller, C., Brouwer, F., Heckelei, T., Janssen, S., Knapen, R., Kuiper, M., Louhichi, K., Olsson, J. A., Turpin, N., Wery, J., Wien, J. E., Wolf, J., 2009. A methodology for integrated assessment of policy impacts in agriculture Environmental Science & Policy In press,
Falconer, K., Hodge, I., 2000. Using economic incentives for pesticide usage reductions: responsiveness to input taxation and agricultural systems. Agricultural Systems 63 (3), 175-194.
Falconer, K., Hodge, I., 2001. Pesticide taxation and multi-objective policy-making: farm modelling to evaluate profit/environment trade-offs. Ecological Economics 36 (2), 263-279.
FAO, 1976. A framework for land evaluation. FAO Soils bulletin 32, FAO and agricultural organization of the United Nations, Rome, pp. 72.
FAO, 2008. FAOSTAT. http://faostat.fao.org/.
Farquhar, A., Fikes, R., Pratt, W., Rice, J., 1995. Collaborative Ontology Construction for Information Integration. Knowledge Systems Laboratory, Department of Computer Science, Stanford University, http://citeseer.ist.psu.edu/103741.html, pp. 32.
Feathers, M., 2004. Working Effectively With Legacy Code. Prentice Hall PTR, pp. 456.

Flanagan, K., Stevens, R., Pocock, M., Lee, P., Wipat, A., 2005. Ontology for genome comparison and genomic rearrangements. Comparative and Functional Genomics 5 (6-7), 537-544.

Flichman, G., 1996. L'evaluation micro-economique de la reforme de la PAC par un modele bio-economique: une illustration sur la region de Toulouse. In: Reyniers, F. N., Benoit-Cattin, M. (Eds.), Linking-agricultural-models:-seminar-proceedings, IAM, Montpellier, France., pp. 71-75.

Flichman, G., Jacquet, F., 2003. Couplage des modèles agronomiques et économiques: intérêt pour l'analyse des politiques. Cahiers d'économie et sociologie rurales 67, 1-20.

Freund, R. J., 1956. The Introduction of Risk into a Programming Model. Econometrica 24 (3), 253-263.

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

Gallopin, G. C., 1997. Indicators and their use: information for decision-making. In: Billhartz, S. B., Matravers, R. (Eds.), Sustainability indicators: a report on the project on indicators of sustainable development. John Wiley and Sons, Chichester, pp. 13-27.

GAMS, 2008. General Algebraic Modeling System (GAMS). http://www.gams.com/.

Gangemi, A., Grimm, S., Mika, P., Oberle, D., Lamparter, S., Sabou, M., Staab, S., Vrandecic, D., 2008. Core Software Ontology - Core Ontology of Software Components - Core Ontology of Services. http://cos.ontoware.org/.

Garrett, P., 1984. The relevance of structural variables for farming systems research. Rural-Sociology 49 (4), 580-589.

GeoServer, 2008. GeoServer http://geoserver.org.

Gibbons, J. M., Sparkes, D. L., Wilson, P., Ramsden, S. J., 2005. Modelling optimal strategies for decreasing nitrate loss with variation in weather - a farm-level approach. Agricultural Systems 83 (2), 113-134.

Gibson, L., Kingwell, R., Doole, G., 2008. The role and value of eastern star clover in managing herbicide-resistant crop weeds: A whole-farm analysis. Agricultural Systems 98 (3), 199-207.

Gobin, A., Jones, R., Kirkby, M., Campling, P., Govers, G., Kosmas, C., Gentile, A. R., 2004. Indicators for pan-European assessment and monitoring of soil erosion by water. Environmental Science & Policy 7 (1), 25-38.

Goble, C., Stevens, R., 2008. State of the nation in data integration for bioinformatics. Journal of Biomedical Informatics 41 (5), 687-693.

GO-Consortium, 2007. The Gene Ontology Project. www.geneontology.org/.

Greeuw, S. C. H., Van Asselt, M. B. A., Grosskurth, J., Storms, C. A. M. H., Rijkens-Klomp, N., Rothman, D. S., Rotmans, J., 2000. Cloudy crystal balls- An assessment of recent European and global scenario studies and models. Environmental issues series No 17: Experts' corner report: Prospects and Scenarios No 4, European Environment Agency, Copenhagen, http://reports.eea.europa.eu/Environmental_issues_series_17/en/envissue17.pdf, pp. 112.

Gregersen, J. B., Gijsbers, P. J. A., Westen, S. J. P., 2007. OpenMI: Open modelling interface. Journal of Hydroinformatics 9 (3), 175–191.
Groot, J. C. J., Rossing, W. A. H., Jellema, A., Stobbelaar, D. J., Renting, H., Van Ittersum, M. K., 2007. Exploring multi-scale trade-offs between nature conservation, agricultural profits and landscape quality--A methodology to support discussions on land-use perspectives. Agriculture, Ecosystems & Environment 120 (1), 58-69.
Gruber, T. R., 1993. A translation approach to portable ontology specifications. Knowledge Acquisition 5, 199-220.
Gutierrez-Aleman, N., De Boer, A. J., Hart, R. D., 1986a. A bio-economic model of small-ruminant production in the semi-arid tropics of the Northeast Region of Brazil: Part 1--Model description and components. Agricultural Systems 19 (1), 55-66.
Gutierrez-Aleman, N., De Boer, A. J., Kehrberg, E. W., 1986b. A bio-economic model of small-ruminant production in the semi-arid tropics of the Northeast region of Brazil: Part 2--Linear programming applications and results. Agricultural Systems 19 (3), 159-187.
Hardaker, J., Huirne, R. B. M., Anderson, J., 1997. Coping with risk in agriculture. CAB International, Wallingford.
Harris, G., 2002. Integrated assessment and modelling: an essential way of doing science. Environmental Modelling & Software 17 (3), 201-207.
Hazell, P. B. R., Norton, R. D., 1986. Mathematical programming for economic analysis in agriculture. Macmillan Publishing Company, New York, pp. 400.
Hazeu, G., Elbersen, B., Andersen, E., Baruth, B., Van Diepen, C. A., Metzger, M. J., 2009. A biophysical typology for a spatially-explicit agri-environmental modeling framework. In: Brouwer, F., Van Ittersum, M. K. (Eds.), Environmental and agricultural modelling: integrated approaches for policy impact assessment. Springer Academic Publishing.
Heckelei, T., Britz, W., 2001. Concept and explorative application of an EU-wide regional agricultural sector model (CAPRI-Projekt). In: Heckelei, T., Witzke, H. P., Henrichsmeyer, W. (Eds.), 65th EAAE Seminar Vauk Verlag, Kiel, Germany, pp. 281–290.
Helming, K., Pérez-Soba, M., Tabbush, P. (Eds.), 2008. Sustainability Impact Assessment of Land Use Changes. Springer Publishing, pp. 508.
Hendriks, K., Stobbelaar, D. J., Mansvelt, J. D. v., 2000. The appearance of agriculture an assessment of the quality of landscape of both organic and conventional horticultural farms in West Friesland. Agriculture, Ecosystems and Environment 77 (1/2), 157-175.
Hengsdijk, H., Bouman, B. A. M., Nieuwenhuyse, A., Jansen, H. G. P., 1999. Quantification of land use systems using technical coefficient generators: a case study for the Northern Atlantic zone of Costa Rica. Agricultural Systems 61 (2), 109-121.
Hengsdijk, H., Van Ittersum, M. K., 2002. A goal-oriented approach to identify and engineer land use systems. Agricultural Systems 71 (3), 231-247.

Hengsdijk, H., Van Ittersum, M. K., 2003a. Dynamics in input and output coefficients for land use studies: a case for nitrogen in crop rotations. Nutrient Cycling in Agroecosystems 66, 209-220.
Hengsdijk, H., Van Ittersum, M. K., 2003b. Formalizing agro-ecological engineering for future-oriented land use studies. European Journal of Agronomy 19 (4), 549-562.
Herrero, M., Gonzalez-Estrada, E., Thornton, P. K., Quiros, C., Waithaka, M. M., Ruiz, R., Hoogenboom, G., 2007. IMPACT: Generic household-level databases and diagnostics tools for integrated crop-livestock systems analysis. Agricultural Systems 92 (1-3), 240-265.
Hertel, T. W. (Ed.) 1997. Global Trade Analysis: Modeling and Applications. Cambridge University Press, Cambridge, pp. 424.
Hillyer, Bolte, Evert, V., Lamaker, 2003. The ModCom modular simulation system. European Journal of Agronomy 18 (3-4), 333-343.
Hinkel, J., 2008. Transdisciplinary Knowledge Integration - Cases from Integrated Assessment and Vulnerability Assessment. Wageningen University, PhD thesis, Wageningen, pp. 198
Hinkel, J., 2009. The PIAM approach to modular integrated assessment modelling. Environmental Modelling & Software 24 (6), 739-748.
Holsapple, C. W., Joshi, K. D., 2002. A collaborative approach to ontology design. Communications of the ACM 45 (2), 42-47.
Howitt, R., 1995. Positive Mathematical Programming. American Journal of Agricultural Economics 77, 329-342.
INSPIRE, 2008. INSPIRE. http://inspire.jrc.it.
IPCC, 2000. Special Report on Emissions Scenarios. Intergovernmental Panel on Climate Change (IPCC), UNEP and WMO, http://www.grida.no/climate/ipcc/spmpdf/sres-e.pdf, pp. 27.
IPCC, 2007. Climate Change 2007: Synthesis Report. Intergovernmental Panel on Climate Change (IPCC), Geneva, http://www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4_syr.pdf, pp. 104.
ISO, 2008. ISO/tc211. http://www.isotc211.org
ISO, 2009. International Organization for Standardization. http://www.iso.org/iso/home.htm.
Jakeman, A. J., Letcher, R. A., Norton, J. P., 2006. Ten iterative steps in development and evaluation of environmental models. Environmental Modelling & Software 21 (5), 602-614.
Jakobsen, C. H., McLaughlin, W. J., 2004. Communication in Ecosystem Management: A Case Study of Cross-Disciplinary Integration in the Assessment Phase of the Interior Columbia Basin Ecosystem Management Project. Environmental Management 33 (5), 591-605.
Jannot, P., Cairol, D., 1994. Linear Programming as an Aid to Decision-making for Investments in Farm Equipment for Arable Farms. Journal of Agricultural Engineering Research 59 (3), 173-179.
Jansen, M. J. W., 1997. Validation of models at different scales. In: Ten Berge, H. F. M., Stein, A. (Eds.), Model-based decision support in agriculture :

