Centre for Policy Modelling, MMUBS, Manchester
Stockholm Environmental Institute, Oxford Office
Universität Kassel
Politechnika Wroclawska
IIASA, Vienna
Macaulay Institute, Aberdeen
Uniwersytet Wroclawski

CAVES Initiation Report

Deliverable No. 1 of Project 012816: CAVES – Complexity, Agents, Volatility, Evidence and Scale

Duration: 2005-2008

Funded under the EU 6FP NEST programme.

Table of Contents

1 Introduction 5

2 Team integration 5

3 Modelling issues 6

3.1 Centre for Policy Modelling 6

3.1.1 Modelling tools 6
3.1.2 Ontology of Scottish Case Study 9
3.1.3 Fine-Grained Model Prototype 10
3.1.4 Directions for Future Work 10
3.1.5 References 10

3.2 Universität Kassel 11

3.2.1 Organisational setup 11
3.2.2 Towards coarse grained models of land and water use 12
3.2.3 Outlook: Further steps 16

3.3 Macaulay Institute 17

3.3.1 Ontology of South African Case Study 17
3.3.2 Workshop on Ontologies 18
3.3.3 Enhancement of FEARLUS Model 19
3.3.4 Glossary 19
3.3.5 Directions for future work 19
3.3.6 References 20

3.4 Politechnika Wroclawska 20

3.4.1 Introduction 20
3.4.2 Description of work 20

4 Case studies 23

4.1 South African case study (SEI Oxford) 23

4.1.1 Background to the case study 23
4.1.2 CAVES briefing note 23
4.1.3 Ontology development 24
4.1.4 Conceptual Mapping 25
4.1.5 Links with other projects 27
4.1.6 References: 27

4.2 Grampian case study (Macaulay Institute) 28

4.2.1 Recruitment of Social Scientist 28
4.2.2 Specification of Research Questions 28
4.2.3 Source Assessment and Evidence Gathering Exercise 31
4.2.4 Ontology (Designed by Ruth Meyer of MMU) 31
4.2.5 Directions for Future Work 32

4.3 Odra Valley case study (Uniwersytet Wroclawski) 32

4.3.1 Evidence 32
4.3.2 GIS 34
4.3.3 Literature 35
4.3.4 Steps to be undertaken in the nearest future 35

5 Schedule of meetings 35

6 Amendments to the project work plan 36

1 Introduction

This report is Deliverable No. 1 of the CAVES Project. Its purpose is to set out a more detailed scheme of work for the project and initial state of preparedness for that work. Specifically, it describes the plans for the development of the project modelling and knowledge elicitation capabilities.

A key element of the project proposal was the dominance of evidence in the design and implementation of models and the use of the models to support rather than to dominate the development of policy scenarios by stakeholders. Whilst several of the project partners – the Centre for Policy Modelling, the Macaulay Institute and the University of Kassel – bring to the project relevant competence in agent based modelling, their respective approaches have been different. Consequently, an important early step in the life of the project will be to bring these approaches together to produce modelling technology and process that will be (at least in the first instance) particularly well suited to the case studies of the Odra Valley in Poland, the Limpopo Basin in South Africa and the Grampian Region of Scotland in the United Kingdom.

In parallel with the initialisation of the modelling capacity of the project, initial work is being undertaken on the case studies.

2 Team integration

Because the project team members come from a wide range of disciplines, we have decided to establish a glossary that will be developed throughout the lifetime of the project. The glossary will provide a focus for agreement on what we in the project shall mean by the use of particular words and phrases even though they have a variety of meanings across disciplines. The initial glossary was compiled by Gary Polhill of MI in consultation with all partners and is to be appended to the final version of this document.

Different approaches to the modelling of complexity are taken by physicists and the agent based social modellers in the project. In order to further the discussion about complexity and how complex processes are to be modelled and generally to explicate the differences between social and physical complexity, Piotr Magnuszewski will write a paper on different approaches to complexity in relation specifically to modelling.

One issue that is of considerable importance to researchers in the area of land use change and which has patently crucial policy implications is that of resilience. In order to bring this issue into the work domain of the project, and to develop a shared understanding that might be expressed in an ontological visualisation, Jan Sendzimir (IIASA) will write a review of resilience measures. The use of such measures will not only inform the modellers' understanding of the relevant issues, but might also provide statistical measures for the process of validating simulation models at macro level.

Although there are differences between the social and physical scientists' approaches to complexity, all participants in the CAVES project agree that the core of the models will be network representations of social relations. Joanna Stefanska is being asked to review social network data to bring prevailing empirical understandings and approaches to social networks into the project discussions.

3 Modelling issues

Different modellers and domain experts have their own views of the relationships among concepts and how these can be modelled. The issues here are considered by philosophers under the heading of *ontology*. The meaning of *ontology* in the community of agent based computer scientists is rather more specific than the philosophical meaning of the word in that it relates mainly to vocabulary and grammar in the protocols for communication amongst agents. In this case (see glossary), ontology means the structure of relations among relevant concepts. For the purposes of the CAVES project, ontologies will be developed relating concepts that are shared by modellers and domain experts and then expressing those relationships in a model design.

The following sections report on the work undertaken by the three different modelling teams during the first six months of the CAVES project.

3.1 Centre for Policy Modelling

3.1.1 Modelling tools

The focus of work in the first six month of the CAVES project has been on modelling tools, in particular tools that enable declarative modelling. It is our understanding that declarative modelling is often the most appropriate technique to capture social phenomena (Moss and Edmonds 2005) whereas many physical or biological processes are best described by numerically based formalisms. Since the models developed in the CAVES project will need to represent both, social and physical processes, it is therefore important that a modelling environment should support declarative modelling as well as imperative modelling.

Our first approach was to develop a package, which supports the implementation of declarative (that is, rule-based) features in otherwise conventional models. The package is written in Java and is especially designed to integrate with any agent-based social simulation software, as long as they are implemented in Java. Thus, the declarative package is not a complete simulation system; rather it extends other Java-based systems, such as Repast[1] or Swarm[2], to include declarative features.

The system uses a MySQL[3] database as a back end to manage the (potentially) large volume of data that can accumulate in social simulations, but its use is transparent to the modeller. The database stores both the knowledge of the agents and the rules that the

[1] http://repast.sourceforge.net/
[2] http://www.swarm.org/
[3] http://www.mysql.com/

agents use to manipulate that knowledge. The modeller uses a graphical user interface to specify both the types of knowledge that the agents can have and the rules that manipulate this knowledge. The system then translates these into tables and data in the database. At run time, at each time step, the system determines which rules are applicable for each agent (that is, which rules have conditions that can be evaluated), and applies them (that is, performs the knowledge updates that appear in the actions).

The development interface provides additional tools to allow the modeller to experiment with the clauses and queries that will be used in rules, and also to inspect the database. A standard set of operators are defined for operations on clauses (including `and`, `or`, `count` and so on), but advanced modellers also have the option to add their own operator definitions, by sub-classing the standard operator definitions. The basic classes of the declarative package are illustrated in figure 1.

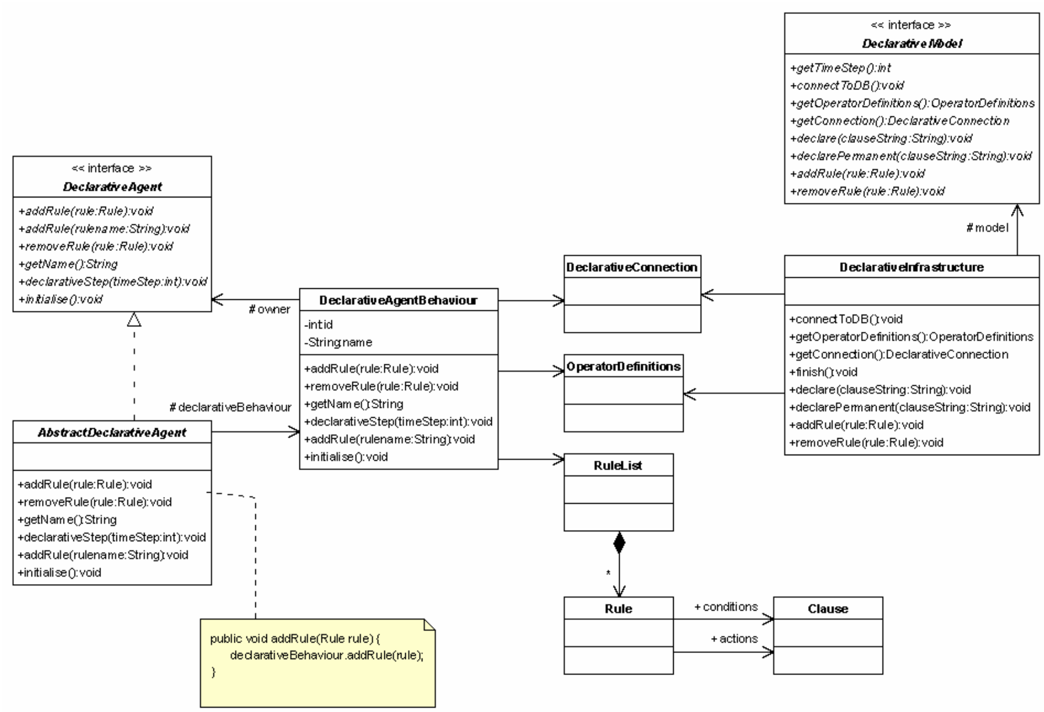

Fig. 1: The key classes in the declarative package.

