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

CAVES Periodic Activity Report

Project Month 18 • Project 012816: CAVES – Complexity, Agents, Volatility, Evidence and Scale

Duration: 2005-2008

Funded under the EU 6FP NEST programme.

Table of Contents

1 Executive Summary

The CAVES project is producing computational social models based on detailed case studies of social processes in three areas: the Grampian Regions of Scotland in the UK, the Odra River Valley in Poland and the Sekhukhune district of South Africa. All of the models are *agent-based* meaning that small, autonomous computer programs are implemented to represent the behaviour of individual stakeholders in these three regions. We know from past modelling experience that models with agents that interact with each other and influence each other exhibit complex outcomes. These outcomes are complex in the sense that (i) the overall behaviour of the model cannot be predicted by statistical means, (ii) knowledge of individual behaviour alone does not predict the model outcomes and (iii) they entail clusters of changes in behaviour and aggregate outcomes that look like the changes we see in asset prices in financial markets. For this reason, the CAVES project is not designed to produce forecasting models. It is designed to produce models that help stakeholders and, in particular, policy makers to clarify the concepts they use, the assumptions underlying their analyses and the expectations they form about the value and consequences of different policy measures. In the first 18 months of the CAVES project, substantial progress has been achieved in gathering evidence about the societies and problems of each of the regions studied and also in the development and application of tools to model the relevant aspects of those societies and to analyse their problems.

The project is predicated on the proposition that individuals tend to follow routines unless they have some reason to abandon or change a routine, individuals are socially embedded in the sense that they interact with and influence one another but they do not engage in widespread and mindless imitation. These are conditions that give rise to complexity. Societies might be more complex than physical systems because individuals change their behaviour and the institutional structures that guide and constrain that behaviour over time. The changes are frequently responses to critical events – especially the unpredictable and large changes that occur from time to time. A core question of the CAVES project is how *social networks* evolve and how and why they change both endogenously and in response to significant events. The three case studies of the CAVES project are identifying and documenting relevant social networks and the phenomena leading to changes in those networks.

It is clear in all three case studies that there is no single social network. Any individual can be part of each of several networks based on neighbourhood, family, work, hobbies, religious organisations or mutual assistance societies as well, no doubt, as many others. In the Sekhukhune district of South Africa, for example, individuals are networ-ked with family, with household clusters, with savings societies and with burial societies. In the Odra Valley, our partners from the University of Wroclaw and Wroclaw University of Technology identified networks based on neighbourhood and common or related activities. The identification of social networks in the Grampian region is at an earlier stage but the evidence so far has identified social networks based on the National Farmers Union and community involvement of various types. Acquaintances and friends in one of these networks might also populate other networks. Whilst we find such overlaps, it is not usually the case that the same individuals populate all social networks.

The CAVES project is unusual in basing formal specifications of social networks on detailed *evidence* from stakeholders. In the South African study, we are using interviews with householders and officials and also have access to a longstanding and highly detailed database from the RADAR project which is concerned with AIDS and gender violence. In the Grampian study, our colleague Lee-Ann Small is conducting extensive interviews with landowners, farmers and others affected by changing land use and is also drawing on longitudinal data gathered in several previous studies by the Macaulay Institute. MI also has access to, and is using, several substantial public data sources on land use and land ownership. The Odra case study is also relying on both extensive fieldwork with current stakeholders and previous studies. Every one of these studies is identifying whether social networks have been evolving as a response to social and political change or, indeed, drivers of social and political change or both in the sense that social and political institutions are co-evolving with the social networks.

The *modelling* effort proceeds in tandem with the case study developments. The purpose of the modelling effort is to determine and demonstrate different approaches to representing the social structures and processes being studied, to identify sources and elements of complexity and to develop procedures. A major task of the first half of the CAVES project has been to identify a range of means of formalising the case study data as agent based social simulation models. In line with the project design, one of the models (the Grampian model) will be implemented in a procedural environment previously developed by our partners in the Macaulay Institute while the other two will have strong declarative elements. Procedural programming is well suited to the implementation of algorithms that essentially impose the process of solution on a model. Declarative programming is rule based and is well suited to capturing the emergence of process in a model. In order to develop the declarative models, the modelling teams at both Manchester and Kassel are developing rule bases to implement agent behaviour in JESS, the Java Expert System Shell developed by the Sandia National Laboratory in the USA. We have integrated JESS with Repast, one of the standard agent based modelling environments comprising a set of Java libraries. Both approaches have been effective and models from all three modelling teams were presented at the World Congress on Social Simulation in Kyoto in August, 2006. All of these models are based on qualitative and numerical data provided by the corresponding case study teams: the Wroclaw partners for the Kassel model and the Stockholm Environment Institute for the Manchester model.

A departure that had not been anticipated in the project design was a formal approach to the specification of model structure. The proposal was concerned with formal representations of behavioural and social processes. However, our partners from the Macaulay Institute have been using formal *ontologies* which describe the structure of relationships within a program to describe the structure of social relations to be represented in our social simulation models. Two of the three modelling teams, Macaulay and Manchester, have developed these ontologies and are using them to inform and guide model design. The process is time consuming so that the remaining modelling team, Kassel, have avoided that step, thereby to provide a control benchmark for the value of the use of ontology. As a result, we have one team using the ontology development with a procedural modelling approach, one using ontology with a declarative element in their modelling approach and one not using ontology with the declarative modelling approach. That is the widest possible spread of combinations among three teams and, it is hoped, will give us a basis to evaluate ontologies and declarative modelling approaches.

At the half-way mark in the project, the models and case studies are not yet ready to be rolled out in a policy formation context. At the same time, it is important to consider the issues that the models will be used to address and the procedures to be used in addressing them. To this end, our partner from IIASA has produced a review of the literature on social *resilience* and related it to complex adaptive systems. A key point in this deliverable is how to approach an understanding of the ability of a society to absorb disturbance, i.e. reorganise in a process of ongoing change whilst continuing to function. In more traditional approaches, the key concepts were based on social equilibrium and the possibility of multiple equilibria. An essential aspect of social (as well as natural) complexity is the occurrence of change that is unpredictable in its timing, magnitude, duration and even the outcome once volatility has passed. The resilience literature entails the wholesale use of metaphor which might serve as a stimulus to thinking about the problems of complexity and attendant volatility but is either formalised in ways that depend on inappropriate concepts such as equilibrium or lack precision. The IIASA paper is drafted from the standpoint of scholars who accept and have participated in the development of resilience theory. It thus provides a benchmark and perhaps a target for the assessment of how the case study with modelling techniques developed in the CAVES project can lend precision and constraint to the analysis of social phenomena associated with a loose and frequently contentious concept of resilience and its complement, vulnerability.

A benchmark of a different kind is being provided by our partners at WUT. A core feature of the CAVES project is that it starts from evidence and problems as specified by stakeholders independently of any prior theoretical constraint. This is arguably a different approach from that pursued in the complexity and social simulation literature by physicists who adopt and possibly adapt formal techniques from physical science. This approach is being followed by WUT and will form a key element in our evaluation of the relative strengths and weaknesses of an evidence and problem driven approach.

In summary, a range of approaches and techniques and combinations thereof have been investigated and developed in the first half of the CAVES project. The modelling techniques have become settled and the case studies are providing the data required to test the ability of the models to identify social complexity and its consequences in a changing natural and political environment. The results to day continue to support the promise of agent based social simulation as a means of lending precision to the development of narrative scenarios to inform policy development in conditions of uncertainty.