proceedings of the INRA-KCW workshop on decision support systems, France (Laon), October 1997, pp. 93-99.
Jansson, T., Bakker, M., Boitier, B., Fougeyrollas, A., Helming, J., Van Meijl, H., Verkerk, P. J., 2008. Cross sector land use modelling framework. In: Helming, K., Pérez Soba, M., Tabbush, P. (Eds.), Sustainability Impact Assessment of Land Use Changes. Springer, Berlin, pp. 159-180.
Jarrige, R. (Ed.) 1988. Alimentation des Bovins, Ovins et Caprins. INRA Paris, France, pp. 476.
Jarrige, R. (Ed.) 1989 Ruminant Nutrition. Recommended Allowances and Feeding Tables. INRA, Paris, France. , pp. 389.
Jarrige, R., Demarquilly, C., Dulphy, J. P., Hoden, A., Robelin, J., Beranger, C., Geay, Y., Journet, M., Malterre, C., Micol, D., Petit, M., 1986. The INRA "Fill Unit" System for Predicting the Voluntary Intake of Forage-Based Diets in Ruminants: A Review. Journal of Animal Science 63, 1737-1758.
JBOSS, 2008. Hibernate. www.hibernate.org.
Jeffrey, S. R., Gibson, R. R., Faminow, M. D., 1992. Nearly optimal linear programming as a guide to agricultural planning. Agricultural Economics 8 (1), 1-19.
Jones, R. J. A., Hiederer, R., Rusco, E., Loveland, P. J., Montanarella, L., 2005. Estimating organic carbon in the soils of Europe for policy support. European Journal of Soil Science 56, 655-671.
JRC, 2008. Meteorological data Source JRC/AGRIFISH Data Base - EC - JRC. http://agrifish.jrc.it/marsstat/datadistribution/.
Judez, L., Chaya, C., Martinez, S., Gonzalez, A. A., 2001. Effects of the measures envisaged in "Agenda 2000" on arable crop producers and beef and veal producers: an application of Positive Mathematical Programming to representative farms of a Spanish region. Agricultural Systems 67 (2), 121-138.
Kachele, H., Dabbert, S., 2002. An economic approach for a better understanding of conflicts between farmers and nature conservationists--an application of the decision support system MODAM to the Lower Odra Valley National Park. Agricultural Systems 74 (2), 241-255.
Kanellopoulos, A., Berentsen, P. B. M., Heckelei, T., Ittersum, M. K. v., Lansink, A. G. J. M. O., 2009. Assessing the forecasting performance of a generic bio-economic farm model calibrated with two different PMP variants. Submitted In review,
Keating, B. A., Carberry, P. S., Hammer, G. L., Probert, M. E., Robertson, M. J., Holzworth, D., Huth, N. I., Hargreaves, J. N. G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J. P., Silburn, M., Wang, E., Brown, S., Bristow, K. L., Asseng, S., Chapman, S., McCown, R. L., Freebairn, D. M., Smith, C. J., 2003. An overview of APSIM, a model designed for farming systems simulation. European Journal of Agronomy 18 (3-4), 267.
Keeney, R. L., Raiffa, H., 1976. Decisions with Multiple Objectives: Preferences and Value Trade-offs. John Wiley & Sons, New York.

Kingwell, R. S., Ghadim, A. K. A., Robinson, S. D., Young, J. M., 1995. Introducing Awassi sheep to Australia: an application of farming system models. Agricultural Systems 47 (4), 451-471.

Kingwell, R. S., Pannell, D. J. (Eds.), 1987. MIDAS, A Bioeconomic Model of a Dryland Farm System. Pudoc, Wageningen, pp. 207.

Klein Haneveld, W. K., Stegeman, A. W., 2005. Crop succession requirements in agricultural production planning. European Journal of Operational Research 166 (2), 406-429.

Knapen, M. J. R., Athanasiadis, I. N., Jonsson, B., Huber, D., Wien, J. J. F., Rizzoli, A. E., Janssen, S., 2009. Use of OpenMI in SEAMLESS. In: Proceedings of Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Policy and Society (AgSap), Egmond aan Zee.

Knapen, R. J. R., Verweij, P. J. F. M., Wien, J. J. F., 2007. Applying Enterprise Application Architectures in Integrated Modelling. In: Proceedings of MODSIM, International Congress on Modelling and Simulation, University of Canterbury, Christchurch, New Zealand.

Knublauch, H., 2005. Protege OWL. http://protege.stanford.edu/.

Kok, K., Verburg, P. H., Veldkamp, T., 2007. Integrated Assessment of the land system: The future of land use. Land Use Policy 24 (3), 517-520.

Kopke, E., Young, J., Kingwell, R., 2008. The relative profitability and environmental impacts of different sheep systems in a Mediterranean environment. Agricultural Systems 96 (1-3), 85-94.

Kruseman, G., Bade, J., 1998. Agrarian policies for sustainable land use: bio-economic modelling to assess the effectiveness of policy instruments. Agricultural Systems 58 (3), 465-481.

Leeuwis, C., 1999. Science and integral design in agriculture and resource management. Identification of issues and organisation of the book. In: Leeuwis, C. (Ed.) Integral Design: innovation in agriculture and resource management. Mansholt Studies, Wageningen, pp. 1-11.

Leontief, W., 1986. Input-output economics. Oxford University Press, New York, pp. 436.

Letcher, R. A., Croke, B. F. W., Jakeman, A. J., 2007. Integrated assessment modelling for water resource allocation and management: A generalised conceptual framework. Environmental Modelling & Software 22 (5), 733-742.

Li, H., Louhichi, K., Janssen, S., Rizzoli, A. E., Athanasiadis, I. N., Meuter, E., David, H., 2007. Development and Application of a Component-based Generic Farm System Simulator Implementing a Semantically Enriched Integrated Modelling Framework. In: Proceedings of the International Symposium on Environmental Software Systems 2007, Prague.

Li, M., Yost, R. S., 2000. Management-oriented modeling: optimizing nitrogen management with artificial intelligence. Agricultural Systems 65 (1), 1-27.

Liu, Y., Gupta, H., Springer, E., Wagener, T., 2008. Linking science with environmental decision making: Experiences from an integrated modeling approach to supporting sustainable water resources management. Environmental Modelling & Software 23 (7), 846-858.

Louhichi, K., Belhouchette, H., Wery, J., Therond, O., Flichman, G., Casellas, E., Traoré, B., Rapidel, B., Barbier, B., Lahmar, R., 2008. Application of FSSIM in two Test Case regions to assess agro-environmental policies at farm and regional level. PD6.3.2.2, SEAMLESS integrated project, EU 6th Framework programme, contract no. 010036-2, www.SEAMLESS-IP.org, Wageningen, pp. 62.

Louhichi, K., Flichman, G., Zekri, S., 1999. A bio-economic model for analysing the impact of soil and water conservation policies applied to a Tunisian farm.OT: Un modele bio-economique pour analyser l'impact de la politique de conservation des eaux et du sol. Le cas d'une exploitation agricole tunisienne. Economie Rurale 252, 55-64.

Louhichi, K., Hengsdijk, H., Janssen, S., Bigot, G., Perret, E., 2009a. EU dairy farming in the face of change: An exploration using a bio-economic farm model. In: Proceedings of Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Policy and Society (AgSap), Egmond aan Zee.

Louhichi, K., Janssen, S., Li, H., Borkowski, N., Kanellopoulos, A., Van Ittersum, M. K., Flichman, G., Hengsdijk, H., Zander, P., Blanco, M., Stokstad, G., Athanasiadis, I., Rizzoli, A. E., Huber, D., 2009. A generic farming system simulator. In: Brouwer, F., Van Ittersum, M. K. (Eds.), Environmental and agricultural modelling: integrated approaches for policy impact assessment. Springer Academic Publishing.

Lu, C. H., Van Ittersum, M. K., Rabbinge, R., 2004. A scenario exploration of strategic land use options for the Loess Plateau in northern China. Agricultural Systems 79 (2), 145-170.

Majewski, E., Was, A., Belhouchette, H., Louhichi, K., Mouratiadou, I., 2009. Impact Assessment of Policy Changes On the Arable Sector Using the FSSIM Model : the case study of the Zachodniopomorskie NUTS2 region In: Proceedings of Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Policy and Society (AgSap), Egmond aan Zee.

Makowski, D., Hendrix, E. M. T., Van Ittersum, M. K., Rossing, W. A. H., 2001. Generation and presentation of nearly optimal solutions for mixed-integer linear programming, applied to a case in farming system design. European Journal of Operational Research 132 (2), 425-438.

McCown, R. L., 2001. Learning to bridge the gap between science-based decision support and the practice of farming: Evolution in paradigms of model-based research and intervention from design to dialogue. Australian Journal of Agricultural Research 52 (5), 549-572.

McEvoy, J., 1972. Multi- and interdisciplinary research—Problems of initiation, control, integration and reward. Policy Sciences 3 (2), 201-208.

McGuinness, D., Van Harmelen, F., 2004. OWL Web Ontology Language Overview. www.w3.org/TR/owl-features/.

Metsker, S. J., 2002. Design Patterns JavaTM Workbook. Adisson-Wesley, Boston, pp. 475.

Metzger, M. J., Bunce, R. G. H., Jongman, R. H. G., Mucher, C. A., Watkins, J. W., 2005. A climatic stratification of the environment of Europe. Global Ecology and Biogeography 14 (6), 549-563.

Metzger, N., Zare, R. N., 1999. SCIENCE POLICY:Interdisciplinary Research: From Belief to Reality. Science 283 (5402), 642-643.