Once the clauses and rules have been defined, the modeller simply sets up a model in an existing simulation framework. Each model class using the declarative package has to implement the `DeclarativeModel` interface, ensuring that the model contains an instance of `DeclarativeInfrastructure`, which provides the connection to the database (via `DeclarativeConnection`). The functionality for rule-based agent behaviour is encapsulated in the `DeclarativeAgentBehaviour` class. It keeps a reference to the database connection to store and retrieve an agent's rules. All agents must implement the `DeclarativeAgent` interface and use an instance of

`DeclarativeAgentBehaviour`. For convenience, they may directly subclass `AbstractDeclarativeAgent`, which already does both, if the applied simulation framework does not require the use of its own agent classes. For each agent, the modeller must specify which rules initially apply using the `addRule()` method (more can be added via rules at run time), and then the model can be run using the default scheduler.

This framework was tested by re-implementing existing models. The first was Jiggle, one of the standard demonstration models that are bundled with both Repast and Swarm. The second was a model of the emergence of symbiotic groups, which was developed by Bruce Edmonds in SDML (Edmonds 2005) and has subsequently been re-implemented using Repast. Our re-implementations using the declarative package make use of Repast's standard scheduler, using both the graphical and batch-mode features of the scheduler. These experiments allowed us to compare the efficiency characteristics of the declarative package with those of standard Repast. Unfortunately, it turned out that the data structures and algorithms used for the declarative package are very inefficient. This is due to the fact that the package borrowed some concepts and their realisation from the SDML language developed and extensively applied at the CPM over the past ten years. Whereas SDML has an underlying list structure, this is not one of Java's strengths.

For efficiency reasons, we therefore decided to investigate existing alternatives. Combining Repast with JESS, the Java Expert System Shell[4], was deemed the most promising candidate. JESS is a rule engine and scripting environment developed at Sandia National Laboratories by Ernest Friedman-Hill. Although not open source, it is available free of cost for academic purposes, including the source code. Since it is written entirely in Java and allows for calling Java methods from rules, it integrates well with any Java software.

JESS consists of a rule interpreter which can apply both forward and backward chaining, using an improved version of the fast but memory-intensive RETE algorithm (Forgy 1982) to match facts from the fact base to rules in the rule base. Declaring facts and rules is done via a script language with a LISP-like syntax. This language supports not only the manipulation of symbolic facts but also method calls on arbitrary Java objects, thus facilitating the combination of declarative modelling and imperative modelling.

As a first test of this combination, we re-implemented the Jiggle model once more, this time specifying the behaviour of the Jiggle agents as JESS rules. Each agent has its own instance of the rule engine, which results in a longer model setup phase compared to the original Jiggle model.[5] Apart from that, the re-implementation with JESS runs similarly fast and scales rather well, as shown in figure 2. Since JESS is a memory-intensive application, its performance is sensitive to the behaviour of the Java garbage collector. Forcing the garbage collector to run more often (e.g. by invoking it directly or indirectly via reducing the model's heap size) improves the overall performance. These experiments led to the conclusion that combining JESS with Repast (or an equivalent

[4] http://herzberg.ca.sandia.gov/jess/

[5] The newest version of JESS (7.0xb) allows agents to share one rule engine. This might speed up the model setup phase considerably and probably even cut down on memory requirements. Further experiments will have to prove if this expectation will hold true.

agent-based simulation tool like MASON[6] or MadKit[7]) will be the route of choice within the CPM to pursue for the CAVES project.

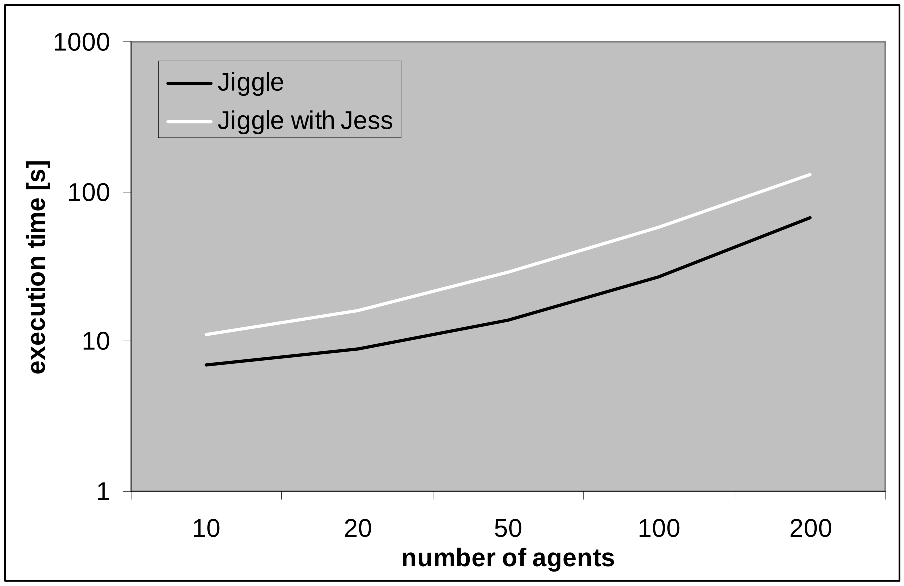

Fig. 2: Comparing execution times with regard to the number of simulated agents.

3.1.2 Ontology of Scottish Case Study

As agreed at the project meeting in Aberdeen at the end of June, we also started work on an ontology for the Scottish case study and a prototype fine-grained model based on the demonstrator ontology Gary Polhill developed for the South African case study. This included becoming acquainted with the software tool Protégé[8], a free, open source ontology editor and knowledge-base framework developed at Stanford University.

The ontology for the Scottish case study is based on a number of journal articles describing results of several surveys of farm households that where conducted in the Grampian region between 1987 and 2002 (Shucksmith and Winter 1990, Shucksmith and Smith 1991, Shucksmith 1993, Shucksmith and Herrmann 2002, Burton 2004). Additional material was provided by the Macaulay Institute, comprising informal notes of a meeting discussing the latest survey and the draft of a conference paper.

A first version of the ontology has been developed and is currently in the process of being refined based on the feedback of our domain expert at the Macaulay Institute. We followed the methodology for developing ontologies as established by the Macaulay team (see section 3.3) as close as possible. Since the ontology is meant to aid in the process of

[6] http://cs.gmu.edu/~eclab/projects/mason/
[7] http://www.madkit.org/
[8] http://protege.stanford.edu/

model building, the focus during its development has been on identifying not only the main concepts and actors in the domain but especially their actions (responses) and the factors that trigger them (stressors). While trying to relate these to each other, we have found that ontologies are a good tool to capture the static structure of a model domain but are not very well suited to model dynamic processes like agent's behaviours. We will take this into account during the revision of the Grampian ontology.

3.1.3 Fine-Grained Model Prototype

The development of a fine-grained model has been affected by lack of necessary data. While the South Africa ontology provides a comprehensive object model, defining detailed attributes of agents and other entities, it does not specify any actions of the agents, be they reactive (responses to a certain stimulus) or pro-active (pursuing goal-directed strategies). Empirical data on the physical environment has also not been available. Therefore, we decided to start with a fairly abstract model that captures the basic processes, nevertheless. Features like physical characteristics of the land or the climatic conditions are currently modelled as homogenously or randomly distributed; this can easily be replaced by the accurate empirical data.

For a model of the processes in the domain, we referred to the FEARLUS model developed at the Macaulay Institute. We also obtained a crop model from SEI Oxford, which we integrated into our model. At present, the fine-grained model prototype comprises a 2D grid environment, where each grid cell represents a land parcel, and a group of agents representing the stakeholders. Stakeholders own different land parcels, choose crops to grow on them, and buy or sell land parcels if the crop yield enables or forces them to do so, respectively. New stakeholders can enter the system via buying a land parcel, whereas other stakeholders may leave when they sold all of their land. We are currently working on incorporating different strategies for production and trade of land parcels.

3.1.4 Directions for Future Work

- Refinement of the Grampian ontology with respect to the input from the domain experts.
- Enhancement of the prototype fine-grained model by incorporating greater diversity in the physical characteristics of the environment, introducing stakeholder-stakeholder interaction, and implementing different stakeholder strategies to cope with the climatic variability.
- Evaluation of other Java tools for agent-based social simulation regarding their combination with JESS.

3.1.5 References

Burton, R. (2004): Seeing through the 'good farmer's' eyes: towards developing an understanding of the social symbolic value of 'productivist' behaviour. In: Sociologica Ruralis 44(2), pp 125ff

Edmonds, B. (2005): The emergence of symbiotic groups resulting from skill-differentiation and tags. In: Proceedings of the AISB 2005 Symposium on Emerging Artificial Societies.