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 Macaulay Institute in consultation with all partners.

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) has written 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 has reviewed 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 half of the CAVES project.

3.1 Centre for Policy Modelling

Project Month 1–6

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 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.

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.

---

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

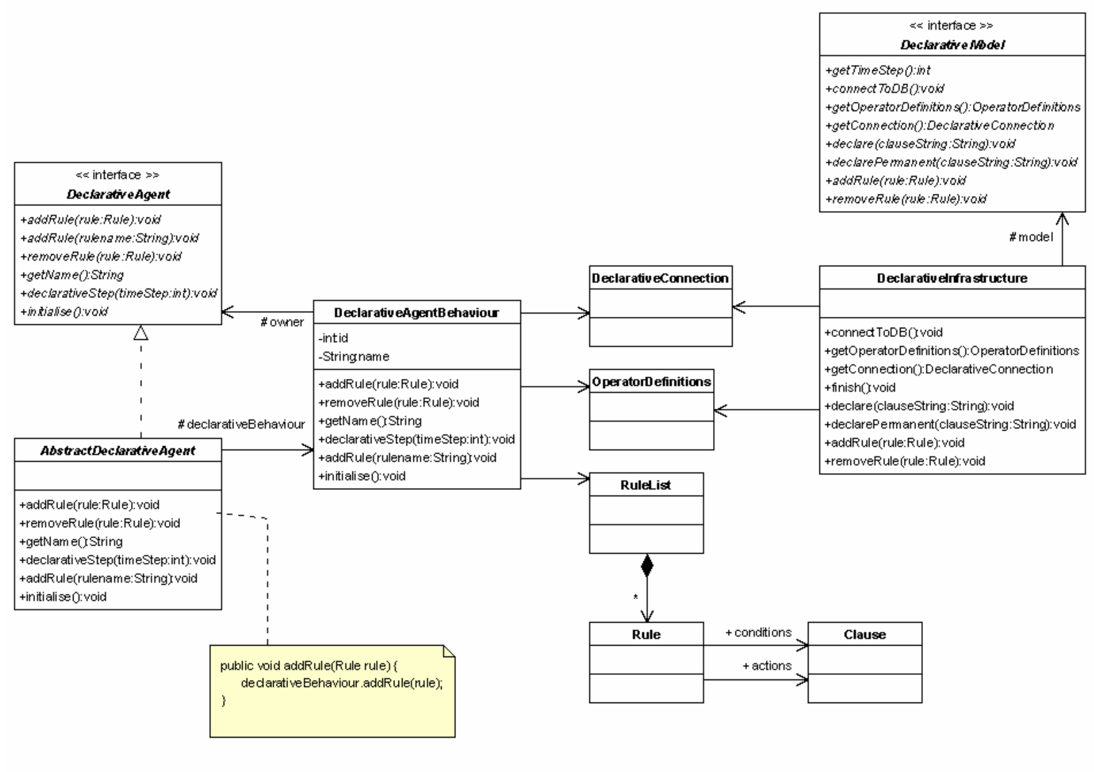

Figure 1: The key classes in the declarative package

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

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

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 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.

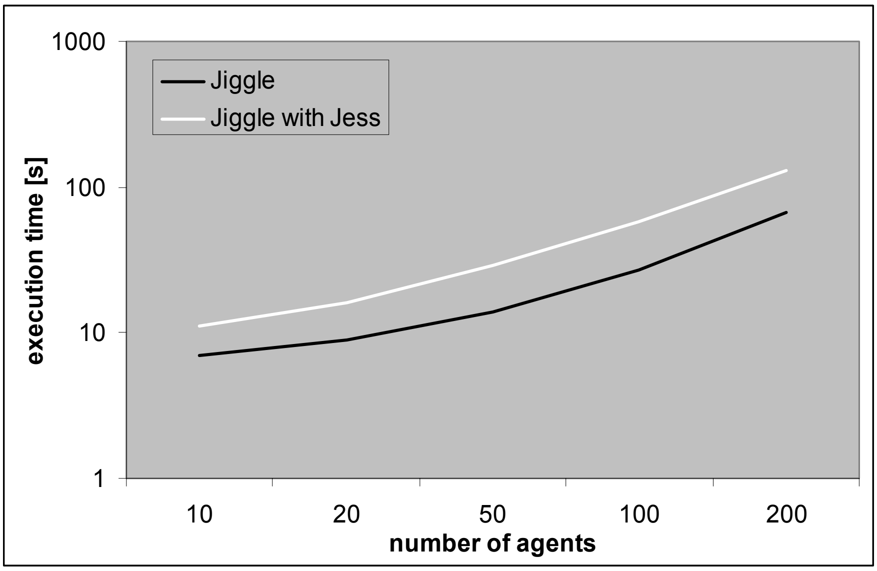

Figure 2: Comparing execution times with regard to the number of simulated agents.

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

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 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

[8] http://protege.stanford.edu/

land parcels.

Project Month 7-12

3.1.4 Prototype Model for the South African Case Study

Work has concentrated on the development of a prototype model for the South African case study. As suggested by the case study team (SEI) we focused on the social impacts of HIV/AIDS in villages in the Vhembe district, Limpopo Province, South Africa. Since field work for knowledge elicitation had to be done in parallel (see section 4.1), this prototype model is largely based on existing data from previous projects undertaken in the same region; namely the UNRAVEL project (Understanding resilient and vulnerable livelihoods in Malawi, South Africa and Zambia; Ziervogel et al. 2005) and FIVIMS-ZA (du Toit and Ziervogel 2004), a project assessing the feasibility of developing a Food Insecurity and Vulnerability Information and Mapping System (FIVIMS[9]) for South Africa. Within the framework of the latter, a pilot study was undertaken in the Sekhukhune district, whose data obtained from a detailed questionnaire was made available to our team. Unfortunately, it turned out that the data pertaining to social networks of individuals and/or households was either missing completely or aggregated in such a way that it was impossible for us to infer the necessary information. For the prototype model we thus had to make a number of assumptions, which we checked with our domain experts from the case study team.

In the case study area, HIV/AIDS is one of the major stressors for people's livelihoods, together with climate variability and food insecurity, leading to a high vulnerability. Most people in the area rely on state grants such as pensions or child/orphan grants and remittances from migrant workers for their subsistence, since agriculture alone is not sufficient. Death of the family member receiving the grant or sending money home can therefore have a devastating effect on a household, to the point of dissolution. Orphan children are usually accommodated by a household in the extended family. Other strategies for coping with stressors such as resource sharing or pooling of finances also rely on social networks in the community.

The model adopts the multi-layer network approach proposed by the Kassel team at the last project meeting in Wroclaw. So far, two network layers are considered, one on the level of individuals and one on the level of households. Individuals are represented as agents with a network of friends. Each individual is member of a household, with one of the household members acting as the household head. Households have a network of social neighbours with whom they interact. Since actual empirical data on the structure of these networks is yet missing, we assumed a small-world network. This assumption is supported by several other studies conducted in the region (e.g. Quinlan et al. 2005, Salomon et al. 2000).