Meuter, E., Janssen, S., Van Ittersum, M. K., 2009. A Graphical User Interface for a generic bio-economic farm model (FSSIM). In: Proceedings of Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Policy and Society (AgSap), Egmond aan Zee.

Meyer-Aurich, A., 2005. Economic and environmental analysis of sustainable farming practices - a Bavarian case study. Agricultural Systems 86 (2), 190-206.

Meyer-Aurich, A., Zander, P., Werner, A., Roth, R., 1998. Developing agricultural land use strategies appropriate to nature conservation goals and environmental protection. Landscape and Urban Planning 41 (2), 119-127.

Moore, A. D., Holzworth, D. P., Herrmann, N. I., Huth, N. I., Robertson, M. J., 2007. The Common Modelling Protocol: A hierarchical framework for simulation of agricultural and environmental systems. Agricultural Systems 95 (1-3), 37-48.

Moore, R. V., Tindall, C. I., 2005. An overview of the open modelling interface and environment (the OpenMI). Environmental Science & Policy 8 (3), 279-286.

Morrison, D. A., Kingwell, R. S., Pannell, D. J., Ewing, M. A., 1986. A mathematical programming model of a crop-livestock farm system. Agricultural Systems 20 (4), 243-268.

Mouratiadou, I., Topp, C., Russell, G., 2009. Investigating the economic and water quality effects of the 2003 CAP reform on arable cropping systems: a Scottish case study. In: Proceedings of Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Policy and Society (AgSap), Egmond aan Zee.

MSC, 2009. Marine Stewardship Council. http://www.msc.org/.

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

Niemeijer, D., 2002. Developing indicators for environmental policy: data-driven and theory-driven approaches examined by example. Environmental Science & Policy 5 (2), 91-103.

NMP, 2006. Integrated modelling of global environmental change. An overview of IMAGE 2.4. Report no. 500110002, Netherlands Environmental Assessment Agency, http://www.mnp.nl/en/publications/2006/Integratedmodellingofglobalenvironmentalchange.AnoverviewofIMAGE2.4.html, pp. 228.

Norgaard, R. B., 1992. Coordinating disciplinary and organizational ways of knowing. Agriculture, Ecosystems & Environment 42 (1-2), 205-216.

Novak, J. D., Cañas, A. J., 2006. The Theory Underlying Concept Maps and How to Construct Them Technical Report IHMC CmapTools 2006-01, Florida Institute for Human and Machine Cognition http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf.
Oglethorpe, D. R., 1995. Sensitivity of Farm Planning under Risk-Averse Behaviour: a Note on the Environmental Implications. Journal of Agricultural Economics 46 (2), 227-232.
Oglethorpe, D. R., Sanderson, R. A., 1999. An ecological-economic model for agri-environmental policy analysis. Ecological Economics 28 (2), 245-266.
Onate, J. J., Atance, I., Bardaji, I., Llusia, D., 2007. Modelling the effects of alternative CAP policies for the Spanish high-nature value cereal-steppe farming systems. Agricultural Systems 94 (2), 247-260.
Oomen, R., Hengsdijk, H., Therond, O., Bergez, J. E., Russell, G., Casellas, E., Janssen, S., Van Ittersum, M. K., 2009. Development of generic management rules for crop growth simulation models In: Proceedings of Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Policy and Society (AgSap), Egmond aan Zee.
OpenMI, 2009. The OpenMI Association. www.openmi.org.
Oxley, T., ApSimon, H. M., 2007. Space, time and nesting Integrated Assessment Models. Environmental Modelling & Software 22 (12), 1732-1749.
Pacini, G. C., 2003. An environmental-economic framework to support multi-objective policy-making : a farming systems approach implemented for Tuscany. Wageningen University, PhD thesis, Wageningen, pp. 173
Pala, M., Stockle, C. O., Harris, H. C., 1996. Simulation of durum wheat (Triticum turgidum ssp. durum) growth under different water and nitrogen regimes in a mediterranean environment using CropSyst. Agricultural Systems 51 (2), 147-163.
Pandey, S., Hardaker, J. B., 1995. The role of modelling in the quest for sustainable farming systems. Agricultural Systems 47 (4), 439-450.
Pannell, D. J., 1997. Sensitivity analysis of normative economic models: theoretical framework and practical strategies. Agricultural Economics 16 (2), 139-152.
Pannell, D. J., Malcolm, B., Kingwell, R. S., 2000. Are we risking too much? Perspectives on risk in farm modelling. Agricultural Economics 23 (1), 69-78.
Parker, P., Letcher, R., Jakeman, A., Beck, M. B., Harris, G., Argent, R. M., Hare, M., Pahl-Wostl, C., Voinov, A., Janssen, M., Sullivan, P., Scoccimarro, M., Friend, A., Sonnenshein, M., Barker, D., Matejicek, L., Odulaja, D., Deadman, P., Lim, K., Larocque, G., Tarikhi, P., Fletcher, C., Put, A., Maxwell, T., Charles, A., Breeze, H., Nakatani, N., Mudgal, S., Naito, W., Osidele, O., Eriksson, I., Kautsky, U., Kautsky, E., Naeslund, B., Kumblad, L., Park, R., Maltagliati, S., Girardin, P., Rizzoli, A., Mauriello, D., Hoch, R., Pelletier, D., Reilly, J., Olafsdottir, R., Bin, S., 2002. Progress in integrated assessment and modelling. Environmental Modelling & Software 17 (3), 209-217.

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

Payraudeau, S., Van der Werf, H. M. G., 2005. Environmental impact assessment for a farming region: a review of methods. Agriculture, Ecosystems and Environment 107 (1), 1-19.

Pérez Domínguez, I., Bezlepkina, I., Heckelei, T., Romstad, E., Oude Lansink, A., Kanellopoulos, A., 2009. Linking farm and market models by means of response functions. Environmental Science & Policy Accepted, in press.

Pérez Domínguez, I., Gay, S. H., M'Barek, R., 2008. Agricultural Modelling for European Policy Makers: A User's Perspective. In: the EAAE Congress Ghent, Belgium.

Pervanchon, F., Bockstaller, C., Girardin, P., 2002. Assessment of energy use in arable farming systems by means of an agro-ecological indicator: the energy indicator. Agricultural Systems 72 (2), 149-172.

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

Pfister, F., Bader, H. P., Scheidegger, R., Baccini, P., 2005. Dynamic modelling of resource management for farming systems. Agricultural Systems 86 (1), 1-28.

Ponsioen, T. C., Hengsdijk, H., Wolf, J., Van Ittersum, M. K., Rotter, R. P., Son, T. T., Laborte, A. G., 2006. TechnoGIN, a tool for exploring and evaluating resource use efficiency of cropping systems in East and Southeast Asia. Agricultural Systems 87 (1), 80.

PostGIS, 2008. PostGIS. http://postgis.refractions.net/.

PostgreSQL, 2008. PostgreSQL. www.postgreSQL.org.

Rahman, J. M., Seaton, S. P., Perraud, J.-M., Hotham, H., Verrelli, D. I., Coleman, J. R., 2003. It's TIME for a New Environmental Modelling Framework. In: Proceedings of MODSIM 2003 International Congress on Modelling and Simulation, Townsville.

Ramsden, S., Gibbons, J., Wilson, P., 1999. Impacts of changing relative prices on farm level dairy production in the UK. Agricultural Systems 62 (3), 201-215.

Refsgaard, J. C., Henriksen, H. J., 2004. Modelling guidelines--terminology and guiding principles. Advances in Water Resources 27 (1), 71-82.

Refsgaard, J. C., Nilsson, B., Brown, J., Klauer, B., Moore, R., Bech, T., Vurro, M., Blind, M., Castilla, G., Tsanis, I., Biza, P., 2005. Harmonised techniques and representative river basin data for assessment and use of uncertainty information in integrated water management (HarmoniRiB). Environmental Science & Policy 8 (3), 267-277.

Rehman, T., Romero, C., 1993. The application of the MCDM paradigm to the management of agricultural systems: Some basic considerations. Agricultural Systems 41 (3), 239-255.

Reinds, G. J., Van Lanen, H. A. J., 1992. Crop production potential of rural areas within the European Communities. II. A physical land evaluation procedure

for annual crops and grass. Publication of the Scientific Council for Government Policy, Scientific Council for Government Policy, The Hague.
Reus, J., Leendertse, P., Bockstaller, C., Fomsgaard, I., Gutsche, V., Lewis, K., Nilsson, C., Pussemier, L., Trevisan, M., Van der Werf, H., Alfarroba, F., Blumel, S., Isart, J., McGrath, D., Seppala, T., 2002. Comparison and evaluation of eight pesticide environmental risk indicators developed in Europe and recommendations for future use. Agriculture, Ecosystems & Environment 90 (2), 177-187.
Riesgo, L., Gomez-Limon, J. A., 2006. Multi-criteria policy scenario analysis for public regulation of irrigated agriculture. Agricultural Systems 91 (1-2), 1-28.
Rizzoli, A. E., Donatelli, M., Athanasiadis, I. N., Villa, F., Huber, D., 2008. Semantic links in integrated modelling frameworks. Mathematics and Computers in Simulation 78 (2-3), 412-423.
Rossing, W. A. H., Meynard, J. M., Van Ittersum, M. K., 1997. Model-based explorations to support development of sustainable farming systems: case studies from France and the Netherlands. European Journal of Agronomy 7 (1-3), 271-283.
Rotmans, J., 1998. Methods for IA: The challenges and opportunities ahead. Environmental Modeling and Assessment 3 (3), 155.
Rotmans, J., Asselt, M., 1996. Integrated assessment: A growing child on its way to maturity. Climatic Change 34 (3), 327.
Rounsevell, M. D. A., Ewert, F., Reginster, I., Leemans, R., Carter, T. R., 2005. Future scenarios of European agricultural land use: II. Projecting changes in cropland and grassland. Agriculture, Ecosystems & Environment 107 (2-3), 117-135.
Ruben, R., Moll, H., Kuyvenhoven, A., 1998. Integrating agricultural research and policy analysis: analytical framework and policy applications for bio-economic modelling. Agricultural Systems 58 (3), 331-349.
Ruben, R., Van Ruijven, A., 2001. Technical coefficients for bio-economic farm household models: a meta-modelling approach with applications for Southern Mali. Ecological Economics 36 (3), 427-441.
Russel, G., 1990. Barley knowledge base. An agricultural information systems for the European Community. Commission of the European Community, Brussels.
Schilizzi, S. G. M., Boulier, F., 1997. `Why do farmers do it?' Validating whole-farm models. Agricultural Systems 54 (4), 477-499.
Schoemaker, P. J. H., 1993. Multiple Scenario Development: Its Conceptual and Behavioral Foundation. Strategic Management Journal 14 (3), 193.
Scholten, H., 2008. Better Modelling Practice: an ontological perspective on multidisciplinary, model based problem solving. Wageningen University, PhD thesis, Wageningen, pp. 313
Scholten, H., Kassahun, A., Refsgaard, J. C., Kargas, T., Gavardinas, C., Beulens, A. J. M., 2007. A methodology to support multidisciplinary model-based water management. Environmental Modelling & Software 22 (5), 743-759.