Forgy, C. (1982): Rete: A Fast Algorithm for the Many Pattern/Many Object Pattern Match Problem. *Artificial Intelligence* 19, pp 17-37

Moss, S., Edmonds, B. (2005): Sociology and simulation: statistical and qualitative cross-validation. *American Journal of Sociology* 11(4), pp 1095-1131

Shucksmith, M. (1993): Farm Household Behaviour and the Transition to Post-Productivism. In: J Ag Econ 44(3), pp 466-478

Shucksmith, M.; Herrmann, V. (2002): Future Changes in British Agriculture: Projecting Divergent Farm Household Behaviour. J Ag Econ 53(1), pp 37-50

Shucksmith, M.; Smith, R. (1991): Farm Household Strategies and Pluriactivity in Upland Scotland. In: J Ag Econ 42(3), pp 340-353

Shucksmith, M.; Winter, M. (1990): The Politics of Pluriactivity in Britain. In: J Rural Studies 6(4), pp 429-435

3.2 Universität Kassel

3.2.1 Organisational setup

The Centre for Environmental Systems Research is responsible for the development of coarse grained models of land and water use with special attention to the Odra River case study and for coordinating the project work package 3 (modelling). In Kassel there are two people working on the CAVES project. In May 2005, Friedrich Krebs has joined the team. He studied Mathematics and Environmental Systems Analysis and is responsible for the development of agent-based models with focus on the evolution of network structures. In mid July 2005, Claudia Zehnpfund started working in the project. She is a mathematician as well and is planning to do her doctorate within the CAVES project with focus on social network analysis and modelling. Andreas Ernst is a full professor at the Centre for Environmental Systems Research at the University of Kassel. He leads and organises the working group in Kassel.

During the last six months the Kassel team attended several CAVES meetings:

- The kick-off workshop, 28 February - 02 March 2005, that was held in Manchester, where the course was set for the next months,
- The modellers' meeting in Aberdeen, 21-22 June 2005, mainly to talk about ontologies,
- Informal consultations at the ESSA (European Social Simulation Association) conference at the beginning of September in Koblenz, to discuss further developments,
- From September 26th to September 30th, the biannual project meeting with all the

participating partners will take place in Wroclaw/Poland. There will be a presentation of the first results and a discussion about further developments.

3.2.2 Towards coarse grained models of land and water use

Network theory

Mathematical network theory provides a wide variety of network analysis methods. In order to get a good grasp of the concepts relevant to our modelling tasks, some time was spent to learn about social network analysis and the small world concept. There are many ways to describe social network data mathematically. We focussed mainly on the graph theoretic approach and the algebraic approach. Of special interest were concepts about central actors in the social network and cohesive subgroups. Cohesive subgroups are subsets of agents among whom there are comparatively strong, direct, intense or frequent ties.

Interfacing with the case studies: Ontologies

In our context, ontology means the structure of relations among relevant concepts. An ontology is typically a hierarchical data structure containing all the relevant entities and their relationships and rules within the particular domain. The main objective for using ontologies was to find a common ground and vocabulary for the different teams and approaches. At the Modellers' Meeting in Aberdeen in June, all the participants were introduced to Protégé, the standard ontology editor, in order to learn how to apply the ontology development procedures and software. In Kassel, the ontology of the South African case study was used, among other sources, to design first simulation models. Nevertheless we decided not to develop an own ontology for the Odra River case study. Ontologies are primarily designed to provide a static view on relationships, whereas we are also very much interested in processes (i.e. actions and perceptions using network links) and in dynamic network evolution. This information has probably to be transferred from the case studies to the modelling teams in a somewhat less formalised way.

Modelling concepts: Multi-layer networks of intelligent agents

In the modelling approach for the coarse grained models we propose here the strict distinction between physical environment and social environment of the agents play a key role. This distinction focuses on a separation between physical and social spaces both in terms of semantics and techniques used for their representation.

For various reasons that are detailed below, the simulation of the agent's physical environment uses a (traditional) grid based approach. The social "location" of an agent is given by his position within a social network context, where an agent is viewed as a node and social relations are represented by edges. Since agents are usually considered in more than one social context an agent's social environment generally consists of more than one network layer. The modelled agents' perceptions vary related to their physical or social environment (or context). Both perceptions are locally bounded in terms of a perceivable section of the surrounding physical space and in terms of network edges and neighbouring nodes. In the same way, the agents' repertoire of actions differs relating to their respective environment.

Locality in physical environment and in social contexts

We feel that the underlying concept of locality both in the physical and in the social space is crucial for the modelling concept that we are developing. Therefore, this section focuses on the different concepts of locality in a physical and a social network sense.

The modelled physical environment is represented as a grid topology. This allows for straightforward integration of static geo-referenced data from GIS layers as well as of dynamic processes in the physical environment. "Locality" on a grid topology has to be understood in the context of a certain metrics, i.e. concept of distance. Given an agent located on a position in the physical environment and a certain metrics, it is easy to define measures for the perceivable area of the environment or the area an agent may interfere with. It has to be noted that the extent of an agent's local physical environment is bounded by the grid topology. E.g. on a classical regular 2D (checkerboard) grid torus, there are exactly eight direct neighbouring cells for each location. This is a distinct difference to a network topology where a node may have an arbitrary number of neighbouring nodes (of one edge distance). Besides being spatially bounded an agent's perception could also be semantically bounded, e.g. concerning the access to different GIS layers.

The agents' social environments are modelled as networks. An agent may be seen as a node in different social network contexts. Technically, an agent object has slots that are nodes representing potential or actual social roles in different networks. Unlike other modelling approaches agents do actively perceive their social environment and are enabled to act in their social network. To illustrate this further, consider the following example: An agent has two semantically different nodes: one in an acquaintances network and one in an adviser/supporter network. This would represent a two layered social network environment. The acquaintances network could be initialised with empirical data or with an assumed small-world topology. The adviser/supporter would initially consist only of the respective nodes without interconnections. Under certain conditions (perhaps given by an individual internal state) an agent would poll her neighbours in her acquaintances network to find out if any of those agents has a more suitable strategy to cope with (physical) environmental challenges. If the strategy provided by the network neighbour has been successfully applied, the agent would build up an edge to the respective node in the adviser/supporter network context. Next time the agent would first poll the adviser/supporter context when looking for a new strategy. It has to be noted that the sole interconnection between the social network layers exists inside the cognitive units of the agents (see below). The same holds for interconnections between physical and social layers.

On a network topology the concept of "locality" is quite different from a grid topology. First, the number of network neighbours of an agent may be arbitrary, or at least it may be arbitrarily scaled by assuming a maximum capacity of outgoing edges a node may have. Second, when using multiple layers of networks, locality may be multi-dimensional, i.e. in our example an agent's social locality may consist of network neighbours in the acquaintances network and a different set of neighbours in the adviser/supporter context.

In the following section, we will make some remarks on an agent architecture that uses the described concepts of physical and social environments.

Agent architecture

In order to describe the proposed agent architecture we take the rather technical approach to separate the agent into functional components. Thus, we imagine that a simple agent has some kind of perception unit that will provide the agent with information about both his physical environment and his social environment. Second, there is a (in a first step fixed) repertoire of actions that the agent may apply to his environments. Third, the agent is equipped with a cognitive unit that will provide him with some means (e.g. a simple rule base and an internal state) to decide about his next action based on his current perceptions. The remainder of this section consists of the descriptions of these functional agent components.

The **perception unit** generates information about the agent's physical and social environment. The perception of the physical environment provides local information about environmental attributes like resource availability, types of land cover, or the locations of other agents. As mentioned earlier, the accessible attributes of the physical environment may differ depending e.g. on the type of the agent. The perceived social network environment is represented by lists of network neighbour nodes. Generally, these lists of nodes originate from multiple network layers. The agent "knows" about the semantics of each of those lists (as in the example above, it is known whether a network perception relates to the acquaintances network or the adviser/supporter network). Again, it seems to be a reasonable assumption that network perception is strictly local, i.e. no agent within the network has a global, bird's eye view of the whole network.

The **actions** an agent may execute in his physical environment are straightforward: The simulated environment should allow the agents to move/migrate, to access resources, or to apply land use strategies. Side effects of actions in the physical environment on the social environment should be explicitly considered (e.g. migration may cause social network connections to break). Thus, an agent receives delayed and implicit feedback of the consequences of his actions in the physical environment. Actions in the agent's social environment are network-related modifications like strengthening/weakening of outgoing and/or incoming edges, establishment of new edges in already established networks/social contexts or access to new networks/social contexts by building first edges in a new network layer. Again side effects in the physical environment should be considered.

For the coarse grained models we propose that the agent's **cognitive unit** is based on rule execution, which can model as well routine actions as deliberate decision making. Learning through basic adaptation mechanisms occurs in response to feedback of actions. Feedback may originate from the agent's physical environment or from his social contexts. Furthermore, knowledge can be acquired through more advanced mechanisms either by observation or by action feedback.

A first South African and Polish model

Within the participating partners there was found the consensus to use the Recursive Porous Agent Simulation Toolkit (Repast) as the principal modelling software. As a starting point for the modelling activities, a small demonstrator was developed using the Repast framework. The purpose of the demonstrator is to show how evidence from the case studies can be used to construct a simple model that follows the modelling concept outlined above. The model is coarse grained in the sense that it represents knowledge derived from the case studies on a more abstract level. On the abstract level, we then apply the above concepts of simple intelligent agents, layered social networks and a common physical environment. The demonstrator mainly shows how the proposed modelling concepts are applied. We have not done yet any detailed analysis or further validation of the model.

For the demonstrator we start out with a physical environment in which each grid cell only has one attribute that stands for the type of land cover on that location. The simulated environment is split into five equally sized areas that each have one distinct land cover. During the course of the simulation, the land cover on a randomly chosen land parcel shifts to a different type to model shocks on the system.