These social neighbours are the basis for informal savings clubs, known as *stokvels*. Members of a stokvel pay a mutually agreed sum into the club every month. The cumulative savings of the group are then rotated to each member of the group on a regular basis. After everyone has had their turn in receiving the contributions, the group

[9] http://www.fivims.net/

may disband or start another cycle. Female household heads with higher literacy are usually the coordinators of these savings clubs (Verhoef 2001). We model this by introducing the role of 'innovators' for a certain proportion of agents. Innovators are able to initiate a savings club by inviting other agents and run the club after its formation.

To cover all agricultural aspects we obtained an agent-based cropping model from SEI (Bharwani et al. 2005), which models crop choices dependent on climate experiences and weather forecasts, the growth of the chosen crop, harvest and subsequent market transactions. We adapted this model to a package to be used as a component in other agent-based models. This included debugging the code and solving some of the implementation problems already identified by the SEI team. The imperatively implemented decision process was extracted and transformed into rules, so that the agents' decisions about which crop to plant when is now modelled declaratively.

The prototype model is implemented in Java/Repast, using the Repast scheduler, graphical user interface and network library. The declarative component integrates JESS (see below). Figure 3 shows a snapshot of a simulation run with the social networks of the households in the top right corner. The chart on the bottom left plots the number of savings clubs currently in existence.

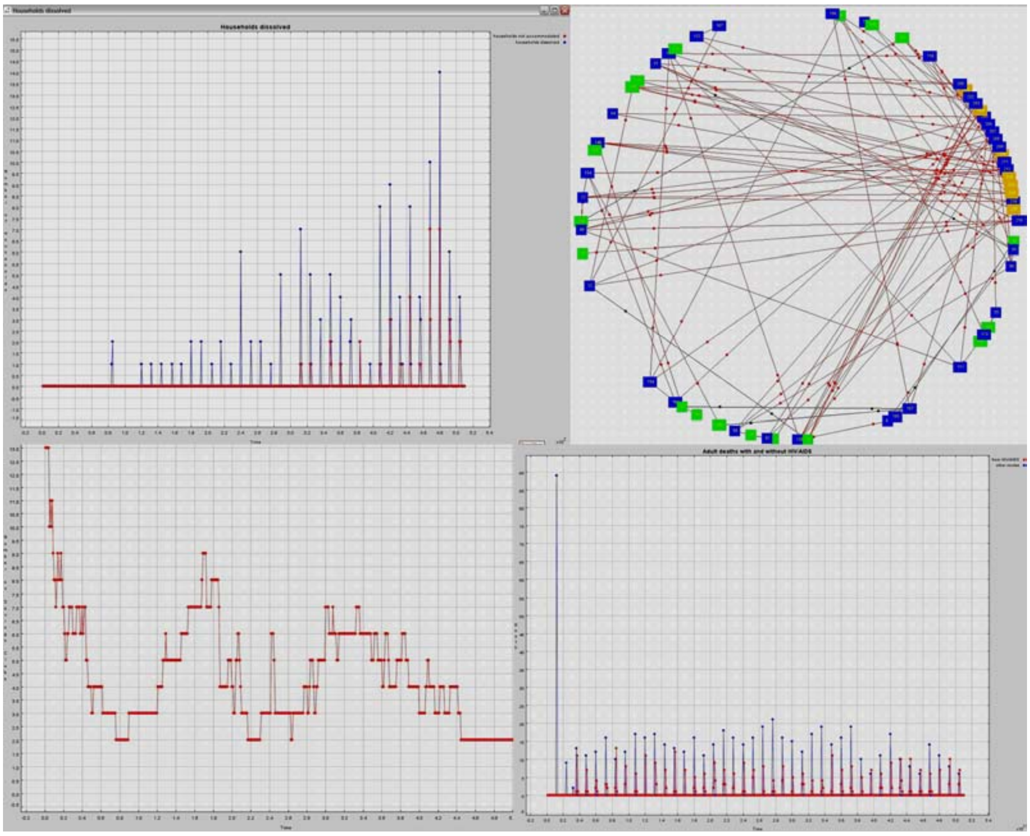

Figure 3: Snapshot of a simulation run.

3.1.5 Modelling Tools

JESS/Java integration

Another focus of work has been JESS[10] and its integration with Java/Repast. This combination was identified during the project initiation phase as the most promising candidate to integrate declarative features with agent-based social simulation software.

To investigate this further we adopted a model of land use and migration developed at the CPM before CAVES, which also integrates the modified crop model. Our first approach has been to let each agent have its own rule interpreter (engine). This allows the rules describing the agent's behaviour to be specified in the respective agent class. Global facts like the state of the agents' environment, however, have to be copied for each engine. Because of the high memory requirements of this approach we have been researching the possibilities of sharing one rule engine amongst all agents.

Several options exist, with different advantages and disadvantages:

- Structuring the joint rule base into modules for each agent. This allows agents to share facts while using rules specific to a particular agent. The focus of the rule engine has to be set to the appropriate module before running the engine; this can easily be done in an agent's `step()` method. So far, the module approach has proven successful for a simplified land use model with 3 rules per agent (module).

  The advantage of this approach is its conceptual clarity: agents "own" their rules and are solely responsible for their execution. On the other hand, a large part of the rule base is made up of more or less identical rules.

- Defining rules per agent type and collecting them in the model class. This ensures there is only one copy of each rule in the rulebase. Agents are best declared as shadow facts so that rules can refer to individual agents easily. In every run of the engine, each rule will fire once per matching agent fact. Thus the engine has to be run only once per time-step to update all agents.

  This approach keeps the rulebase as small as possible. Its main disadvantage is the gap between conceptual model (rules describe an agent's behaviour and therefore belong to an agent) and implementation (rules belong to the model and are shared between agents). This could probably be overcome by developing a framework that allows a modeller to specify agent-specific rules and then transforms them to the model-specific approach.

- Differentiating rules by adding an identifier (agent ID) to their name, e.g. `plant-maize-rule-1`, `plant-maize-rule-2`, etc. This results in each agent having its own set of rules in a joint rule base. The engine is still run only once per time step.

  This approach is a mix between the first two: It allows agents to "own" their rules again but the execution of rules is initiated from the model. Depending on the number of agents and rules per agents, the rulebase can become very large. This

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

approach is the least elegant of the options.

As a control and to provide a parallel development path in case the single-engine approach proves unsatisfactory, we have also been developing a trading model with one engine for each agent. The number of agents that can be implemented with suitable values for heap size is (so far) acceptable for a fine grain model – on the order of 500 – 700 agents.

GIS interface

We have also started to investigate GIS interfaces to models. Although Repast claims to have integrated GIS support with version 3[11], this is so far insufficient. The library for vector data (`anl.Repast.gis`) has two major drawbacks: it focuses on vendor-specific software (Esri's ArcView/ArcInfo) and it represents geographic shapes as agents instead of incorporating them into a space object.

Due to a lack of relevant GIS data for the South African case study, the development of a package to incorporate GIS data as a spatial component into a model has been suspended.

Project Month 13–18

Work at the CPM has concentrated on further developing the prototype model for the South African case study. This included obtaining and evaluating new evidence data from the case study area and incorporating the elicited knowledge into the model. Other foci of work have been investigating declarative modelling techniques and researching measures for social network analysis. In addition, work related to the CAVES project has been presented at several conferences and workshops (Norling et al. 2006, Werth et al. 2006, Werth 2006, Alam 2006).