Schuler, J., Kachele, H., 2003. Modelling on-farm costs of soil conservation policies with MODAM. Environmental Science & Policy 6 (1), 51-55.
Sedgewick, R., 1977. Permutation Generation Methods. Computing Surveys 9 (2), 137-164.
Semaan, J., Flichman, G., Scardigno, A., Steduto, P., 2007. Analysis of nitrate pollution control policies in the irrigated agriculture of Apulia Region (Southern Italy): A bio-economic modelling approach. Agricultural Systems 94 (2), 357-367.
SenterNovem, 2009. Besluit Subsidies Investeringen Kennisinfrastructuur (Bsik) (ICES/KIS-3).
Smith, N. J., Sage, A. P., 1973. An introduction to hierarchical systems theory. Computers & Electrical Engineering 1 (1), 55-71.
Stöckle, C. O., Donatelli, M., Nelson, R., 2003. CropSyst, a cropping systems simulation model. European Journal of Agronomy 18 (3-4), 289-307.
Stöckle, C. O., Martin, S. A., Campbell, G. S., 1994. CropSyst, a cropping systems simulation model: Water/nitrogen budgets and crop yield. Agricultural Systems 46 (3), 335-359.
Strassert, G., Prato, T., 2002. Selecting farming systems using a new multiple criteria decision model: the balancing and ranking method. Ecological economics 40 (2), 269-277.
Szyperski, C., Gruntz, D., Murer, S., 2002. Component Software - Beyond Object-Oriented Programming. ACM Press, Ñew York, NY.
Taylor, M. C., 2009, End the University as We Know It. New York Times, New York, 26 April 2009
Ten Berge, H. F. M., Van Ittersum, M. K., Rossing, W. A. H., Van de Ven, G. W. J., Schans, J., 2000. Farming options for The Netherlands explored by multi-objective modelling. European Journal of Agronomy 13 (2-3), 263-277.
Thérond, O., Belhouchette, H., Janssen, S., Louhichi, K., Ewert, F., Bergez, J.-E., Wery, J., Heckelei, T., Olsson, J. A., Leenhardt, D., Ittersum, M. K. V., 2009. Methodology to translate policy assessment problems into scenarios: the example of the SEAMLESS Integrated Framework. Environmental Science & Policy In press,
Thomann, R. V., Muller, J. A., 1982. Verification of Water Quality Models. Journal of Environmental Engineering 108 (EE 5), 923-940.
Thompson, A. M. M., 1982. A farm-level model to evaluate the impacts of current energy policy options. Lincoln College, Canterbury, pp. 159.
Thorne, P. J., Hengsdijk, H., Janssen, S., Louhichi, K., Van Keulen, H., Thornton, P., 2008. Modelling Livestock Component in FSSIM. PD 3.3.4, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org, pp. 65.
Thornton, P. K., Herrero, M., 2001. Integrated crop-livestock simulation models for scenario analysis and impact assessment. Agricultural Systems 70 (2-3), 581-602.
Ticehurst, J. L., Newham, L. T. H., Rissik, D., Letcher, R. A., Jakeman, A. J., 2007. A Bayesian network approach for assessing the sustainability of coastal

lakes in New South Wales, Australia. Environmental Modelling & Software 22 (8), 1129-1139.

Topp, C. F. E., Mitchell, M., 2003. Forecasting the environmental and socio-economic consequences of changes in the Common Agricultural Policy. Agricultural Systems 76 (1), 227-252.

Torkamani, J., 2005. Using a whole-farm modelling approach to assess prospective technologies under uncertainty. Agricultural Systems 85 (2), 138-154.

Tress, G., Tress, B., Fry, G., 2006. Publishing integrative landscape research: analysis of editorial policies of peer-reviewed journals. Environmental Science & Policy 9 (5), 466-475.

Tress, G., Tress, B., Fry, G., 2007. Analysis of the barriers to integration in landscape research projects. Land Use Policy 24 (2), 374-385.

Turner, R. K., Ledoux, L., Cave, R., 2001. The use of scenarios in integrated environmental assessment of coastal-catchment zones. Report for the EUROCAT project. EC-contractnr. EVK1-CT-2000-00044, School of Environmental Sciences, University of East Anglia, Norwich, http://www.cs.iia.cnr.it/EUROCAT/publications/EUROCAT%20WD01.pdf, pp. 25.

Turner, R. K., Ledoux, L., Cave, R., 2001. The use of scenarios in integrated environmental assessment of coastal-catchment zones. Report for the EUROCAT project. EC-contractnr. EVK1-CT-2000-00044, School of Environmental Sciences, University of East Anglia, Norwich, http://www.cs.iia.cnr.it/EUROCAT/publications/EUROCAT%20WD01.pdf, pp. 25.

UN, 2005. Investing in Development: a practical plan to achieve the Millennium Development Goals. UN Millennium Project, United Nations, New York, http://www.unmillenniumproject.org/documents/overviewEngLowRes.pdf.

Uthes, S., Sattler, C., Reinhardt, F.-J., Piorr, A., Zander, P., Happe, K., Damgaard, M., Osuch, A., 2008. Ecological effects of payment decoupling in a case study region in Germany. Journal of Farm Management 13 (3), 219-230.

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

Van de Ven, G. W. J., de Ridder, N., Van Keulen, H., Van Ittersum, M. K., 2003. Concepts in production ecology for analysis and design of animal and plant-animal production systems. Agricultural Systems 76 (2), 507.

Van der Burgt, G., Oomen, G., Habets, A., Rossing, W., 2006. The NDICEA model, a tool to improve nitrogen use efficiency in cropping systems. Nutrient Cycling in Agroecosystems 74 (3), 275-294.

Van der Goot, E., 1997. Technical description of interpolation and processing of meteorological data in CGMS. Joint Research Centre of the European Commission, Ispra, Italy, pp. 23.

Van der Maden, E., 2007. Preparations for Implementation of FSSIM -- Implementation of the bio-economic farm model FSSIM for assessing arable conventional farming in the Flevopolder region, The Netherlands--.

Wageningen University, Plant Production Systems group, MSc thesis, Wageningen, pp. 112
Van der Ploeg, J. D., 1994. Styles of farming: an introductory note on concepts and methodology. In: Van der Ploeg, J. D., Long, A. (Eds.), Born from within: practices and perspectives of endogenous development. Van Gorcum, Assen, pp. 7-30.
Van Ittersum, M. K., Donatelli, M., 2003. Special Issue of the European Journal of Agronomy: Modelling Cropping Systems. European Journal of Agronomy 18 (3-4), 187-394.
Van Ittersum, M. K., Ewert, F., Heckelei, T., Wery, J., Alkan Olsson, J., Andersen, E., Bezlepkina, I., Brouwer, F., Donatelli, M., Flichman, G., Olsson, L., Rizzoli, A., Van der Wal, T., Wien, J.-E., Wolf, J., 2008. Integrated assessment of agricultural systems- a component based framework for the European Union (SEAMLESS). Agricultural Systems 96, 150-165.
Van Ittersum, M. K., Rabbinge, R., 1997. Concepts in production ecology for analysis and quantification of agricultural input-output combinations. Field Crops Research 52 (3), 197-208.
Van Ittersum, M. K., Rabbinge, R., Van Latesteijn, H. C., 1998. Exploratory land use studies and their role in strategic policy making. Agricultural Systems 58 (3), 309-330.
Van Latesteijn, H. C., Rabbinge, R., 1992. Grond voor keuzen ook grond voor discussie. Werkgroep Landbouwpolitiek, Landbouw Universiteit Wageningen, Wageningen.
Van Meijl, H., Van Rheenen, T., Tabeau, A., Eickhout, B., 2006. The impact of different policy environments on agricultural land use in Europe. Agriculture, Ecosystems & Environment 114 (1), 21-38.
Van Notten, P. W. F., Rotmans, J., Van Asselt, M. B. A., Rothman, D. S., 2003. An updated scenario typology. Futures 35 (5), 423.
Van Tongeren, F., Van Meijl, H., Surry, Y., 2001. Global models applied to agricultural and trade policies: a review and assessment. Agricultural Economics 26 (2), 149-172.
Vatn, A., Bakken, L. R., Bleken, M. A., Baadshaug, O. H., Fykse, H., Haugen, L. E., Lundekvam, H., Morken, J., Romstad, E., Rorstad , P. K., Skjelvag, A. O., Tr, S., 2003.
Vatn, A., Bakken, L. R., Lundeby, H., Romstad, E., Rorstad , P. K., Vold, A., Botterweg, P., 1997. Regulating nonpoint-source pollution from agriculture: An integrated modelling analysis. European Review of Agricultural Economics 24 (2), 207-230.
Verburg, P. H., Rounsevell, M. D. A., Veldkamp, A., 2006. Scenario-based studies of future land use in Europe. Agriculture, Ecosystems & Environment 114 (1), 1-6.
Verburg, P. H., Schot, P. P., Dijst, M. J., Veldkamp, A., 2004. Land use change modelling: current practice and research priorities. GeoJournal 61 (4), 309-324.