The social environment consists of two network layers: One layer is an acquaintances network that is initialised with a pre-generated small-world network with a given average node degree. This layer remains fixed over the whole simulation run. The second network layer is an advisor/supporter network where agents actively construct network connections to other agents that have already provided them with useful information (see below). In this layer, unused edges slowly decay in strength and disappear once their weight becomes zero.

The simulated agents apply a random land use strategy at their position in the environment. There are five such strategies. Exactly one strategy fits one type of land cover. Agents possess an internal state that reflects their individual level of satisfaction with their land use strategy. It decays over time. Periodically, the internal level of satisfaction is increased if the land use strategy is appropriate for the type of land cover. If an agent however uses an inappropriate land use strategy, the satisfaction level will eventually drop below a certain threshold value. In that case the agent will look in its social environment for a more appropriate land use strategy. The manner in which this is done depends on the type of the agent. There are two types of agents that we refer to as imitators and innovators, respectively. Innovators try out different strategies by themselves until they are satisfied. Imitators use their social context to investigate useful strategies: They first poll their advisor/supporter network to get help from agents that have previously been helpful. If there is an agent in that network layer that lives on the same land cover, uses a different land use strategy, and has a higher level of satisfaction, the agent will simply copy (imitate) this agent's strategy and refresh the weight of the network edge to the supplier of the strategy. If polling the advisor/supporter context does not yield useful information, the agent will use his acquaintances network. If in that case a strategy is copied from a network neighbour, the agent that copied the strategy will build a network edge to the supplier of the strategy in the advisor/supporter context.

The simulation is initialised with randomly distributed agents with randomly initialised strategies. A small proportion of agents (1%-2%) are innovators, the remainder are imitators. The dynamics of the model is roughly as follows: Given a sufficient connectivity of the acquaintances network (e.g. an average node degree of 20), a sufficiently large number of agents, and a reasonable proportion of innovators, the information about useful strategies will quickly spread over the two network layers allowing almost all of the agents to apply the correct land use strategy. During this the advisor/supporter network builds up. Once the correct strategies are found the advisor/supporter network fades (because it is no more used) and disappears. When a shift of land covers occurs agents on the "shocked" land parcel have to look for a new strategy. Again the advisor/supporter context will build up. The semantic relation to historic macro-level phenomena in the Odra region will have to be discussed in more detail throughout the project.

3.2.3 Outlook: Further steps

Further steps in the Kassel project will relate to

- Incorporating data and further concepts from the Polish case study into the coarse grained model architecture as described above.
- Integrating rule based decision making mechanisms into the architecture to allow for incorporating more complex yet modular knowledge structures. To that end, the Java library Jess will be tested. Jess is a rule engine and scripting language and serves as a tool for adding rules technology to Java-based software systems.
- Analysing networks with advanced techniques. In the network analysis we deal with the social network layers that will emerge from the simulation. The most important issue in the construction of our models will be to make sure that our empirical data will be reflected qualitatively in the models. A major task is therefore to find a type of network that fulfils certain previously defined conditions. Among others these measures might include cohesiveness, connectedness, density, structural balance or centrality.

  What we hope for is empirical data from the Polish case study about the actual social networks, so that measures of the actual social networks can be compared with the simulated ones on a qualitative level to validate emergent network structures. In connection with the case study it has therefore to be analysed which type of network should be used. In the study, there are several surveyed villages. We expect that the social connection within the villages is very close, whereas there might be only few contacts between people from different villages. Hence one possibility might be to use small-world-networks to illustrate the reality within the villages and then to connect the villages by a few edges.

  Based on the Polish case study the actors that are connected to other actors in another village might be important or central in the sense that they are the intermediaries between two villages and therefore the only way that actors of different villages get in contact is a path that includes one or several of these actors. These actors might even be cut points. A vertex is called a cut point, if deleting the vertex results in a graph with more components. However there are also other possibilities to define the concept of an actor's centrality or prestige, such as

- visibility, i.e. not only direct edges but also within reach via paths from different vertices
- vertices that are connected to many other vertices
- high prestige, in directed graphs: The arrows point to this actor
- closeness: how close is an actor to all the other actors in the network. The idea is that an actor is central if he can communicate/interact with the others quickly.

It remains to be seen which of these concepts will be of avail in the actual implementation or if it will be necessary to define new theoretical concepts.

3.3 Macaulay Institute

3.3.1 Ontology of South African Case Study

As agreed at the inception meeting in March 2005, the Macaulay modelling team developed an ontology of some earlier work by colleagues at the Stockholm Environmental Institute in South Africa, covering similar issues to the South African case study proposed for CAVES. The purpose of this ontology was to act as a demonstrator for other modelling teams to use as a basis for developing ontologies of the CAVES case study areas.

Ontologies are defined by Gruber (1993) as "formal, explicit specifications of a shared conceptualisation": formal in that they are machine processable; and explicit in that all pertinent concepts and relationships between them are represented. A conceptualisation is a model—an abstract representation of some real-world phenomenon, and Gruber stipulates that this abstract representation should be shared by some particular group, though more recent work on ontologies in Computer Science has cast doubt on whether this is a necessary condition of ontology formulation.

Ontologies are based on description logics, such as SHIQ. Ontological statements, such as one class of concept being a particular kind of another class of concept, form axioms in that logic, allowing them to be used as the basis for making formal inferences from the ontology, using software developed for this purpose. This allows the consistency and satisfiability of concepts in the ontology to be formally verified, and also permits instances to be classified in an ontologically consistent manner. It was suggested at the inception meeting that ontologies offered an intermediary between texts and computer programs that were nonetheless machine processable, and that they might be useful for comparing case studies.

The Macaulay team developed a methodology for developing ontologies from textual evidence (in this case, a number of journal articles written about the case study). This methodology is a seven-stage process as follows:

1. Assembling evidence. All materials containing information that should be explicitly described in the ontology is gathered together.
2. Preliminary examination of evidence. The evidence is examined for key concepts and central themes, which will suggest (respectively) classes and properties in the ontology.
3. Detailed analysis. All occurrences of key concepts are highlighted in the source evidence, and surrounding text (adjectives in noun phrases, or even whole sentences) is used to suggest subclasses, data type and object properties, and supporting classes.
4. Ontologies for supporting classes. Supporting classes describe concepts that are not discussed in detail in the source evidence, but are nonetheless assumed to exist as common knowledge or common sense. Examples are concepts of geographical areas (e.g. biophysical versus administrative regions) or social networks. Ideally supporting classes would be imported from existing ontologies developed elsewhere.
5. Defining classes. So-called ‘defined classes’ have necessary and sufficient conditions (rather than just necessary conditions) for class membership defined, enabling reasoning software to infer class membership for particular individuals. Defined classes offer the potential for ontologies to be used to consistently classify case study evidence.
6. Ontology testing. The ontology is tested for consistency using the reasoning software. Individuals are created and classified to see if the inferences made are as expected. The ontology is treated as a database for a case study to look for concepts that are superfluous or are missing.
7. Iteration. The ontology is refined by repeating from step 1 in the light of information received from the tests. New evidence can also be brought in at this stage, and other ontologies linked to.

This methodology was used to develop an ontology of work on climate information dissemination following the South African food security crisis of 2002-03, by CAVES colleagues Sukaina Bharwani and Gina Ziervogel. The ontology focused on describing various classes of farmer discussed in their work, and in particular definitions of qualitative, but politically sensitive terms such as ‘vulnerable farmers’. A user guide was prepared for the ontology, and Gary Polhill met with Gina, Sukaina, and other members of the SEI team in July 2005 to discuss the ontology. Evidence from the original case study has been entered in to the ontology, and we are now ready for the next phase of iteration in refining it.

3.3.2 Workshop on Ontologies

A workshop on ontologies was organised at the Macaulay Institute in June 2005. The purpose of this workshop was to explain ontologies to the rest of the modelling team and to disseminate the South African case study. The workshop featured talks by Nick Gotts on the Web Ontology Language (OWL), and the South African case study ontology developed using OWL; by Gary Polhill explaining the methodology used to develop the ontology of the South African case study, introducing Ontologies and giving training in the use of the Protégé Java application for developing them; and an invited lecture by Dr. Alun Preece from the University of Aberdeen on wider aspects of ontologies.

One of the issues raised with OWL ontologies at this workshop was the lack of a formalism for describing processes (dynamics) in a case study. Nonetheless, in terms of describing concept hierarchies and relationships between concepts, ontologies offer a more transparent formal description of a scenario than the source code of a more traditional object-oriented computer model, with at least the potential for the evidence base to be rigorously established.

3.3.3 Enhancement of FEARLUS Model

Work on enhancing the FEARLUS agent-based model of land use change has progressed in two threads. In the first thread, the FEARLUS model has been updated with a prototype land market model, which was presented at this year's meeting of the European Social Simulation Association. The land market model, in contrast to currently existing such components of agent-based land use change models, does not assume that farmers are profit-maximisers. This assumption cannot be justified on the basis of evidence gathered by Dr. Robert Burton at the Macaulay Institute during earlier work in the UK. However, in dropping the assumption, several questions have to be answered about how real farmers make decisions to buy and sell land.

In the second thread, designs are being developed for generalising the biophysical component of FEARLUS, to incorporate the use of look-up tables. These lookup tables are designed to be flexible enough to allow incorporation of any biophysical factors that are believed to influence farm decision making and/or yield in a symbolic form, and are intended to specify, for all combinations of land use/management decisions and influencing factors, the outcomes that affect farmers and policy-makers. Though the approach is resource-intensive in terms of memory, it has the advantage of facilitating integration with existing biophysical models.