3.1.6 Model for the South African Case Study

The prototype model for the South African case study was delivered at the 12-month meeting in Aberdeen. This model was presented at WCSS'06 (Alam et al. 2006). With the basic framework thus implemented, several additional social processes like mutual help amongst villagers or migrating for work have been incorporated, whereas some of the already existent processes could be fine-tuned in close collaboration with our domain experts.

The declarative component, which focussed on the agricultural aspects, has been extended considerably. The decision making of the household heads concerning the distribution of resources amongst household members and the joining of social clubs such as stokvels or burial societies has been extracted into rules and implemented in Jess. For this we adopted the approach which had been identified in the last report as being the most efficient. Agents are represented as shadow facts, and rules are collected in the

[11] See http://Repast.sourceforge.net/index.html

model class ensuring that there is only one copy of a rule in the joint rule base. Since the number of household heads is fairly small at any time during a simulation run compared to the total number of agents, there was no significant impact on the runtime of a simulation.

Other social processes are dealt with by all agents individually. These are e.g. decisions with respect to finding a job, borrowing food or marrying and starting a new household. To be able to model the individual behaviour declaratively, all (adult) agents have to be represented as shadow facts in Jess' working memory. While this has slowed down the execution speed considerably, simulation runs with several thousands agents are still feasible.

Regarding mutual help, which is the crux of social networks in the system, we have focused on the borrowing and lending of food among agents. The model now incorporates a more realistic abstraction of the food-intake and its subsequent impact on the health status of an individual agent. This is further differentiated for individuals according to their age and sex. There are rules for household heads to decide how much to spend for food for the household members. The amount spent depends upon the number of children, adults and senior people in the household, whether the household has extra health expenditures, and whether the household is a member of a burial club or not.

A detailed description of the model structure and the modelled social processes will soon be available online as part of the Centre for Policy Modelling discussion papers series. We are also planning to submit a paper about the model in the Journal of Artificial Societies and Social Simulation (JASSS).

3.1.7 Data Analysis

In June the CPM modelling team attended a meeting at SEI Oxford with our case study partners and two external domain experts from South Africa (see section 4.1.10). We were able to agree on a collaboration with RADAR (Rural AIDS and Development Action Research Programme[12]), giving us access to longitudinal data from the case study area. Specifically, we obtained data pertaining to several surveys, which were conducted in the context of the IMAGE study (Intervention with Microfinance for AIDS & Gender Equity[13]). This research initiative aims at targeting the prevention of HIV/AIDS and gender-based violence through alleviating poverty via a microfinance-based lending scheme and is based in 8 villages in the Sekhukhuneland District.

So far, we were able to extract information relevant to our model from two of the six surveys. These include detailed information about household structure, age distribution, income and migrants, which have been used to enhance the model's validity. Example findings are that household size varies from 1 to 19, with a mean of 6.98. The size distribution can be approximated by a Normal distribution with a mean of 7 and standard deviation of 3. On average, households are made up of the household head, 3.5 children, 1.7 grandchildren, 0.5 spouses of the head (i.e. half of the household heads are married) and 0.59 other relations, ranging from nieces/nephews and siblings of the head to relations of the spouse or a child in-law. The majority of grandchildren (88%) are

[12] http://www.wits.ac.za/radar/Home.htm

[13] http://www.wits.ac.za/radar/IMAGE_study.htm

children of a daughter – this probably means that if an unmarried young woman, who is still living at her parents', has a child, this child stays with her rather than the father's family. Nearly all of the household members (apart from the head and spouse) are unmarried. In the rare cases that married children are still living with their parents they are usually living with the man's family.

There are a total of 830 households in the survey, 394 (47.5%) with a male head and 436 (52.5%) with a female head. Almost all male heads are married (96.4%), whereas female heads are mostly widowed (58%), or have either never married (22.4%) or divorced/separated (16.2%). Married female heads are very rare (3.4%). This leads to the conclusion that if a couple are married, the man is the official household head. Still, nearly a third of male household heads go away for work (27.4% didn't sleep at home in the last month) whereas practically all female household heads (95.6%) stay at home.

The extracted data has been especially useful for the model setup, in constructing the initial social networks. What is still missing, though, is more information about the individuals' behaviour and decision processes. We hope to gain more insight into these through further data analysis and continued collaboration with our domain experts as well as a field trip to South Africa in January.

3.1.8 Towards Network Analysis and Statistical Signatures

We have incorporated methods to export the generated social networks data in two formats, for the commonly used network analysis tools Pajek and UCINET, respectively. As the social processes in the model generate overlapping social networks, both at the individual and the household level, the simulation is able to export data in the form of 2-mode affiliation networks. The social network analysis measures can be applied at different time-intervals of the simulation.

With regard to dynamically evolving networks, we are currently working on finding 'suitable' analysing techniques and macro-level indicators. This work is in a preliminary stage; albeit some suggestions have already been presented at the Forum for Social Network Analysis in Leeds (Alam and Meyer 2006).

How evidence-based networks can contribute towards complexity science as a test for complex social systems, is certainly an open question. A strong collaboration among two modelling teams in CAVES, University of Kassel and Centre for Policy Modelling, is underway to address this issue (Alam et al.; forthcoming). The idea is to abstract from the case study models within the CAVES framework and contribute towards helping the stakeholders in dealing with complex policy making decisions and adding further precision to their understanding.

3.1.9 Modelling Techniques

Another focus of work has been investigating ways of using Jess for declarative modelling within Java/Repast models. As presented in the last progress report, there are several options, from using one rule engine per agent to sharing not only the rule engine but the complete rule base amongst all agents within a model. During the last six months our experiences with trying different approaches in several models have let us to decide on the latter approach as the most efficient and feasible one.

Moreover, the proportion of Jess versus Java for the implementation of a model has been explored. Knowing the best way of employing Jess, a modeller still has to decide how much of the model is to be implemented declaratively and how exactly the declarative part is to be translated into facts and rules. With the South African prototype model we started with a procedural implementation and added declarative components little by little, keeping the major part in Java.

Since this let to constraints regarding the formulation of rules, we are currently in the process of re-implementing the model; this time starting from a declarative approach. It is already clear that this will result in a much more extensive use of Jess, both in terms of the rule base and the working memory (number of facts). A comparison with the more procedurally implemented model will hopefully show the benefits and costs of each approach. It is planned to submit a respective paper to the Third International Model-to-Model Workshop, Marseille, France, March 2007.

3.1.10 Internet Portal

The CAVES internet portal (Deliverable No. 2) was enhanced by a Wiki[14] component. This is planned to be used for the knowledge base, glossary and discussions about terms. So far, a first version of the glossary is available As agreed at the project meeting in Aberdeen the glossary contains links to relevant Wikipedia entries for terms without a project-specific definition. The Wiki component can be found at http://caves.cfpm.org/wiki/pmwiki.php.

3.1.11 Future Work

- Adapt and expand the South Africa model. This includes incorporating biophysical models to determine water availability and land use as core elements of the environment in which social vulnerability dynamics emerge.
- Continue work on exploring and facilitating ways to use Jess in combination with Java/Repast for declarative modelling.
- Add logging functionality to Jess/Repast so that the trace of a simulation run is stored in a data base. This will enable modellers and/or stakeholders to use the trace not only for validation purposes but also for the analysis of simulation results.
- Continue work on social networks and statistical signatures.