Verburg, P., Eickhout, B., Van Meijl, H., 2008. A multi-scale, multi-model approach for analyzing the future dynamics of European land use. The Annals of Regional Science 42 (1), 57-77.
Verburg, P., Lesschen, J.-P., 2006. Agriculture, Climate and Future Land Use Patterns: Potential for a Similation-based exploration. In: Brouwer, F., McCarl, B. A. (Eds.), Agriculture and climate beyond 2015; A new perspective on future land use patterns. Springer, pp. 5-32.
Vereijken, P., 1997. A methodical way of prototyping integrated and ecological arable farming systems (I/EAFS) in interaction with pilot farms. European Journal of Agronomy 7, 235-250.
Veysset, P., Bebin, D., Lherm, M., 2005. Adaptation to Agenda 2000 (CAP reform) and optimisation of the farming system of French suckler cattle farms in the Charolais area: a model-based study. Agricultural Systems 83 (2), 179-202.
Villa, F., Athanasiadis, I. N., Rizzoli, A. E., 2009. Modelling with knowledge: A review of emerging semantic approaches to environmental modelling. Environmental Modelling & Software 24 (5), 577-587.
Villa, F., Ceroni, M., Krivov, S., 2007. Intelligent databases assist transparent and sound economic valuation of ecosystem services. Environmental Management 39, 887-899.
Wallace, M. T., Moss, J. E., 2002. Farmer Decision-Making with Conflicting Goals: A Recursive Strategic Programming Analysis. Journal of Agricultural Economics 53 (1), 82-100.
Wang, Z.-M., Zhang, B., Li, X.-Y., Song, K.-S., Liu, D.-W., Zhang, S.-Q., 2006. Using CropSyst to Simulate Spring Wheat Growth in Black Soil Zone of Northeast China. Pedosphere 16 (3), 354-361.
Warren, R., de la Nava Santos, S., Arnell, N. W., Bane, M., Barker, T., Barton, C., Ford, R., Füssel, H. M., Hankin, R. K. S., Hinkel, J., Klein, R., Linstead, C., Kohler, J., Mitchell, T. D., Osborn, T. J., Pan, H., Raper, S. C. B., Riley, G., Schellnhüber, H. J., Winne, S., Anderson, D., 2008. Development and illustrative outputs of the Community Integrated Assessment System (CIAS), a multi-institutional modular integrated assessment approach for modelling climate change. Environmental Modelling & Software 23 (5), 592-610.
Wear, D. N., 1999. Challenges to interdisciplinary discourse. Ecosystems 2, 299–301.
Weersink, A., Jeffrey, S., Pannell, D., 2004. Farm-Level Modeling for Bigger Issues. Review of Agricultural Economics 24 (1), 123-140.
Weyant, J., Davidson, H., Dowlatabadi, H., Edmonds, J., Grubb, M., Richels, R., Rotmans, J., Shukla, P., Tol, R. S. J., Cline, W., Fankhauser, S., 1996. Integrated Assessment of climate change: An overview and comparison of approaches and results. In: Bruce, J. P., Lee, H., Haites, E. F. (Eds.), Climate Change 1995 - Economic and Social Dimensions. Cambridge University Press, Cambridge, pp. 367-396.

White, D. S., Labarta, R. A., Leguia, E. J., 2005. Technology adoption by resource-poor farmers: considering the implications of peak-season labor costs. Agricultural Systems 85 (1), 183-201.
Wien, J. J. F., Rizzoli, A. E., Knapen, M. J. R., Athanasiadis, I. N., Janssen, S. J. C., Ruinelli, L., Villa, F., Svensson, M., Wallman, P., Jonsson, B., Van Ittersum, M. K., 2009. A web-based software system for model integration in agro-environmental impact assessments. In: Brouwer, F., Van Ittersum, M. K. (Eds.), Environmental and agricultural modelling: integrated approaches for policy impact assessment. Springer Academic Publishing, pp. 24.
Wolf, J., Boogaard, H. L., Van Diepen, C. A., 2004. Evaluating crop data in CGMS system. Crop data for winter wheat and spring barley. MARSOP-2 Project report no. 1, Alterra, Wageningen.
Wossink, G. A. A., de Koeijer, T. J., Renkema, J. A., 1992. Environmental-economic policy assessment: A farm economic approach. Agricultural Systems 39 (4), 421-438.
Wossink, G. A. A., Oude Lansink, A. G. J. M., Struik, P. C., 2001. Non-separability and heterogeneity in integrated agronomic-economic analysis of nonpoint-source pollution. Ecological Economics 38 (3), 345-357.
Wossink, G. A. A., Renkema, J. A., 1994. Analysis of future change in Dutch arable farming: A farm economics approach. European Review of Agricultural Economics 21, 95-112.
WU, 2009. Interdisciplinary Research and Education Fund. http://www.wi.wur.nl/UK/services/International+research+at+Wageningen+UR/INREF.
Yu, P. L., 1973. A class of solutions for group decision problems. Management Science 19, 936-946.
Zander, P., Borkowski, N., Hecker, M., Uthes, S., Stokstad, G., Rørstad, P. K., Bellocchi, G., 2009. Conceptual approach to identify and assess current activities. PD3.3.9, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org, Wageningen, pp. 111.
Zander, P., Kächele, H., 1999. Modelling multiple objectives of land use for sustainable development. Agricultural Systems 59, 311-325.
Zeleny, M., 1973. Compromise programming. In: Cochrane, J. L., Zeleny, M. (Eds.), Multiple Criteria Decision-making. University of South Carolina Press, Columbia, pp. 262-301.

Summary

Interdisciplinary research is required if science wants to assess policy measures and technological innovations targeting complex multi-scale and multi-dimensional problems, e.g. climate change, poverty, hunger, biodiversity loss. Ex-ante assessment through science based methods can provide insight into the impacts of potential policy measures or innovations to manage these complex problems. Integrated Assessment and Modelling (IAM) is a method that supports ex-ante assessment through modelling and modelling tools. One type of IAM links models focusing on particular processes on a particular scale into model chains covering multiple scales across multiple disciplines. Such model chains simulate the impact of policy measures and technological innovations on sustainability indicators on multiple scales and across disciplines.

To achieve an operational model chain for IAM, interdisciplinary integration is required of models, data sources, indicators and scenarios. Methodological, semantic and technical integration are the focus of this thesis. These three modes of integration are developed for an IAM project codenamed SEAMLESS with ca. 30 research groups and over 100 researchers involved. SEAMLESS aimed to (1) achieve a model linking for a group of models from the agricultural domain, (2) develop a framework for integrated assessment that can be executed on a computer, (3) develop procedures for integrated assessment of agricultural systems and (4) deliver applications of the linked models in the modelling framework using the developed procedure.

In this thesis, methodological, semantic and technical integration focuses on two case studies. The first case study is on integration within bio-economic farm models covering two hierarchical systems levels (e.g. field to farm) involving a relatively small team of approximately ten agronomic and farm economic scientists. The second case is modelling European agricultural systems. In this case, the integration covers five hierarchical systems levels (e.g. field, farm, region, country and continental) and different types of models were linked by a large team of about

hundred scientists from agronomy, economics, landscape ecology, information technology and environmental sciences.

A bio-economic farm model (BEFM) links decisions on management of a farm's resources to current and alternative production possibilities describing input-output relationships and associated externalities. The basis for a methodological integration for BEFMs is a review of descriptions, theoretical considerations and applications as found in the scientific literature. Chapter 2 introduces a classification of different types of BEFMs, discusses strengths and weaknesses and outlines a research agenda. According to this research agenda, future research must incorporate an adequate model description, availability of source code or program, evaluation of model results and development of an easily usable generic BEFM. Also, incorporation of alternative agricultural activities, strategic decision making and social environment are lacking in BEFMS.

The methodological integration of different methods to define alternative agricultural activities for BEFMs is presented in Chapter 3. Alternative activities are activities that are not currently practiced on farms, but that might be suitable alternatives in the future. The integrated method to define such activities represents an approach based on production ecological principles, that can be applied throughout Europe. It attempts to be inclusive to avoid that suitable alternatives are excluded a priori. The integrated method is based on generating alternatives as permutations and filtering the possible permutations to feasible alternatives using heuristic filters.

Both methodological and technical integration are described for the development of the Farm System SIMulator (FSSIM) as a generic BEFM (Chapter 4). FSSIM is a bio-economic farm model, that was developed to simulate farm responses of European agricultural systems (e.g. livestock and annual and perennial cropping). It consists of two parts, FSSIM-Mathematical Programming and FSSIM-Agricultural Management. FSSIM-Mathematical Programming is a positive static risk programming model supplemented with different calibration methods. FSSIM-Agricultural Management describes current activities, generates alternative activities and quantifies the activities through all the required technical coefficients. A generic BEFM must be applicable to a range of agri-environmental zones, for a range of farm types, for both assessments of technological innovations and policy questions, for applications that require different level of detail in input or output data and for linking to other models at different scales. All applications of FSSIM are evaluated according to these criteria and using this evaluation it is argued that FSSIM fulfills these criteria.

The case study on modelling European agricultural systems requires integration of multiple models, data sources, indicators and scenarios and builds on the integration achieved within the BEFM. An important prerequisite for model integration is the semantic and technical integration of data sources into an integrated database with a shared database scheme (Chapter 5). The integrated database stores the input and output data for the models, the assessment projects and scenarios as defined through a modelling framework, the data derived from the original data-sources and data on indicators required to quantify the impacts of the assessment. The original data sources part of the integrated database include sources on economic farm performance, soil, climate, policy, agricultural management and trade. The development of the integrated database required the shared definition of concepts across data-sources into an ontology.