3.3.4 Glossary

A glossary of terms that are likely to have a special meaning within CAVES has been developed and released on the CVS server, using initial definitions taken from a dictionary and other sources such as Wikipedia. It is hoped that over time these definitions will be refined to reflect the specific meaning that they have to colleagues in CAVES.

3.3.5 Directions for future work

- Establishing the role of ontologies in developing consistent descriptions of case studies as a tool for social scientists to use.
- Linking ontologies to models, and developing means of formally describing processes within ontologies.

- Refinement of the South African ontology, insofar as this provides a basis for the ontology of the CAVES South African case study.
- Gathering evidence on real farmers' decision-making processes with respect to land markets.
- Implementation of look-up table biophysical model in FEARLUS.
- Updates to the glossary as required.

3.3.6 References

Gruber, T. R. (1993): A translation approach to portable ontology specifications. *Knowledge Acquisition*, **5**, 199-220.

3.4 Politechnika Wroclawska

3.4.1 Introduction

The team form the Wrocław University of Technology is responsible mainly for the development of abstract models and theories based on results of the CAVES project (WP4). We also have a supporting role in building models for the Odra case study (WP3) and data gathering process in the Odra river case study region (WP2). For the last six months our team has been working on three themes: development of abstract models, cooperation with the Odra case study team and exploration of the Repast library. Members of our team attended two CAVES meetings: the kick-off meeting held in Manchester between the 28th of February and the 2nd of March 2005 and the modellers' meeting held in Aberdeen between the 21st and the 22nd of June 2005.

3.4.2 Description of work

Development of abstract models

Although development of abstract models needs to be based on models developed for different case studies, it is necessary from the beginning to work on possible concepts, approaches and research questions. The following themes were considered during a series of meetings of WP4 group in WUT:

- Literature review: Self Organized Criticality, Small World Networks and Scale-Free Networks, models of Sznajds and Deffuant.
- Volatility, its definitions and its occurrence in social phenomena. The relation between SOC and volatility. Conditions for obtaining volatile behaviour in social models.
- Threshold models: Granovetter's model, Rolfe's extension.
- Brock's and Durlauf's model of decision making: mathematical formulation and

mean field approximation. Social interpretation of physical terms as phase transition, hysteresis and temperature.

- Numerical simulations of Brock–Durlauf model: discussions on methods (Monte Carlo study of system dynamics vs. Metropolis-like algorithm); nearest neighbours' interactions vs. long range interactions; going beyond the uniform model: varying parameters between individuals.
- Extension of Brock–Durlauf model: conditions for multi-stability; discussion about correspondence between social and physical models, meaning of "social temperature", microscopic background of statistical mechanics and non-existence of its social analogue.
- The difference between the models with individual utility functions and the "social planner": existence of hysteresis ("social traps").
- Our considerations concerning utility function that would describe the process of joining and leaving a collective action corresponding to the Odra Valley case study.
- Further numerical analysis of Brock–Durlauf model: observation of hysteresis.
- Discussions about obtaining a two-maximum utility function in the mean field approximation and nearest-neighbours model; considerations about modelling SOC phenomena in social systems: how to introduce cumulative character of interactions, the nature of external tuning.

The more comprehensive report on these issues will be available within next 2 months.

Cooperation with the Odra case study team

Throughout the last six months we have been supporting the Odra case study team in their data gathering activities. We have taken part in the workshops with members of local communities organised in Zaborów (10th of May) and Kwiatkowice (11th of May). We have also participated in the workshops with domain experts conducted in Wrocław (8th of June).

The case study team has identified three major problems of the local communities: land improvement, flood protection and social initiatives. It was our responsibility to estimate the extent to which each of these problems can be a base for a computer model. As the land improvement issue seemed to be the most promising one, in terms of creating an interesting and policy relevant model, we started our investigation with examining various software tools for modelling hydrology, nutrient transport and vegetation. We have looked at the following libraries: SWAT (Soil & Water Assessment Tool), ANSWERS2000, AGNPS (AGricultural Non-Point Source), ACRU (Agricultural Catchments Research Unit), LISFLOOD, KINEROS2 and TOPOG. These libraries differ in the scope of phenomena that they can simulate. SWAT seems to have the widest scope ranging from surface runoff, evapotranspiration and groundwater to climate, nutrient and pesticide transport, erosion, land cover and management practises. Other libraries focus mainly on erosion, rainwater distribution and floods. They all, however, share the same problematic feature – they are written in C or FORTRAN. This makes them less portable, than if they were written in Java and a little harder to integrate with a Java program. Fortunately these problems are not very significant. An issue of greater significance emerged while we learned more about the use of those libraries. It was the amount and type of data necessary to run a simulation. We are still uncertain if we will be capable of gathering such data.

Exploration of the Repast library

Repast (Recursive Porous Agent Simulation Toolkit) was chosen as the base programming platform for the implementation of models created in the CAVES project. As our team will undoubtedly work with these models, we decided to get acquainted with the Repast library, and get to know its abilities, limitations and overall characteristics. Repast is a relatively large library, so the acquaintance process is not yet over, however some possibly significant findings have already been made.

The only really effective way of exploring the possibilities of any library is to test it in a real-life or near real-life application. So we looked around for a simple model of land use in agriculture that we could implement in Repast in relatively short time. We found the Von Thünen model. It was created, in its original form, in the 19th century and it describes the relationship between agricultural land use, production and distance from market. Initially it was an analytical model, but it can be easily transformed into an agent based model in a discrete two dimensional space. A more detailed description of the Von Thünen model can be found at http://people.hofstra.edu/geotrans/eng/ch6en/conc6en/vonthunen.html.

We knew that Repast contained special classes for working with raster data and that it even provided a visual component for displaying them. Thus, we were quite optimistic about our little test-implementation of the Von Thünen model and expected a short and simple development process. To make the test slightly more demanding we decided to create a grid of 10^6 (1000x1000) cells. We also decided to introduce 3 markets and 100 farmers.

Before creating our test model we played around a bit with the GisBugsModel that comes with Repast as a demo, to see how the RasterSpace, Object2DGrid and display classes cope with spaces containing one million cells. The result was very clear - they don't. The RasterSpace class uses only 64 bit double precision numbers, which means that a raster of 10^6 cells takes little under 8 MB of RAM. The raster that we used was made from an 8 bit gif file and there was no reason to transform it into a 64 bit raster. The display classes were not suited to deal with large spaces as well. Moreover the DisplaySurface class has at least two quite obvious bugs in it. One makes it impossible to view the menu during simulation as it is being constantly overdrawn by the so called "painter". The second bug causes the whole application to hang, when an attempt is made to use the "zoom" feature of the display.

After this little ordeal we decided to create our own space and display classes. This obviously made our attempt to implement the Von Thünen model "a bit" longer, but in the end we could at least view the simulation. However the only part of Repast that we ended up using was the schedule and the basic GUI. Obviously we could have used classes from the "analysis" package, but the quality of graphs generated by these classes is somewhat disappointing. Hence, if we really wanted to have graphs, we would probably either create our own diagram plotting components, or adapt some third party components. So the main question that troubles us right now is: What is the point in using Repast if, even for such a simple model, you have to make the spaces and the display components on your own?

Final note: A description of our implementation of the Von Thünen model along with source code is available on demand.

4 Case studies

In this chapter the different case study teams report on the work undertaken during the first six months of the CAVES project.

4.1 South African case study (SEI Oxford)

4.1.1 Background to the case study

The South African case study is based in Vhembe district, Limpopo Province, South Africa. Previous SEI work has revolved around two projects. CLOUD (Climate outlooks and agent-based simulation of adaptation in Africa) focused at the village level, investigating the utility of seasonal climate forecasts for communal garden users. A variety of methods were used including interviews and knowledge elicitation tools (KnETS) that enabled decision rules to be explored around agricultural strategies and the drivers of strategies. This fed in to an agent-based model which is documented in Bharwani *et al.* (in press) and Ziervogel and Bharwani (submitted). This built on work that Emma Archer had undertaken as documented in Archer and Easterling (Draft) and Archer (2003). The other research was for FAO and focused on district level institutions in Vhembe (and Chikwawa, Malawi) and the support for food security and how appropriate it was given the high prevalence rates of HIV/AIDS.

4.1.2 CAVES briefing note

Previous field experience was used to inform a Briefing note for CAVES, submitted in July 2005 by Gina Ziervogel, entitled “Non Governmental Organisations in Vhembe, Limpopo Province, South Africa”: Their emergence from and consequent changes to social networks. This document explored the nature and emergence of organisations at village and district level. The information was based on experience in the district gained from involvement in fieldwork in other projects rather than specifically for this purpose.

4.1.3 Ontology development

An ontological model was developed by Gary Polhill in collaboration with Gina Ziervogel. Gary interpreted documentation related to the case study to create a model of farmers in Mangondi. It was an iterative process, where Gina would answer his questions and provide clarifications where necessary. The Ziervogel and Bharwani (submitted) paper formed the bulk of the information used for developing the ontology, with the other papers supplying supplementary information. An Ontology User Guide was developed to detail the ontological process and how it was developed for the Mangondi case. Figure 3 illustrates the Role class hierarchy used in the model. The reader is referred to [the Macaulay modelling report] for more information on the ontology development process.