3.1.12 References

Alam, S.J. (2006): Towards better socioeconomic policies in the Sub-Saharan region. Poster presentation at the ESRC Research Methods Festival, Oxford, UK, July 2006.

Alam, S.J. and Meyer, R. (2006): Finding Suitable Analysis Techniques for Agent-Based Networks Generated from Social Processes. In: Proceedings of Social Network

[14] See e.g. http://en.wikipedia.org/wiki/Wiki

Analysis: Second Forum on Advances and Empirical Applications, Leeds, UK, July 2006

Alam, S.J., Meyer, R. and Ziervogel, G. (2006): Modelling Socio-Economic Impact of HIV/AIDS in a South African Village. In: Proceedings of the First World Congress on Social Simulation (WCSS'06), Kyoto, Japan, August 2006

Alam, S.J., Ernst, A., Krebs, F., Meyer, R., Moss, S. and Zehnpfund, C. (forthcoming): Evidence-Based Social Networks: Testing for Complexity. Submitted to Multi-Agents for Modelling Complex Systems (MA4CS'06), satellite workshop of the Second European Conference on Complex Systems (ECCS '06), Oxford, UK

Bharwani, S., Bithell, M., Downing, T.E., New, M., Washington, R and Ziervogel, G. (2005): Multi-agent modelling of climate outlooks and food security on a community garden scheme in Limpopo, South Africa. *Philosophical Transactions of the Royal Society B*, 360 (2005), pp. 2183–2194

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

du Toit, A. and Ziervogel, G. (2004): Vulnerability and Food Insecurity – Background concepts for informing the development of a national FIVIMS for South Africa.. <http://www.agis.agric.za/agisweb/FIVIMS_ZA>

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

Norling, E., Edmonds, B. and Alam, S.J. (2006): Complexity in Socially-Inspired Simulation. Poster presentation at the EPSRC NANIA Project Summer Meeting, Ambleside, UK, August 2006

Quinlan, T., Ziervogel, G., and O'Brien (2005): Assessing Vulnerability in the Context of Multiple Stressors – The Southern Africa Vulnerability Initiative (SAVI), IFPRI Conference, Durban.

Salomon, J.A., Gakidou, E.E, Murray, C.J.L. (2000): Methods for Modeling the HIV/AIDS Epidemic in Sub-Saharan Africa. Geneva, World Health Organization, GPE Discussion Paper Series: No. 3, EIP/GPE/EBD <https://www.who.int/healthinfo/paper03.pdf>

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

Verhoef, G. (2001): Informal Financial Service Institutions for Survival: African Women and Stokvels in Urban South Africa, 1930-1998. Enterprise & Society 2, pp. 259-296.

Werth, B., Moss, S., Ziervogel, G., Downing, T.E. (2006): Modelling Migration in the Sahel: an Alternative to Cost-Benefit Analysis. In: Proceedings of the First World Congress on Social Simulation (WCSS'06), Kyoto, Japan, August 2006

Werth, B. (2006): Harnessing Uncertainty in IT Projects of Financial Institutions. Poster presentation at the Eighth European Agent Systems Summer School (EASSS'06), Annecy, France, July 2006

Ziervogel, G. et al. (2005): UNRAVEL – Understanding resilient and vulnerable livelihoods in Malawi, South Africa and Zambia. Final Research Report, September 2005

3.2 Universität Kassel

Project Month 1–6

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 first 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.

Project Month 7–12

3.2.3 The Context of Our Model

In numerous real-world situations, people are confronted with tasks that they are unable to fulfil alone. Often, such tasks are characterised by the necessity to include a number of different expertises to their accomplishment. Consequently, people organise themselves into networks aimed at the completion of some specific task. Examples of such situations are to be found in virtually any domain, such as science, economy, or in the context of managing and maintaining natural resources.

The Polish case study (with input provided by the Wroclaw Institute of Technology and Wroclaw University) focuses on those parts of the Odra river region that are at risk of regular flooding due to neglected or damaged dikes and the lack of maintenance of an old land reclamation system and also more generally on land use in the Odra river region. Social mobilisation or collective action by the individual farmers is required to maintain or re-establish the system of channels, ditches and dikes of the land reclamation system. Between the farmers, acquaintance or friendship links exist. When looking for collaborators to accomplish a maintenance related task however, the friendship network may serve as a starting point to build up a collaborator network, but the friendship network may not suffice to get all needed expertises together. By word of mouth, additional persons in the collaborator network (i.e. collaborators of collaborators) with the necessary expertise are sought, until the task can be solved. Such existing networks tend to be used again and again, thus leading to cliques of collaborators with complementtary expertises.

In a more abstract way, situations like those just described can be characterised by the following features: They include multiple social networks representing multiple social contexts that interact, like friends vs. collaborators. People show goal or task-directed behaviour and use the networks at their disposition to fulfil their tasks. The conditions of the emergence of such multiple networks, their long term evolution, characteristics, interaction and their dynamics over time is of theoretical as well as practical interest to social science as well to complexity science. We will report on this dynamics by contrasting different social networks resulting from an agent-based model of task-oriented behaviour in a collective action situation. Specific measures have been designed to analyse the behavioural and structural efficiency of the networks and knowledge that is accumulated by the agents over time when solving tasks of varying difficulty.

3.2.4 Basic Modelling Concepts

In order to model the above mentioned situation characteristics, core features of the case study are abstracted. We follow a rather strict distinction between physical environment and social environment of the agents. 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, the simulation of the agents' 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 considered here 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. 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. In the model version presented in this paper, the focus is on the development of the social networks and the actions related to the natural or physical environment have been reduced to abstract tasks.

The agents' social environment is modelled as networks. An agent may be seen as a node in different social network contexts. Technically, an agent has slots that are nodes representing potential or actual social roles in different networks, so the networks actually reside in the agents' memory. Unlike in other network modelling approaches, agents do actively perceive their social environment and are enabled to act in their social network. In the model considered here, an agent has two semantically different nodes: One in a friendship or acquaintances network and one in an advisor or collaborator network.

The friendship network can be initialised with empirical data or in a more abstract way with an assumed small-world topology. A collaborator network does not exist initially. Once a task is assigned to an agent, it polls its social friendship network for expertise needed to accomplish the specific task additionally to its own. The search is started in the direct social neighbourhood of the agent. If the collected expertise provided by the network neighbour has been successfully applied, the agent builds up an edge to the respective node in the collaborator network. Next time the agent would first poll the collaborator context when looking for a new strategy. If the agent cannot find all the necessary expertise in the directly neighbouring links of the collaborator network it will pursue the search in the neighbourhood of collaborators, i.e. collaborators of collaborators to find additional expertises.

In the following section, a description of the agent architecture that uses the described basic concepts will be given.

3.2.5 The SONATA Model

The SONATA model (Social Networks of Abstract Task oriented Agents) has been realised in the Repast agent programming framework[15]. In order to describe the proposed agent architecture we follow the separation of the agent's functional components: perception, action repertoire and cognitive unit.

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, the locations of other agents, or in the more abstract version presented here, information about tasks and their accomplishment. 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 collaborator network). Perception is locally bounded, so no agent within the network has a global, bird's eye view of the whole network.

The action an agent may execute in its physical environment is to solve a task that has been assigned to it. To do so, it has to complement its own expertise by other expertises needed by looking for collaborators accordingly. Additional actions in the agent's social environment are network-related modifications like strengthening or weakening of outgoing and/or incoming edges, the establishment of new edges in already established networks.