Models may be based on different conceptualizations of space and time, are of a type (e.g. dynamic simulation, optimization, estimation) and capture a selection of biophysical, economic or social processes. Also, these models often refer to different names and definitions for the same concepts. The methodological, semantic and technical integration of the models into a model chain is required to enable multi-scale and multi-disciplinary assessment of policy and technology changes (Chapter 6). A cropping system model and a bio-economic farm model are integrated by working with a shared definition of agricultural activities, and providing different parameter values of the activity-concept, while the parameter values set by one of the models are the inputs to the other model. A bio-economic farm model, an econometric regional estimation model and a partial equilibrium market model are integrated through price elasticities, which are estimated from the farm responses of the bio-economic farm model by the econometric model and which serve as input to the partial equilibrium market model. The model linkages are made explicit, reproducible and transparent by explicitly describing the concepts of relevance and the calculation steps occurring to transform one concept into another concept.

Scenarios are an important concept in IAM to structure assessments with IAM tools. The concept 'scenario' is used with many different meanings in different disciplines and can thus create confusion in an IAM project or alternatively, can be used as a container term for unknown issues in IAM project. To achieve semantic integration on scenarios, scenario was clearly defined into a shared conceptualization between many researchers (Chapter 7). Scenario is defined as an experiment in an assessment project, that exists of a policy option, context and outlook. This scenario-definition was established in an iterative collaborative process involving many researchers.

Methodological, semantic and technical integration require each a different approach and are subject to different considerations (Chapter 8). For methodological integration explicit consideration of relevant concepts and calculations to transform concepts into one another lead to a transparent and explicit model linking of a cropping system model, a bio-economic farm model, an econometric estimation model and a partial equilibrium market model. Still a lack of methods to integrate model types (e.g. dynamic simulation, optimization, estimation), process descriptions, space and time across land use models is identified. In semantic integration the use of ontologies and collaborative approaches to develop ontologies provided a powerful combination, that has been extensively used for semantic integration in SEAMLESS. The collaborative approaches had to be geared to the integration issue, e.g. data sources, models and scenarios, and to domain scientists involved. Ontologies built for semantic integration can be combined with advanced modelling frameworks like Open-MI to dynamically execute model chains in a flexible and transparent way. Although powerful IT-solutions are available for technical integration, applying them for model integration is not a trivial task.

Communication is crucial in any integration approach and an explicit strategy to communication is required for integration. Such a communication strategy in integration is often absent in either methodological, semantic or technical integration. Although the case studies demonstrate that methodological, semantic and technical integration is possible, it is concluded that they are only a beginning and demonstrate what is possible. To advance integration and interdisciplinary research in large scale projects, rigorous, tested and documented approaches to any type of integration need to be developed.

Samenvatting

Interdisciplinair onderzoek is vereist voor een effectieve bijdrage van onderzoek aan de beoordeling van beleidsmaatregelen en technologische innovaties. Deze beleidsmaatregelen en technologische innovaties zouden bij kunnen dragen aan het oplossen van complexe problemen op meerdere schaalniveaus en disciplines, zoals klimaatverandering, armoede, honger, biodiversiteitverlies. Beoordeling, vóór de besluitvorming over de maatregelen, door wetenschappelijke methodes kan inzicht geven in de effecten van potentiële beleidsmaatregelen of innovaties om deze complexe problemen te managen. Integrated Assessment and Modelling (IAM) is een methode om beoordelingen vooraf uit te voeren door middel van kwantitatieve modellen. Eén type IAM koppelt modellen die specifieke processen beschrijven op een schaalniveau in modelketens die meerdere schaalniveaus en disciplines beslaan. Zulke modelketens kunnen de effecten simuleren van beleidsmaatregelen en technologische innovaties op duurzaamheidindicatoren op meerdere schaalniveaus en disciplines.

Om een operationele modelketen voor IAM te ontwikkelen, is interdisciplinaire integratie van modellen, databronnen, indicatoren, en scenario's vereist. Methodologische, semantische en technische integratie zijn het onderwerp van dit proefschrift. Deze types van integratie zijn uitgewerkt binnen het IAM project SEAMLESS met ca. 30 onderzoeksgroepen en meer dan 100 betrokken onderzoekers. SEAMLESS had als doelen (1) om een aantal modellen in het landbouw domein te koppelen, (2) een modellenraamwerk te ontwikkelen voor geïntegreerde toetsing, dat op een personal computer uitgevoerd kan worden, (3) procedures te ontwikkelen voor geïntegreerde toetsing van agrarische systemen en (4) toepassingen uit te voeren met de modelketen in het raamwerk volgens de ontwikkelde procedures.

Dit proefschrift beschrijft methodologische, semantische en technische integratie in twee case studies. De eerste case studie is de integratie in zogenaamde bio-economische bedrijfsmodellen op twee schaalniveaus (e.g. bedrijf en veld); aan deze casus nam een relatief kleine groep van tien agronomen en economen deel. De tweede case studie is het modelleren van Europese landbouwsystemen. In deze casus beslaat de integratie vijf schaalniveaus (e.g. veld, bedrijf, regio, land en continent). Verschillende types van modellen worden gekoppeld door een grote groep van ongeveer honderd wetenschappers uit de agronomie, economie, landschapsecologie, informatietechnologie en milieuwetenschappen.

Een bio-economisch bedrijfsmodel (BEFM) koppelt beslissingen over de middelen aanwezig op het landbouwbedrijf aan beschrijvingen van huidige en alternatieve productiemogelijkheden. Deze productiemogelijkheden worden beschreven door middel van input-output relaties, waarin ook milieueffecten meegenomen worden. De basis voor de methodologische integratie van BEFMen is een literatuurstudie van de beschrijvingen, theoretische beschouwingen en toepassingen. Hoofdstuk 2 introduceert een classificatie van de verschillende types van BEFMen, bediscussieert sterke en zwakke kanten en concludeert met een onderzoeksagenda. Volgens deze onderzoeksagenda moet toekomstig onderzoek een adequate modelomschrijving, de beschikbaarheid van sourcecode of programma en de evaluatie van modelresultaten garanderen. Daarbij heeft een generiek en makkelijk te gebruiken BEFM diverse voordelen. Tot slot worden alternatieve productiemogelijkheden voor boeren, en hun strategische beslissingen en sociaal milieu vaak genegeerd in BEFMen.

De methodologische integratie van verschillende methodes om alternatieve agrarische productiemogelijkheden te analyseren wordt beschreven in hoofdstuk 3. Alternatieve productiemogelijkheden zijn activiteiten die niet nu op bedrijven gebruikt worden, maar die wellicht geschikte alternatieven voor de toekomst zouden kunnen zijn, gegeven nieuwe ontwikkelingen. De geïntegreerde methode om alternatieve activiteiten te definiëren is gebaseerd op productie-ecologische principes, die voor heel Europa toepasbaar zijn. De methode tracht inclusief te zijn in het beschouwen van geschikte alternatieven om te voorkomen dat geschikte alternatieven vooraf uitgesloten worden. De methode is gebaseerd op het genereren van alternatieven als permutaties en het vervolgens filteren van deze mogelijke permutaties op toepasbare alternatieven met behulp van heuristische filters.

Zowel methodologische als technische integratie worden beschreven in de ontwikkeling van de Farm System SIMulator (FSSIM) als een generiek BEFM (Hoofdstuk 4). FSSIM is een bio-economisch bedrijfsmodel, dat is ontwikkeld om bedrijfsreacties van Europese landbouwsystemen (bijv. veeteelt of één of meerjarige gewassystemen) te simuleren. Het bestaat uit twee gedeeltes, FSSIM-Mathematical Programming (MP) en FSSIM-Agricultural Management (AM). FSSIM-MP is een positief statisch model dat risicogedrag van bedrijven in beschouwing neemt en gecalibreerd kan worden volgens diverse methodes. FSSIM-AM beschrijft huidige activiteiten, genereert alternatieve activiteiten en kwantificeert de activiteiten door alle input-output coëfficiënten te beschouwen. Een generiek bedrijfsmodel moet toepasbaar zijn voor een reeks van klimaat- en bodemzones, voor een reeks van bedrijfstypes, voor toepassingen op technologische innovaties en beleidsmaatregelen, voor toepassingen die een verschillend detailniveau vereisen en voor het linken met modellen op andere schaalniveaus. Alle toepassingen van FSSIM werden geëvalueerd op deze criteria en met deze evaluatie wordt beargumenteerd dat FSSIM aan deze criteria voldoet en dus een generiek BEFM genoemd kan worden.

De tweede case studie, betreffende het modelleren van Europese landbouwsystemen, vereist de integratie van meerdere modellen, databronnen, indicatoren en scenario's en bouwt voort op de integratie gerealiseerd voor het BEFM. Een belangrijke vereiste voor modelintegratie is de semantische en technische integratie van databronnen in een geïntegreerde database met een gedeeld database schema (Hoofdstuk 5). De geïntegreerde database omvat de input- en outputdata voor de modellen, de applicaties en de scenario's gedefinieerd in het modelraamwerk, de data die van oorspronkelijke databronnen afkomstig zijn en indicatorwaarden berekend in de applicaties van de modelketen. De oorspronkelijke databronnen zijn bronnen over bedrijfseconomische resultaten, bodem, klimaat, beleid, gewasteelt, veehouderij en handel in agrarische producten. De ontwikkeling van deze geïntegreerde database vereiste de gedeelde definitie van concepten in een ontologie.

Modellen kunnen gebaseerd zijn op verschillende definities van plaats en tijd, zijn van verschillende aard (e.g. dynamische simulatie, optimalisatie, regressie) en beschrijven een selectie van fysische, biologische, economische of sociale processen. Deze modellen gebruiken vaak verschillende namen en definities voor dezelfde concepten. De methodologische, semantische en technische integratie van modellen in een modellenketen is vereist om beoordeling van effecten van technologische en beleidsveranderingen mogelijk te maken (Hoofdstuk 6) over meerdere schaalniveaus en disciplines. Een gewasgroei simulatiemodel en een bio-economisch bedrijfsmodel worden geïntegreerd door te werken met een gedeelde definitie van landbouwactiviteiten en ze leveren verschillende parameterwaardes om het 'activiteit' concept te beschrijven. De parameterwaardes berekend door het ene model worden gebruikt als input voor het andere model. Een bio-economisch bedrijfsmodel, een econometrisch regionaal regressiemodel en een partieel evenwichtsmodel voor de agrarische sector worden geïntegreerd door de elasticiteit van de vraag, die geschat wordt met het econometrische regressiemodel op basis van de bedrijfsreacties van het bio-economische bedrijfsmodel. Deze elasticiteit van de vraag dient als input voor het partieel evenwitchtsmodel. De koppelingen tussen de modellen worden expliciet, reproduceerbaar en transparant gemaakt door de belangrijke concepten en berekeningen expliciet te beschrijven. Deze berekeningen zijn nodig om het ene concept in het andere concept te vertalen.