A working session with the SEI staff (Tom, Sukaina, Tak and Gina) and Gary allowed for further explanation and exploration of the developed ontology. A few refinements to the ontology were suggested at this meeting, which have now been implemented. After the meeting, some of the data from previous fieldwork was entered into the ontology to explore how automated reasoning with the ontological model would handle instance classification queries, with a view to validating the ontology. Initial data have been entered but could be explored further in collaboration with the fieldworkers and the model developer.

Fig. 3. The Role class hierarchy, including the various defined subclasses of Farmer.

4.1.4 Conceptual Mapping

A conceptual diagram of the key flows of information from stressor to agents to response was developed as a team. This helped to capture the important factors to consider in a different way to the ontology. One strand of the complex interactions has been representation in Figure 4. This was explored as a potential method for complementing the ontology as it is able to integrate flows of information and processes in a rapid method that allows for multiple stakeholder engagement.

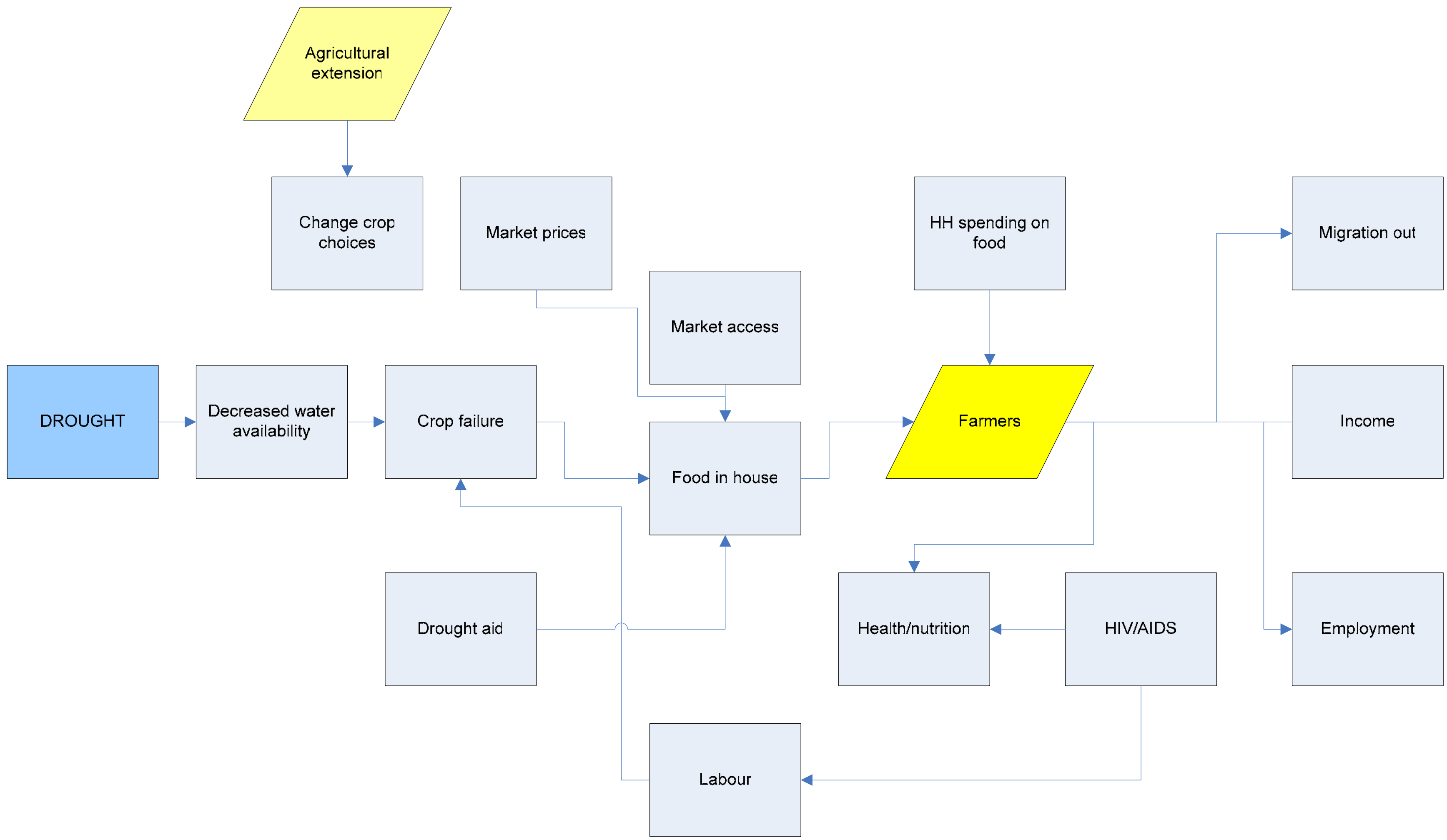

Fig. 4: A conceptual diagram of the impact of drought on small scale farmers in Mangondi.

4.1.5 Links with other projects

CAMP project: Catchment management and poverty[9]

An evaluation of the impacts of alternative forest policy instruments on water resources, economics and poor peoples livelihoods, South Africa, Tanzania, Grenada. This project was completed in 2003 and some of the researchers have expressed an interest in collaborating with the CAVES project (Ian Calder, Rob Hope and Graham von Maltiz).

This project compared and contrasted the application of integrated water resource management (IWRM) and sustainable livelihood (SL) approaches to land and water management within catchments in Limpopo, RSA as a means towards identifying policy instruments which both improve the livelihoods of poor people and protect the resource base. The project employed macro-scale hydrological/economic modelling combined with household-level SL assessment to examine the effects of alternative policy instruments relating to forestry and water allocation.

IFPRI project: Food and Water Security under Global Change: Developing Adaptive Capacity with a Focus on Rural Africa.

The goal of the project is to develop capacity to adapt to global change for vulnerable rural areas in developing countries. The project will provide policymakers and stakeholders in Ethiopia and South Africa, particularly farmers and other rural stakeholders who face the largest impact from global change, with tools to better understand, analyze, and form policy decisions that will allow them to adapt to global change. Their focus in South Africa is Limpopo Province.

This project started in March with a start-up workshop in Pretoria and they are implemented a household farm survey in July. Claudia is aware of CAVES and will ensure we are kept up to date with their project development.

The project partners include:

- IFPRI: Dr. Claudia Ringler, project leader and coordinator, Theme 5, CPWF; Dr. Siwa Msangi, postdoctoral research fellow for global change research and Dr. Mark Rosegrant, policy analysis.
- Center for Environmental Economics and Policy in Africa (CEEPA), University of Pretoria: Dr. Rashid M. Hassan
- Addis Ababa University, Department of Economics (AAU): Dr. Tassew Woldehanna, administration through Ethiopian Development Research Institute (EDRI)
- Centre for Marine and Climate Research, University of Hamburg: Dr. Richard S.J. Tol, Michael Otto Professor of Sustainability and Global Change

4.1.6 References:

Archer, E. (2003): Identifying underserved end-user groups in the provision of climate information. Bulletin of the American Meteorological Society, pp.1525-1532.

Archer, E.; Easterling, W. (Draft 2nd order): Mitigating climate stress: Strengthening the ‘end-to-end’ climate information system in South Africa.

Bharwani, S.; Bithell, M.; Downing, T.E.; New, M.; Washington, R.; Ziervogel, G. (in press): Multi-agent modelling of climate outlooks and food security on a community garden scheme in

[9] http://www.cluwrr.ncl.ac.uk/research_projects/recent_projects/prj_camp.php

Limpopo, South Africa. *Philosophical Transactions of the Royal Society B.*

CLOUD project website: http://www.geog.ox.ac.uk/research/projects/cloud/team.html

Ziervogel, G.; Bharwani, S. (Draft: February 2005): Adapting to variability: Pumpkins, pumps, poverty and the role of climate.

4.2 Grampian case study (Macaulay Institute)

4.2.1 Recruitment of Social Scientist

Dr. Lee-Ann Small was recruited to the Macaulay Institute to undertake the survey work for the Grampian Case study. Dr. Small has recently completed a Ph.D. at the University of Aberdeen entitled "Agriculture-based livelihood strategies in Bulgaria and Southern Russia: implications for agrarian change." She has extensive experience in the qualitative and quantitative research techniques that will be required to conduct the case study. She joined in August 2005.

4.2.2 Specification of Research Questions

Dr. Small began by considering how the statements about the role of case-studies in the project in the project documents could be expressed in terms of specific research questions which would guide her work. In the current draft, three broad questions are suggested as an organising framework:

1. How has (agricultural) land use at a study site in North East Scotland changed over the past twenty years?
2. Why does agricultural land use change?
3. What is the role of land users' social and informational networks in this process?

Each of the three questions was broken down into specific tasks and questions, as laid out below. Some of these questions concern the scope and focus of the case study, others are substantive. A list of these questions, with initial answers to some of them, will give the clearest overview of work within the case study.

How has (agricultural) land use in (North East) Scotland changed over the past 20 years?

Task: Describe land use change in (NE) Scotland over the past twenty or more years.

Related questions:

- What constitutes 'change' in land use?
  - How are we defining 'land use', what properties are we interested in?
    - What categories of land use will we use to determine 'change'?
    - Are we interested in intensity, commodity produced, soil quality, ownership/tenure? Answer: all of these.
    - Are we interested in organics or environmental practices; what about 'improving' land through drainage etc.?
      Answer: this will depend on whether these factors have been or are likely to be choices available to a significant proportion of farmers in the study area.
    - What about non-agricultural land uses – golf, environmental set-aside etc (or recreational agriculture – hobby farming and estates)?