The simulated social environment consists of two network layers. The friendship network the model starts with a pre-generated and stable small-world network with a given average node degree resulting from rewiring of a regular net according to the algorithm by Watts and Strogatz (1998). This network layer remains fixed over the whole simulation run. The second network layer is the collaborator network that builds up during the agent's search for supporters with specific expertises after being assigned a task. Thus, it is actively constructed by connecting to other agents that have already provided the leading agent with useful information, following the algorithm described below. In this layer, unused edges slowly decay in strength and disappear once their weight becomes zero.

A task object is represented by number of different kinds of expertise (know-how, expert knowledge) that is required to perform the task. Tasks are randomly assigned to agents and have a fixed difficulty which results from the expertise necessary to solve them. Expertises are evenly distributed among the agents. Each time step, one agent is assigned with a task for which he needs the expertise of other agents. It will utilise its social environment to compile the required expertise to accomplish the assigned task. An agent first polls its collaborator network to get help from agents that have previously been helpful. If it cannot find enough collaborators among its direct ties, it is able to contact direct collaborators of its collaborators. It will build up edges in the collaborator network to these agents if they supply it with the necessary expertise. Only in the event that polling the collaborator network does not yield the necessary expertise, the agent will use its friendship network. If an expertise looked for can be got from a network neighbour, the agent will build a new network edge in the collaborator context to the supplier of the expertise.

[15] http://repast.sourceforge.net

3.2.6 Discussion

We will compare scenarios where only the initiator of the task builds up arcs to his collaborators (scenario without pairwise linking) with scenarios where all the agents that took part in the task build up arcs to every one of the participating agents (i.e. with pairwise linking). In both scenarios the agents have a maximum in- and out-degree, i.e., they are able to build or receive a limited number of arcs. Special attention will be given to the behavioural efficiency in solving tasks and the structural efficiency (i.e. number of links that are built up). All analyses of the networks generated by the Repast model have been done with the Pajek network analysis tool (de Nooy, Mrvar, & Batagelj, 2004) and R, a free software environment for statistical computing and graphics (http://www.r-project.org/), with methods also discussed by Newman (2003) and Wasserman and Faust (1994).

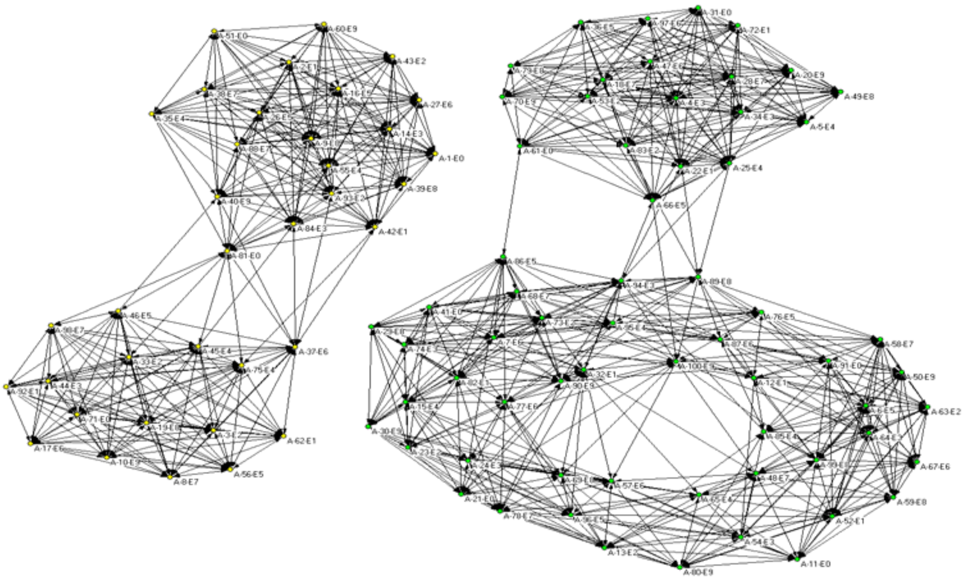

Figure 4: Collaborator network with 100 agents (‘A-1’-‘A-100’) with 10 different expertises (‘E-1’-‘E-10’) and a maximum in- and out-degree of 9, after 10,000 time steps. At this point of the simulation agents with different expertises start to gather into (task oriented) cliques. This network has a clustering coefficient of 0.68.

All the networks discussed here have been produced with the following model parameters: There are 100 agents. Their degree in the (static) friendship network is set to 20. There are 10 expertises needed to solve a task. Accordingly, the maximum degree for the collaborators is set to 9, relating to the number of additional expertises (beside the one the agent possesses). Every time step, 1% of the agents are randomly assigned a task. All agents are cooperative in the sense that they do not turn down a request for joining a task solving group (except they have reached the maximum of in- or outlinks). Links decay over time and disappear after 150 time steps, unless noted otherwise. The simulation stops after 100,000 time steps.

The results of the SONATA model show how a forgetting rate higher than the rate of new tasks coming in causes established links to disappear, so that the collaboration network has to be built repeatedly. The structure of the network thus never stabilises, and efficient cliques never emerge. Stability of network links also depend crucially on the way of linking: If, after having completed a task successfully, all participating agents link to each other in both directions (pairwise linking), stable structures arise that can be used again as soon as the next task is assigned to one of the cluster's members. This can be interpreted as groups remembering the good work they did together and whom they did it with. These pairwise linked networks thus accumulate with each task completed a maximum degree of knowledge relating to possible future collaborators. The knowledge is distributed evenly among collaborators and does not reside only in the agent the task was originally assigned to (no pairwise linking).

The higher efficiency of pairwise linked networks is reflected in a higher degree of tasks successfully completed and a higher degree of connectivity, but it has one drawback. Long-term evolution of such networks shows a segregation of successful cliques over time. While this may be well adapted to the task structure used here (with a constant amount of 10 expertises needed), this system may break down if there are substantial fluctuations in the quality of those tasks. Since the tasks assigned are the abstraction of the problems posed by a natural environment, this may be an important consideration. It will be investigated how shocks on the system, e.g. by changing the task structure, affect the networks with regard to their structure and performance. What does the system need to adapt to new situations? How long does it take to stabilise again, if ever?

Project Month 13–18

3.2.7 Organisational Issues

Claudia Zehnpfund left the Kassel research group for a job with an insurance company at the end of May 2006. A first job announcement to replace her was issued in May, but no qualified person could be found. A second announcement has been issued in the meantime.

At the 1st World Congress on Social Simulation (WCSS) in Kyoto, Japan, the Kassel CAVES group gave a paper titled "Dynamics of task oriented agent behaviour in multiple layer social networks" to present the SONATA model (Ernst, Krebs & Zehnpfund, 2006).

3.2.8 Additional Case Study Meeting

On June, 8th -9th, a meeting with the Polish CAVES project partners was held in Kassel. A central point was the production of useful, yet realistic data on social networks, behavioural options and rules of the farmers in the region and hints to the cost-benefit-structure of farming in general as well as land reclamation with regard to floods and droughts, i.e. the shocks on the natural system.

Storylines as an abstraction of the interviews carried through in the Odra region have meanwhile been abstracted by the Polish case study group. They will serve as a basis for enhancing the SoNARe model with more plausible behavioural rules.