Scenario's zijn een belangrijk concept in IAM om toepassingen met behulp van IAM programma's te structureren. Het concept 'scenario' wordt gebruikt met verschillende betekennissen in verschillende disciplines en kan leiden tot verwarring in een IAM project of het kan gebruikt worden als een containerbegrip voor onbekende en ongespecificeerde zaken. Om semantische integratie van scenario's te bereiken is het begrip scenario duidelijk gedefinieerd in een gedeelde conceptualisatie ('ontologie') van meerdere onderzoekers (Hoofdstuk 7). Scenario in een toepassing van het IAM model is gedefinieerd als een experiment, dat bestaat uit een beleidsoptie, context en verwachtingen m.b.t. trends. Deze scenariodefinitie is tot stand gekomen door een iteratief en gezamenlijk proces waarin vele onderzoekers betrokken waren.

Methodologische, semantische en technische integratie vereisen elk een andere benadering (Hoodstuk 8). Methodologische integratie door middel van een expliciete beschrijving van relevante concepten en benodigde berekeningen om concepten te vertalen leidde tot een transparante en expliciete modelkoppeling van een dynamisch gewasgroei model, een bio-economisch bedrijfsmodel, een econometrisch regressiemodel en een partieel evenwichtsmodel voor de agrarische sector. Echter, er is een gebrek aan methodes om modellen van verschillende types (e.g. dynamische simulatie, optimalisatie en regressie) en met verschillende proces-, plaats- en tijdbeschrijvingen te integreren voor landgebruikmodellering. In semantische integratie is het gebruik van ontologie en gezamenlijke inspanningen om de ontologie te ontwikkelen een sterke combinatie die intensief in het SEAMLESS project gebruikt is. De benaderingen die gebruikt zijn om de gezamenlijke inspanningen vorm te geven waren afhankelijk van het integratievraagstuk (bijv. databronnen, modellen of scenario's) en de onderzoekers uit de verschillende disciplines die daarbij betrokken waren. De ontologie ontwikkeld in de semantische integratie kan gecombineerd worden met krachtige modellenraamwerken, zoals Open-MI, om modelketens dynamisch door te rekenen op een PC, op een flexibele en transparante manier. Hoewel krachtige oplossingen uit de IT beschikbaar zijn voor modelintegratie, is de werkelijke toepassing ervan de integratie van modellen niet triviaal.

Communicatie is cruciaal in elke integratiebenadering en een expliciete communicatiestrategie is dan ook een vereiste. Een dergelijke communicatiestrategie is vaak afwezig in methodologische, semantische of technische integratie. Hoewel de case studies laten zien dat methodologische, semantische of technische integratie mogelijk zijn, wordt geconcludeerd dat ze slechts het begin zijn en demonstreren wat mogelijk is. Om integratie en interdisciplinair onderzoek in grote onderzoeksprojecten verder vorm te geven, moeten rigoureuze, geteste en gedocumenteerde methodes voor elk type van integratie ontwikkeld worden.

Curriculum Vitae

Sander Jansssen was born on 19 September 1980, in Heerlen, Netherlands. Heerlen was also the city where he went to Lyceum, from 1993 to 1999 at Sintermeerten College, where he graduated with distinction in courses physics, chemistry, mathematics, biology, Dutch, French, English and Acient Greek. In 1999 he commenced his academic degree at Wageningen University, then still Landbouw Universiteit Wageningen. In 2005 he completed a masters degree (i.e. ingenieurs diploma) according to a free program focused on agronomy, organic agriculture, systems analysis and environmental economics. Two theses were part of this degree, one on goals and motivations of organic farmers, and the other a review of bio-economic farm models, which became the second chapter in this thesis.

During his student time at Wageningen University, Sander spent time abroad at Kopenhagen University, Unversity of Wales Aberystwyth and Polytechnic University of the Marche, Ancona, where he became fluent in English and Italian. As extra-curriculur activities, he was a board member of AIESEC in 2002-2003 and Stichting RUraal Wageningen (RUW) in 2002-2004. He contributed to the realization of an information portal on master theses at Wageningen University, a conference on the future of European countryside (with RUW) and the organization of week for students on organic agriculture in the Marche, Italy.

His PhD project was a joint project between the business economics and plant production systems group of Wageningen University, which implied having a desk at each of the groups. The PhD project was part of a large research project funded by the 6th framework programme of European Commission, codenamed SEAMLESS. Sander had a crucial role in the development of a farm model, coordinated the database and ontology task force within this project and contributed organizationally and conceptually to the model linking in the SEAMLESS project. In the SEAMLESS project, he experienced interdisciplinary research and learned to work with different cultures and disciplines in research project.

Since November 2008 he works as a researcher at Alterra, Centre for Geo-Information. He has an interest in interdisciplinary research, scientific integration, semantic modelling, simulation modelling, knowledge management, knowledge annotation, agricultural systems and organic agriculture.

Publication list

Journal papers

Athanasiadis, I. N., Janssen, S., 2008. Semantic mediation for environmental model components integration. Information Technologies in Environmental Engineering 1, 3-11.

Ewert, F., Van Ittersum, M. K., Bezlepkina, I., Therond, O., Andersen, E., Belhouchette, H., Bockstaller, C., Brouwer, F., Heckelei, T., Janssen, S., Knapen, R., Kuiper, M., Louhichi, K., Olsson, J. A., Turpin, N., Wery, J., Wien, J. E., Wolf, J., 2009. A methodology for integrated assessment of policy impacts in agriculture. Environmental Science & Policy, In press.

Janssen, S., Van Ittersum, M. K., 2007. Assessing farm innovations and responses to policies: A review of bio-economic farm models. Agricultural Systems 94 (3), 622-636.

Janssen, S., Andersen, E., Athanasiadis, I. N., Van Ittersum, M. K., 2009. A database for integrated assessment of European agricultural systems. Environmental Science & Policy, In press.

Janssen, S., Ewert, F., Li, H., Athanasiadis, I. N., Wien, J. J. F., Thérond, O., Knapen, M. J. R., Bezlepkina, I., Alkan-Olsson, J., Rizzoli, A. E., Belhouchette, H., Svensson, M., Van Ittersum, M. K., In press. Defining assessment projects and scenarios for policy support: use of ontology in Integrated Assessment and Modelling. Environmental Modelling and Software, In press.

Thérond, O., Belhouchette, H., Janssen, S., Louhichi, K., Ewert, F., Bergez, J.-E., Wery, J., Heckelei, T., Olsson, J. A., Leenhardt, D., Ittersum, M. K. V., 2009. Methodology to translate policy assessment problems into scenarios: the example of the SEAMLESS Integrated Framework. Environmental Science & Policy, In press.

Conference papers

Andersen, E., Janssen, S., Athanasiadis, I.N., Rizzoli, A.E., Villa, F., Wien, J., 2008. Data and knowledge management for integrated assessment: A SEAMLESS example. In: Impact Assessment of Land Use Changes, International Conference, April 2008, Berlin

Athanasiadis, I. N., Janssen, S., Andersen, E., Rizzoli, A. E., Wien, J. J. F., 2009. Granular ontologies for integrated assessment of agricultural systems. In: Proceedings of Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Policy and Society (AgSap), Egmond aan Zee.

Athanasiadis, I. N., Rizzoli, A. E., Janssen, S., Van Ittersum, M. K., 2007. Generation of artificial crop rotation schemes as cyclic permutations. In: Donatelli, M., Hatfield, J., Rizzoli, A. E. (Eds.), Farming Systems Design

2007: An International Symposium on Methodologies on Integrated Analysis on Farm Production Systems, Catania, Italy.

Athanasiadis, I.N., Janssen, S., Huber, D., Rizzoli, A., Van Ittersum, M.K., 2007. Semantic modelling in farming systems research: The case of the Agricultural Management Definition Module In: Marx Gómez, J., Sonnenschein, M. ,Müller, M. et al.(eds.) Proceedings of Information Technology in Environmental Engineering (ITEE 2007), March 2007, Pages 417–432

Bezlepkina, I., Van Ittersum, M.K., Ewert, F., Alkan Olsson, J., Andersen, E., Brouwer, F., Donatelli, M., Flichman, G., Janssen, S., Heckelei, T., Olsson, L., Oude Lansink, A., Rizzoli, A., Wery, J., Wien, J.E., Wolf, J. Assessing economic, social and environmental processes in agriculture with a computerized framework for the EU (SEAMLESS-IF) In: Green Week Scientific Conference 2007 "Managing Economic, Social and Biological Transformations". Berlin, Germany

Bezlepkina, I., Alkan-Olsson, J., Athanasiadis, I.N., Janssen, S., Ruinelli, L., Knapen, R., Li, H., Bockstaller, C., Belhouchette, H., and Therond, O., 2007.Integration of models, data and expert knowledge by means of indicators: The SEAMLESS project experience. In: Farming Systems Design 2007: An international symposium on Methodologies for Integrated Analysis of Farm Production Systems M. Donatelli and J. Hatfield and A. Rizzoli (Eds.) Vol. 2 – Field-farm scale design and improvement. Catania, Sicily, Italy, p. 237-238

Borkowski, N., Hecker, J.-M., Louhichi, K., Blanco Fonseca, M., Janssen, S., Van Ittersum, M.K., Stokstad, G., Zander, P. 2007. Surveying crop management data for bio-economic farm models. In: Hatfield, J., Donatelli, M., Rizzoli, A. (Eds.), Farming Systems Design 2007: An international symposium on Methodologies for Integrated Analysis of Farm Production Systems, Catania, Sicily, Italy, 2007, Vol. 1, pp. 33-34.