Answer: this will depend on whether these factors have been or are likely to be choices available to a significant proportion of farmers in the study area.

- What were the land use practices and patterns 20 years ago; what are they now; how has this transition evolved over time (linear, sporadic, etc)?
- Where has land use changed the most? The least?

Issues: What features are we looking for in a study site?

Related questions:

- Are we trying to be representative of NE Scotland?

  Answer: no.

- Are we trying to include a range of land uses?

  Answer: yes, at least enough for farmers to have important land-use choices to make.

- Are we trying to include a range of land tenures?

  Answer: yes if possible: landowning and tenant farmers face different problems, and may well have different kinds of social and information networks.

- Are we looking for an area of high change? Low change? Both?

  The main contrasts in rate and type of change are expected to arise *between* the case studies, but if there were some internally to each that would be an additional advantage.

- Do we need to differentiate between rural and remote areas?

  Answer: we need to be aware that socio-economic ties with cities appear important in allowing family farming to continue in much of NE Scotland, and so be alert for differences between more and less remote parts of the study area.

- Are there specific regions where it would be easier to collect data?

  Answer: almost certainly, and ensuring that it will be possible to collect the desired information from the study site chosen will be crucial.

Why does agricultural land use change?

Task: Define the theoretical perspective utilised in the study, combining theories of complexity, agrarian change, networks, policy studies and others with evidence from recent events. This involves bringing together diverse literatures to form a theoretically consistent framework, a process which will evolve to a degree, over the course of the study.

Related issues:

- Complexity: which theories of complexity are we utilising and why?
  - Which are the key elements of complexity theory that are utilised in the project?
  - What are the impacts of various ‘shocks’ on land use? (i.e. BSE, policy changes etc.)

    Answer: this is of course a focus of the case study, but initial indications are that at least some shocks (the BSE and Foot-and-Mouth epidemics) had surprisingly little impact.
- Agrarian Change
  - farm household adjustment literature
  - commodity systems research

  - others
- Network Theory: which theoretical perspective on networks will be utilised in the study? Possibilities include:
  - Social Capital
  - Long's Actor Oriented Approach to Development Research
  - (Actor-network theory?)
  - (others)
- Policy Studies: Legal issues surrounding tenancy and land use.
  - A specific question here: what have been the effects of the recent Scottish land reform, giving tenants the right to buy in some circumstances?

What is the role of land users' (social) networks in this process?

Task: Evaluate the role of networks in land use change; further refining definitions based on the assumptions identified in question 2.

Related issues:

- What do we mean by 'networks', how are these conceptualised, and what are the primary schools of network conceptualisation?
- What properties of land users are we interested in: demographics; motivations; business practices? How have these changed?

  Answer: we are interested in all of these. The current situation may be relatively easy to assess by interview; changes over 20 or more years may not.

Here there are many general issues (those below are a subset of those being considered), and little concrete information:

- What are land users' networks?
- How are these networks formed?
  - Is this an intentional process?
  - Do new land users to the area join existing networks, or form new ones?
  - How have they changed over time?
  - How are they transferred between generations?
  - Are some networks embedded in the social structure; do you acquire networks based on class or social status?
- How do social norms and their transference interact with them?
- Do different 'types' of farmers have different networks?
- What kind of resources do networks provide – i.e. information, access to markets, access to labour
- What are the structural characteristics of the networks?
- Does everyone belong to networks?
- Are there differentiated positions within networks (i.e. controller – centre, periphery…)?
- Do ties weaken over distance – either geographic, or with disuse?

4.2.3 Source Assessment and Evidence Gathering Exercise

For the first of the three main research questions identified in the previous section, a list of possible information sources has been drawn up, and an initial assessment made of their potential; for the second and third this process is underway. The list of sources for the first main research question, with some comments on availability and usefulness, is as follows:

- Scottish Agricultural Census Summary Sheets by Geographic Area: June 2004
  - Parish level already purchased by the Macaulay for most years – need to apply for permission to use them but it should be free; will require aggregation
- Economic Report on Scottish Agriculture (2004 – annually back to 1999)
  - summaries of performance based on farm type and business size; not disaggregated by region, though
  - they breakdown income by type of farm – but not at the regional level
- Who Owns Scotland? (http://www.whoownsscotland.com)
  - only has data for about 34% of the land in Grampian region – more for the highlands
  - maps of land holdings – requires specific software to view correctly
- Land Cover of Scotland – 1988
  - Macaulay owned therefore free access
  - once the definitive guide, now a bit dated – people tend to use the Land Cover Map 2000 which is UK wide – achieved through automated processing which resulted in some accuracy issues
- Land Capability for Agriculture
  - GIS maps owned (and produced) by the Macaulay Institute
- Trends in Broad Habitats: Scotland 1990-1998 – Scotland National Heritage Commissioned Report
  - data for Scotland on change (and interaction between) 16 habitat types (as set out by the 1992 Earth Summit in Rio de Janeiro) – looks at three areas – lowlands, uplands, and marginal uplands and islands – therefore not Census regions so difficult to disaggregate
  - based on Countryside Surveys
- Field Research Data
  - Mark Shucksmith's 1987 and 1991 Study Data on 300 farms in the Cairngorms – useful for identifying the types of response – and response to question rates – which we might get in a formal survey.
  - Rob Burton's 2002 data from some 80 of the same farms.

A literature review of research papers and books on land use in Scotland is also underway.

4.2.4 Ontology (Designed by Ruth Meyer of MMU)

In response to Ruth Meyer's Grampian Region OWL ontology, Nick Gotts produced a response consisting of:

- An overview, confirming that the ontology captures a great deal of what is expressed in the source materials, and makes a firm foundation for the next stage of the work.
- A list of specific comments and suggestions on specific classes and properties in the

ontology.

- A note on a suggested framework of top-level classes for ontologies related to land use.
- A note on how this framework relates to the upper-level classes in Ruth's ontology.

4.2.5 Directions for Future Work

Given the requirements of the project, the date of Dr. Small's appointment, and the constraints which the farming year places on the timing of surveys of farmers, a timetable for the case study work has been drawn up, as follows:

- Analyse existing data from interview-based surveys of farmers in the region, and from existing sources on the suitability of land for different uses, and on actual land use over the period 1987-date. (August 2005 – February 2006 and July 2006 – September 2006).
- Design a further interview-based survey and decide on the areas in which to recruit interviewees. (September 2005 – April 2006).
- Carry out the interviews for this survey. (May 2006 – July 2006 and October 2006 – December 2006). *Note*: this is the activity around which the others have been arranged, since it is most feasible to carry out interviews with farmers in (roughly) October-March and late May-early July.
- Analyse the results of this survey, in consultation with other members of the CAVES research team. (July 2006 – September 2006 and January 2007 – March 2007).

4.3 Odra Valley case study (Uniwersytet Wroclawski)

4.3.1 Evidence

Quantitative data

We have collected quantitative data relevant for CAVES project's thematic scope and the spatial extent of the Odra River case study area. Information about both social and demographical characteristics of the region as well as land use issues on the level of municipalities/communes were extracted from the Central Statistical Office's (CSO) databases and later processed in accordance with the project's needs[10].

Data features:

- Thematic scope of collected data includes demography, labour issues, administration, land use issues (forestry and agriculture) and contains 220 different variables.
- Data sources: Central Statistical Office's databases contain yearly collected data from local administration, data from National Population and Housing Census and data from National Agricultural Census.
- Data format: We have compiled data for years 2002 and 2003 as Excel spread sheets. Time series for time period 1996-2003 are available as Excel spreadsheets and charts.

[10] Data for the time period of 1996-2004 are available in digital form on the CSO web site; earlier data are available in paper Statistical Annals.

Knowledge elicitation – field trips, interviews and workshops

During field trips we watched the study area to make general reconnaissance and compare the area's present state with maps. We have conducted several meetings and interviews with experts (local authorities, NGOs, scientists) in order to scope our field of interest and find possible modelling topics.

We performed two workshops with local communities in Zaborow and Kwiatkowice villages (see figure 5) on the 10th and 11th of May 2005. Invited people were asked to write down three main problems of the region, each on a separate sheet of paper. Then sheets were collected and sorted by subject. As a result we identified the main problems in the region: degradation of the land reclamation system, poverty and unemployment, flood risk.