The upcoming biophysical model will give more insight in the costs and benefits of farming and land reclamation under certain climatic conditions. It is suspected that land reclamation possesses, under a wide range of conditions, the structure of a social dilemma, and land reclamation therefore represents a collective action (Olson, 1965). If so, existing theoretical insights can be used to investigate this issue further. This also directly points to some leverage point of policy interventions, of which two have been investigated using scenarios produced by the SoNARe model (see below in the result section).

As a result of the meeting, a list of variables to be looked at was compiled. It contains also a classification (A/B/C) of relative importance of these variables to the project and consequently a priority ranking to include them.

Beside these issues, the concept of resilience was discussed with relation to the Odra case study. It seems that up to now, not enough quantitative time series data have been compiled to apply classical measures of resilience. However, central (social) indicators, such as trust being an epiphenomenon of peoples' actions and their networks, could be monitored closely in the case study as well as in the model scenario runs to give hints at upcoming regime shifts.

The technical coupling of the biophysical and the SoNARe (see below) models was also discussed and last points relating to the interface of the two modules were clarified. Skype was proposed as a quick way of communication between the two groups.

3.2.9 Modelling

As a successor to the SONATA (SOcial Networks of Abstract Task-Oriented Agents) model, we developed the SoNARe (Social Networks of Agents' Reclamation of land) model. It is much more concrete in terms of phenomena observed in the Odra case study and allows for the interfacing with the Odra biophysical model developed by the Wroclaw University of Technology group. The agents of the model are situated along an abstracted canal in an upstream-downstream asymmetric dependency. Maintenance of the land reclamation system enables to overcome floods and droughts without loss of harvest, but it requires a collective effort from upstream neighbours in the first place. The model will be described in the next section, followed by some scenario results and their discussion.

Technically, the SoNARe model is a hybrid model. It uses productions rules implemented in JESS and its reasoning engine to represent the cognitive control structure and decision making of the agents. Rules operate on mixed symbolic and numeric structures. Apart from that, large parts of the code are written in Java and the Repast library functions respectively to provide efficient execution of network functionality etc.

The Context of the Model: The Odra Case

The SoNARe model attempts a useful and plausible abstraction of key features of the CAVES Odra case study. Among these are:

- Environmental shocks/extreme weather conditions. A simulated flooding or drought lead to a loss of crop yield, whereas flooding is more severe that drought.
- Maintenance of the land reclamation system (LRS). Following assumptions already made in the SONATA model, this is regarded as a collective task that requires social mobilisation of the participants. The participants are determined by the location of their land parcels along a ditch or a communicating ditch system. Farmers can decide either
    - to participate in the LRS, i.e. maintain the LRS locally on their respective land parcel and thus increase the level of protection against environmental shocks, or
    - to neglect the LRS, leading to degradation and subsequently to a decreased level of protection against environmental shocks.
- Asymmetrical dependency between the agents. Depending on location of the farmer's land parcels along a channel (upstream vs. downstream), the farmer's influence on the functioning of the LRS on his parcel varies: All the upstream LRS sections have to be maintained in order to provide beneficial effects on an upstream land parcel. This dependency entails a social dilemma structure, in turn giving incentives to free riding (i.e. not providing LRS maintenance on one's own parcel when being the first after a row of co-operators). It is expected that it is this social dilemma structure that hinders, and in some cases prohibits the installation of a functioning LRS.

Model Setup

In this section, the model is described as an abstraction of the case study's environment and of its agents. The abstract environment will in a second step be replaced by a more sophisticated spatially explicit biophysical model (developed by WUT).

Environment

- Weather conditions. They can take the three (parameterised) distinct states normal, drought, flooding. They are set yearly and change through a randomly generated weather sequence over the simulation.
- Land parcels. Each agent owns one land parcel. The land parcels are positioned along a channel, and thus define an upstream-downstream neighbouring relationship between their owners.

  The land parcel is parameterised by the level of the quality of the local LRS, which entails a certain protection level against extreme weather. If the LRS is maintained on the respective land parcel then the LRS condition will increase. 100% of LRS condition is reached after a defined period of continuous maintenance (presently 36 months). If the LRS is not maintained then the LRS condition will decrease gradually after a defined degradation off-set period (presently 12 months).

The effective level of protection against flooding/drought of a land parcel is determined by finding the worst LRS condition on land parcels upstream. If all the land parcels upstream are being well maintained then the land parcel still profits somewhat (50%) from the protection upstream even without needing a functioning LRS itself (= free riding). If there are any non-maintained land parcels, the LRS condition of a parcel is set to the minimum of the upstream parcels.

- Crop yield. At month 5 agents plant crop on their land parcel and harvest in month 10. The resulting crop yield depends on the simulated weather conditions which is determined using the formula totalAreaOfLP*(1.0-(1.0-totalLRSCond)*propMaxLoss), where
    - propMaxLoss = 0.0 for years with normal weather,
    - propMaxLoss = 0.6 for years with extremely dry weather,
    - propMaxLoss = 1.0 for years with extremely wet weather.

Agents

- Different LRS maintenance rules. There are two rules, one not to maintain the LRS ("rule 0") and one to maintain it ("rule 4").
- Social networks and social influence. In certain intervals, agents rate their rules by looking at neighbouring agents in the acquaintances network and compare their (recent) economic performance regarding to their harvests. They thus build up a new information network from which they let themselves influence regarding their LRS strategy. The rating of the rule used by the agent is compared with the (subjective, network-local) ratings of known alternative rules (in the current version max. one other rule). Agents switch to a different rule if the difference between the ratings is above a certain threshold value (this threshold represents the degree an agent takes other agents behaviour into consideration).
- Economic balance. It is composed of the investment in planting (at month 5 of each year), the investment in LRS maintenance (each month), the gain from harvest (at month 10), the compensation payments in case of flooding or drought (at month 10) and the allowances given if an agent chooses to start maintaining the LRS. If the balance gets too low, agents stop maintaining their LRS. If the balance does not cover the planting costs, agents stop working on their land parcel forever, i.e. they leave the game.

Simulation Results

This section gives an overview of the first simulation results that were produced with the described model. We start off with some rather trivial parameterisations of the model that are nevertheless realistic and relevant. These scenarios are presented in a descriptive way. A more detailed look will be taken at three very similar scenarios that lead to distinctly different dynamics.

Sparsely linked acquaintances network

In all simulations the mentioned acquaintances network is initialised with a randomly generated Small-World network. This network remains fixed over the whole simulation run and serves as the seed for the networks that agents build up and use when they adjust the rating of their behavioural rules. Obviously, the density of the acquaintances network has an important influence on the accuracy of the rating of the different rules. If the network is sparsely linked it is possible that some agents never "get to know" about rules that are more suitable than the one they currently use. Corresponding simulation runs show isolated agents that are not able to adjust their rule. Caused by the described up-stream/downstream dependencies between the agents this might have a severe impact on the overall performance of the LRS along the channel.

Constantly normal weather conditions

If the simulated weather sequence does not include any extreme conditions over a longer period of time, agents using rule 0 (not maintaining their LRS) will outperform those constantly maintaining their LRS (rule 4). Assuming a reasonably dense network that allows for the required exchange of information, agents using rule 4 will tend to change to rule 0. This will quickly shrink the network of the agents using rule 4 until it completely disappears. In that case the collective knowledge about rule 4 is no more available. This disables the agents to re-adapt to more extreme weather conditions later on.