Ewert, F., M.K. Van Ittersum, I. Bezlepkina, O. Therond, E. Andersen, H. Belhouchette, C. Bockstaller, F. Brouwer, T. Heckelei, S. Janssen, B. Jonsson, M.J.R. Knapen, M. Kuiper, K. Louhichi, J. Alkan Olsson, L. Ruinelli, A.E. Rizzoli, N. Turpin, J. Wery, J.J.F. Wien, J. Wolf, 2009. A methodology for enhanced flexibility of integrated assessment of policy impacts in agriculture. In: Proceedings of Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Policy and Society (AgSap), Egmond aan Zee.

Janssen, S. J. C., Wien, J. J. F., Hongtao, L., Athanasiadis, I. N., Ewert, F., Knapen, M. J. R., Huber, D., Thérond, O., Rizzoli, A. E., Belhouchette, H., Svensson, M., Ittersum, M. K. v., 2007. Defining projects and scenarios for integrated assessment modeling using ontology. . In: MODSIM07 International Congress on Modelling and Simulation "Land, Water and Environmental Management: Integrated Systems for Sustainability", Christchurch, New Zealand.

Janssen, S., Andersen, E., Athanasiadis, I. N., Van Ittersum, M. K., 2008. A pan-European database for integrated assessment and modeling of agricultural systems. In: International Congress on Environmental Modelling and Software, Barcelona.

Janssen, S., Louhichi, K., Kanellopoulos, A., Zander, P., Flichman, G., Hengsdijk, H., Li, H., Heckelei, T., Stokstad, G., Borkowski, N., Hecker, M., Lansink, A. O., Blanco, M., Meuter, E., Van Ittersum, M. K., 2009. Farming Systems SIMulator: first generic bio-economic farm model. In: Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Policy and Society (AgSap), Egmond aan Zee.

Kanellopoulos, A., H. Hengsdijk, S. Janssen, M.K. Van Ittersum, 2009. Ex-ante assessment of the abolishment of the EU set aside policy: Results from a bio-economic farm analysis. In: Proceedings of Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Policy and Society (AgSap), Egmond aan Zee.

Knapen, M. J. R., Athanasiadis, I. N., Jonsson, B., Huber, D., Wien, J. J. F., Rizzoli, A. E., Janssen, S., 2009. Use of OpenMI in SEAMLESS. In: Proceedings of Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Policy and Society (AgSap), Egmond aan Zee.

Li, H., Louhichi, K., Janssen, S., Rizzoli, A. E., Athanasiadis, I. N., Meuter, E., David, H., 2007. Development and Application of a Component-based Generic Farm System Simulator Implementing a Semantically Enriched Integrated Modelling Framework. In: Proceedings of the International Symposium on Environmental Software Systems 2007, Prague.

Louhichi, K., Hengsdijk, H., Janssen, S., Bigot, G., Perret, E., 2009. EU dairy farming in the face of change: An exploration using a bio-economic farm model. In: Proceedings of Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Policy and Society (AgSap), Egmond aan Zee.

Meuter, E., Janssen, S., Van Ittersum, M. K., 2009. A Graphical User Interface for a generic bio-economic farm model (FSSIM). In: Proceedings of Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Policy and Society (AgSap), Egmond aan Zee.

Oomen, R., Hengsdijk, H., Therond, O., Bergez, J. E., Russell, G., Casellas, E., Janssen, S., Van Ittersum, M. K., 2009. Development of generic management rules for crop growth simulation models In: Proceedings of Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Policy and Society (AgSap), Egmond aan Zee.

Thérond, O., Turpin, N., Belhouchette, H., Bezlepkina, I., Bockstaller, C., Janssen, S., Alkan-Olsson, J., Bergez, J.E., Wery, J., Ewert, F., 2008. How to manage models outputs aggregation for indicator quantification within SEAMLESS Integrated Framework. In: Impact Assessment of Land Use Changes, International Conference, April 2008, Berlin.

Turpin, N., O. Therond, F. Ewert, I. Bezlepkina, J. Alkan Olsson, C. Bockstaller, S. Janssen. 2009. Spatial upscaling of on-farm indicators: Concepts to use a scientific knowledge base to match policymakers' requirements. In: Proceedings of Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Policy and Society (AgSap), Egmond aan Zee.

Wien, J. J. F., Knapen, M. J. R., Janssen, S. J. C., Verweij, P. J. F. M., Athanasiadis, I. N., Li, H., Rizzoli, A. E., Villa, F., 2007. Using Ontology to Harmonize Knowledge Concepts in Data and Models. In: MODSIM 2007 International Congress on Modelling and Simulation, Australia and New Zealand, Christchurch, New Zealand.

Wien, J.J.F., A.E. Rizzoli, M.J.R. Knapen, I.N. Athanasiadis, S. Janssen, L. Ruinelli, F. Villa, M. Svensson, P. Wallman, B. Jonsson, 2009. Architecture of SEAMLESS-IF as framework application for integrated modelling. In: Proceedings of Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Policy and Society (AgSap), Egmond aan Zee.

Book chapters

Louhichi, K., Janssen, S., Li, H., Borkowski, N., Kanellopoulos, A., Van Ittersum, M. K., Flichman, G., Hengsdijk, H., Zander, P., Blanco, M., Stokstad, G., Athanasiadis, I., Rizzoli, A. E., Huber, D., 2009. A generic farming system simulator. In: Brouwer, F., Van Ittersum, M. K. (Eds.), Environmental and agricultural modelling: integrated approaches for policy impact assessment. Springer Academic Publishing.

Wien, J. J. F., Rizzoli, A. E., Knapen, M. J. R., Athanasiadis, I. N., Janssen, S. J. C., Ruinelli, L., Villa, F., Svensson, M., Wallman, P., Jonsson, B., Van Ittersum, M. K., 2009. A web-based software system for model integration in agro-environmental impact assessments. In: Brouwer, F., Van Ittersum, M. K. (Eds.), Environmental and agricultural modelling: integrated approaches for policy impact assessment. Springer Academic Publishing, pp. 24.

Reports

Janssen, S., Hengsdijk, H., Oomen, R., Van Ittersum, M.K., 2009. Agricultural Management-module of FSSIM: procedures to generate and quantify alternative agricultural activities -Production Enterprise Generator, Production Technique Generator and Technical Coefficient Generator-, PD3.3.3, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org, Wageningen, pp 108.

Janssen, S., Louhichi, K., Kanellopoulos, A., Zander, P., Flichman, G., Hengsdijk, H., Li, H., Heckelei, T., Stokstad, G., Borkowski, N., Hecker, M., Lansink, A. O., Blanco, M., Meuter, E., Van Ittersum, M. K., 2009. Integration of all

FSSIM components within SEAMLESS-IF and a stand alone Graphical User Interface for FSSIM, D3.3.12, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2), www.SEAMLESS-IP.org, 51 pp.

Janssen, S., I.N. Athanasiadis, E. Andersen, J. Wien, F. Villa, A.E. Rizzoli and F. Ewert, 2009. Knowledge Base for SEAMLESS-IF: Achieving integration through ontologies in interdisciplinary research, PD1.4.2, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org, 62 pp.

Rizzoli, A., Ruinelli, L., Knapen, R., Janssen, S., Athanasiadis, I., 2009. Procedures to integrated new components in SEAMLESS-IF, PD5.6.2.3, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org, wageningen, pp 38.

Thorne, P.J., Hengsdijk, H., Janssen, S., Louhichi, K., Van Keulen, H., Thornton, P., 2009. Modelling Livestock Component in FSSIM, PD 3.3.4, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org pp. 65.

PE&RC PhD Education Certificate

With the educational activities listed below the PhD candidate has complied with the educational requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)

Review of Literature (5.6 ECTS)
- Review article in agricultural systems: Assessing farm innovations and responses to policies: A review of bio-economic farm models (2005-2006)

Writing of Project Proposal (5 ECTS)
- Describing and exploring possibilities in agricultural management for ex-ante assessment of policies and technological innovation through bio-economic farm models (2006/2007)

Laboratory Training and Working Visits (2.5 ECTS)
- Semantic modelling and ontologies; IDSIA (2005)
- Bio-economic modelling in GAMS; IAMM (2005)

Post-Graduate Courses (4.8 ECTS)
- 20 Scientific workshops as part of the SEAMLESS project; SEAMLESS Partners (2005-2009)
- FSSIM Framework expert week (SEAMLESS); WU-PPS, IAMM, IDSIA (2007)
- Understanding global environmental change: processes, compartments and interactions; SENSE (2009)

Competence Strengthening / Skills Courses (4.8 ECTS)
- Time planning and project management; WGS (valley consult) (2006)
- PhD Competence assessment; WGS (2006)
- Organization and leadership for PhD Students UU; Utrecht University (2008)

Discussion Groups / Local Seminars and Other Scientific Meetings (4.2 ECTS)
- Mats & Stats (2006-2008)

PE&RC Annual Meetings, Seminars and the PE&RC Weekend (1.2 ECTS)
- PE&RC Weekend (2006)
- PE&RC Day 'Collapse' (2007)

International Symposia, Workshops and Conferences (6.3 ECTS)

- MODSIM; Christchurch (2007)
- iEMSs; Barcelona (2008)
- AgSAP Conference; Egmond aan Zee (2009)

Acknowledgements and funding

We thank all scientists in the SEAMLESS project who contributed to development described in this thesis (e.g. model linking, database development, FSSIM, ontologies, scenario definition). This work has been carried out as part of the SEAMLESS Integrated Project, EU sixth Framework Programme, Contract No. 010036-2. We gratefully acknowledge access to the multiple data sources by EU-FADN-DG AGRI-G3 for FADN data, JRC for JRC/AGRIFISH Data Base, European Soil Bureau for European Soil Database, University of Bonn for COCO/CAPREG and SEAMLESS partners for agricultural management data.

The research described in this thesis was financially supported by the European Commission.

Financial support from Wageningen University for printing this thesis is gratefully acknowledged. Printed by Wöhrmann Print Service, www.wps.nl in Zutphen