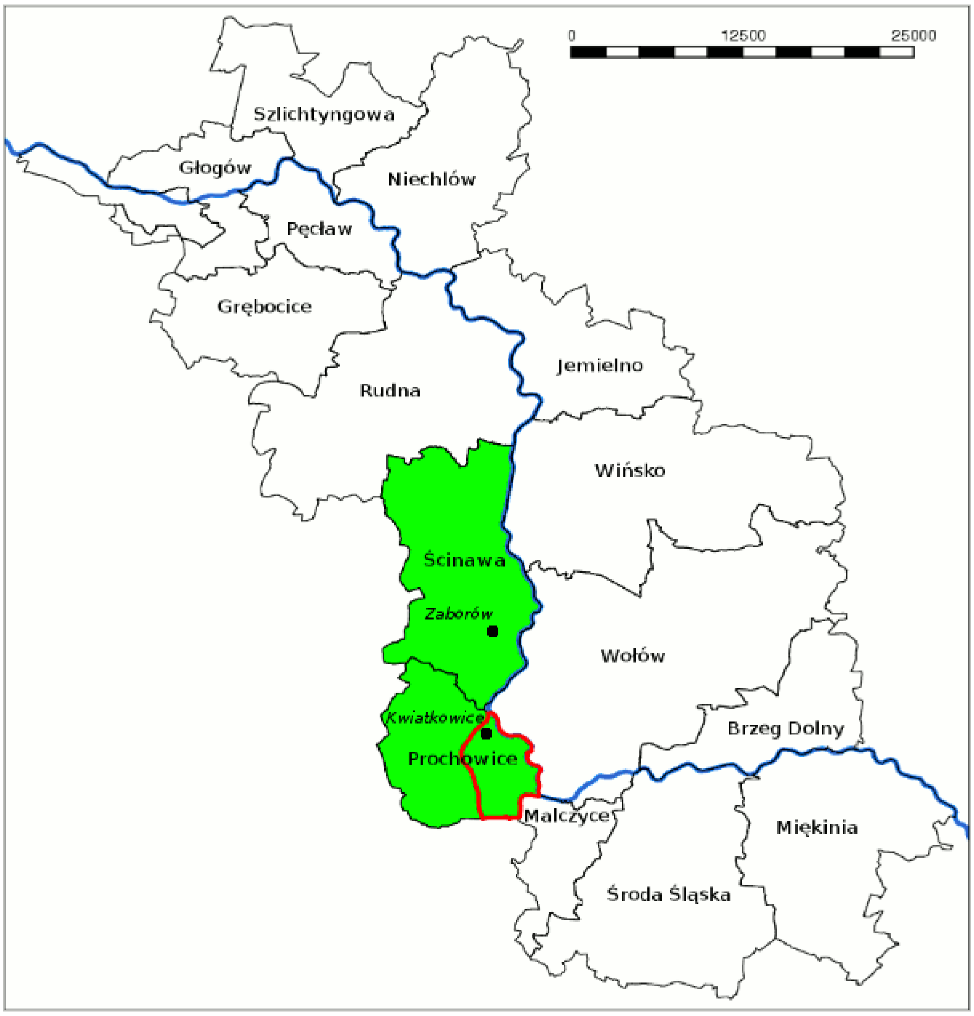

Fig. 5: A map of the Odra Valley case study area, showing the administrative division. Communes of special interest are *coloured* green. The red polygon indicates the area of the social aspects of land reclamation model proposition. Villages where workshops with local community took place are indicated.

A workshop with key domains (land use, water management, social welfare) experts was performed in Wroclaw on the 8th of June 2005 in order to elicit knowledge on main regional development determinants and to find solutions for problems pointed out by local communities. The problems identified during workshops with local communities were presented to experts. The task was to work in groups and find solutions for these problems. The other task was to perform SWOT analysis of regional development issues. Experts worked in groups, writing their comments on paper sheets. Then a public presentation of results and discussion took place.

As a result of meetings, interviews and workshops we elicited general knowledge on the following subjects:

- main relationships among key stakeholders
- important events (external shocks and internal stresses like political transformation and EU accession as well as endogenous changes in agriculture and demography) influencing social and economic situation in the region
- the state of the agriculture and main trends in its transformation, including social, legal and economic drivers of land use change in the case study region
- main problems of local communities as policy relevant issues suitable for modeling (like land reclamation system, flood control, social activity, unemployment and poverty)
- hypothesis about decision making rules among farmers in relation to land use

Social research – questionnaire survey

In cooperation with Joanna Stefanska from the University of Warsaw we have prepared questionnaires and conducted preliminary field research (50 questionnaire interviews in 2 villages) in order to check whether it is possible to find the exact structure of social networks (nodes and links) in the study area. The results are now being processed.

4.3.2 GIS

A GIS database framework for the Odra case study was created using Grass GIS 6.1 in UTM projection. Within this database vector maps of administrative divisions were created for 16 communes encompassing the case study area. Satellite imagery (Landsat TM) from 29.08.1990, 26.09.1991, 13.06.2000, 24.05.2001, and 23.06.2003 were collected for land use/cover change (LUCC) detection and visualization. From this Landsat imagery, the Odra river within the case study area was digitized for the production of maps.

Data collected by the Central Statistical Office regarding land use/land cover (LULC) structure and social phenomena were integrated with the vector maps of communes in the GIS database for spatial analysis and visualisation. Thirty-five thematic maps of analysed phenomena were created. In addition, GIS layers representing the extent of the flood of 1997 were created based on *Atlas obszarow zalewowowych Odry, WWF, 2000*. A thematic map of the flood extent in each commune was created.

An SRTM DEM (digital elevation model) was obtained and integrated into the case study database after re-projection and re-interpolation. Based on this DEM, a topographic analysis of the Odra case study area was performed for a generic topographic description of the case study.

LULC maps of the Odra case study were extracted from the Corine Land Cover 2000 (CLC 2000) provided by the European Space Agency and integrated into the GIS data base. The current LULC data was then analysed based on CLC 2000 as an addition to Central Statistical Office's LULC data.

Topographic maps of different scale were obtained to aid in the field work and the detection of land use change (1:25000 topographic maps from 1930-33, 1980-85 and 1993-94) and to extract detailed information about the area's topography (1:10000). The maps were registered into their original projection and subsequently transformed into the UTM projection and inserted into the existing GIS database of the Odra case study. Topographic maps on paper in the scale 1:25000 are currently being digitised in the area of Prochowice commune for an initial LULC analysis within the period from 1930's till early 1990's in this area. Moreover, elevation contour lines and points as well as watercourses are being digitised from 1:10000 topo maps for creating a detailed DTM (digital terrain model) in the area of Prochowice commune.

4.3.3 Literature

We have revised and gathered relevant scientific publications (about 140 references) on agent-based models of irrigation systems, decision making processes in agriculture, social networks and collective action issues. We have also collected governmental materials about the transformation of rural areas in Poland, which can serve as a source of knowledge about general trends in Polish agriculture.

4.3.4 Steps to be undertaken in the nearest future

As the result of our hitherto work we have general ideas about:

- social characteristics of the case study region
- main relationships among key stakeholders
- characteristics of present land use and land cover (LULC) as well as social, legal and economic drivers of LULC change
- possible modelling topics for fine-grained models (land reclamation issues and general land use issues)

Thus a decision about modelling topics should be made ultimately at the Wroclaw meeting on the 26-29th Sept.2005, so that we can know what detailed variables are necessary for the modelling team and what kind of knowledge representation and data format is expected from us. Then next steps in knowledge elicitation process (like interviews with farmers in order to find detailed decision rules) need to be undertaken respectively to the chosen modelling topic. We will also need to make the LUCC analysis based on collected data sources.

5 Schedule of meetings

The following schedule of project meetings was agreed at the kick-off meeting in March 2005. It has been extended to include a workshop on ontologies at the end of June 2005, hosted by the Macaulay Institute in Aberdeen. One of the major outcomes of this meeting was the need to have more time dedicated to discussing modelling issues across all modelling teams. The next project meeting in Wroclaw at the end of September was therefore extended by one day.

Date	Type	Venue (host)
June 2005	e-meeting	
26th - 30th September 2005	Site meeting	Wroclaw (WUT)
December 2005	e-meeting	
14th - 17th March 2006	Site meeting	Aberdeen (Macaulay)

June 2006	e-meeting	
5th-8th September 2006	Project meeting	Oxford (SEI)
December 2006	e-meeting	
20th-22nd February 2007	Project meeting	Vienna (IIASA)
June 2007	e-meeting	
25th-27th September 2007	Project meeting	Kassel (Kassel)
December 2007	e-meeting	
19th-22nd February 2008	Final project meeting	Manchester (CPM)

6 Amendments to the project work plan

The experiences gained from the modelling work during the first six months have led us to the conclusion that it is best to deviate from the proposed scheme of starting with prototype fine grain models and develop coarse grain models from them. Instead, the modelling teams will develop initial models at a scale natural to the case study in question and scale these up or down, respectively.

This adjustment of approach means that the following deliverables are changed as follows:

Del. no	Deliverable name	WP no.	Modified name	Re-allocated WP
4	Prototype of the descriptive, fine grain model for the Grampian region	3	Prototype of the initial model for the Grampian region	
5	Prototype of the descriptive, fine grain model for the Oder River valley	3	Prototype of the initial model for the Oder River valley	
7	Prototype of the descriptive, fine grain model for the Limpopo region	3	Prototype of the initial model for the Limpopo region	
9	Prototypes of the coarse grain models	3	Prototypes of generalised statistical mechanics models	4

On a more formal note, Deliverable No. 11 (Working paper on modelling and case studies in policy analysis process, lead: MMU) was missing from the work plan as described in the technical annex, p. 34ff. It is to be included as Deliverable 1.4, with the final report being re-numbered 1.5.

Since work in the first six months was delayed due to staff not being appointed in time, an extension of the project by four months will be needed[11]. The extension will cause the following changes in the work plan:

- Milestone 1.3 (international workshop) will be pushed back to month 38.
- All month 36 deliverables (13-16) will be pushed back to month 40.

---

[11] A four months extension requires that CPM will be able to allocate money from other sources to cover 2 months personnel costs; otherwise it will only be possible to extend the project by 2 months.