Spatial distribution of rules

Agents using rule 4 strongly depend on their upstream neighbours also maintaining their LRS. If a simulation run is initialised with a (spatially) cohesive subgroup of agents using rule 4 that have their land parcels located in the very upstream section of the simulated space, then their behaviour will quickly pay off (assuming weather conditions that require an LRS, see above). As agents belonging to the subgroup continue to have success, the knowledge about rule 4 will spread across the network making other agents use the same rule (assuming a reasonable dense acquaintances network, see above).

This effect is sometimes "undermined" by free-riders. Free-riders are direct neighbours of the most downstream agent of a spatially cohesive subgroup of agents maintaining their LRS and do not maintain their local LRS. In the model we assume that such agents partly profit from the maintenance efforts of their upstream neighbours. Inversely, agents located downstream from the free-riders do not get any gain from maintaining their LRS.

An extreme case occurs if, at the start of the simulation run, the subgroup of agents with rule 4 is located more downstream, so that there are upstream agents that use rule 0. In such a case the existing upstream neighbours that neglect their local LRS (rule 0) will prohibit any positive effect on the LRS. Thus, agents using rule 4 will quickly switch to rule 0.

Policy impact

This section gives a more detailed description of three scenarios that explore the effect of the amount of compensation paid to agents that lose part of their harvest due to flooding or drought. All scenarios are based on a fixed set of initialisation settings and only the parameter that defines the amount of the compensation payments is varied. In detail we use a fixed Small-World network with average degree of 10, a fixed weather sequence with 40% drought years, 20% flooding years, and a fixed land parcel map with equal land parcel sizes. The land parcels are managed by 30 agents of whom the 8 most upstream agents start out with a well maintained LRS on their parcels and use rule 4, i.e. maintaining the LRS. The 22 remaining agents are initialised with an un-maintained LRS on their land parcel and use rule 0, i.e. neglecting it. Compensation is paid such that an agent that has lost (all or some of) its harvest by flooding or drought receives a payment that equalises the occurred crop yield loss up to a percentage of the maximum possible yield. If an agent decides to switch from rule 0 to rule 4 then an allowance is paid that covers the investment to get the LRS in full operation. After that the agent itself will pay for the LRS maintenance.

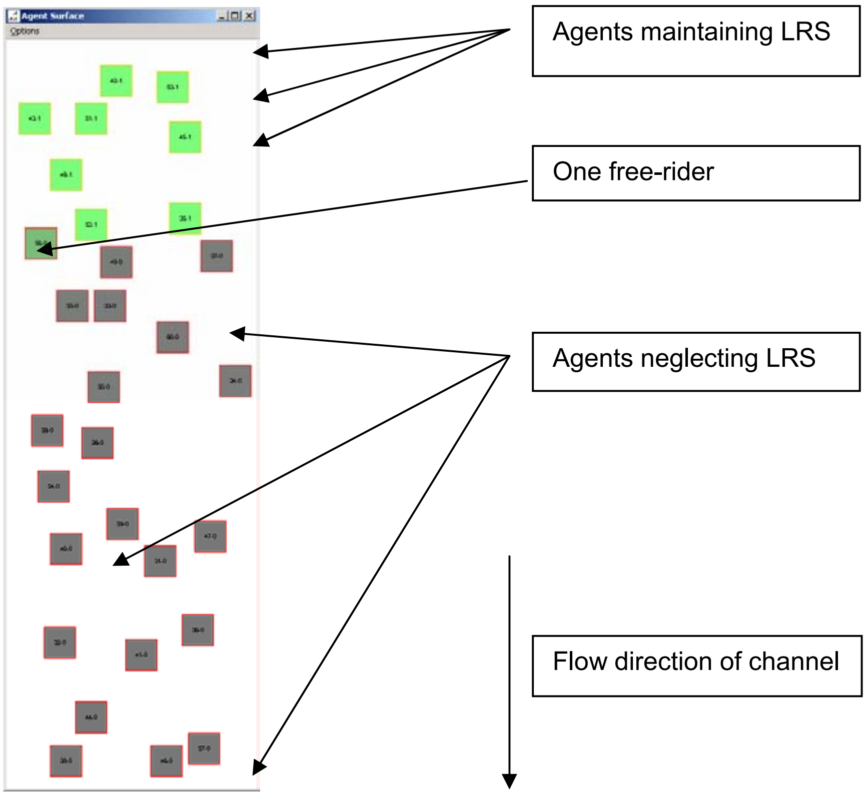

Figure 5: Abstract land parcel map and initial distribution of rules used for the three scenarios. An imaginary channel going from north to south defines the dependency between the agents. The filling colour of the boxes denotes the state of the LRS (green = good condition, grey = bad condition), the frame of the boxes shows the strategy the respective agent uses (yellow = rule 4, red = rule 0).

Scenario 1: Compensation payment of 100%.

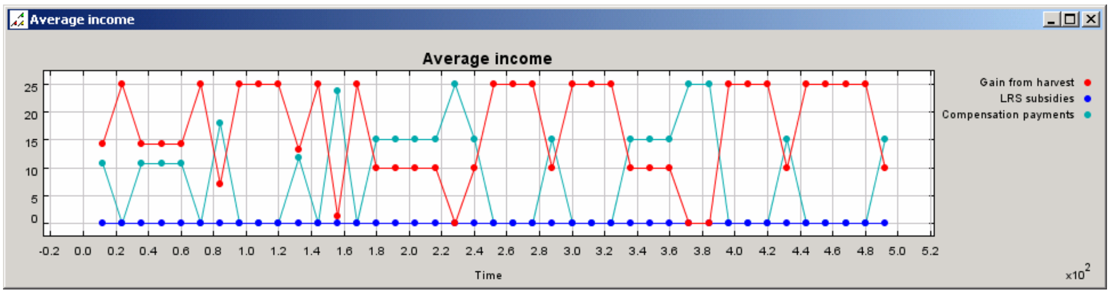

Figure 6: Average composition of the agents' income over time. No LRS allowances (blue) are paid and compensation payments (light blue) completely equalise loss in the crop yield if caused by extreme weather.

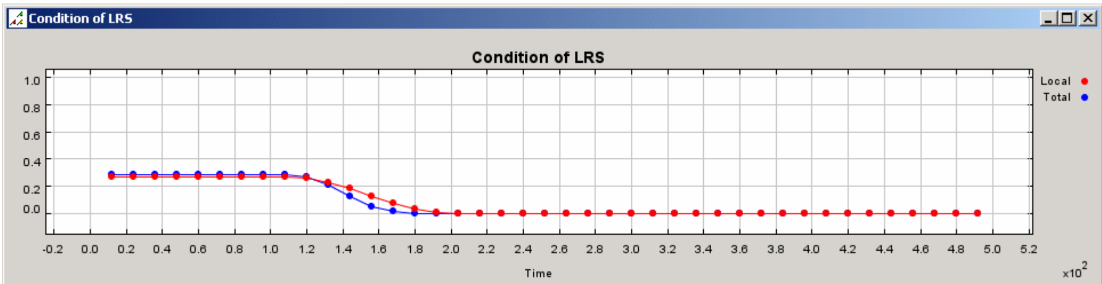

Figure 7: Averages of the local LRS condition and the resulting total LRS condition that takes the dependencies into account over time (months). As the agents change from rule 4 to rule 0, the condition of the LRS drops to 0.

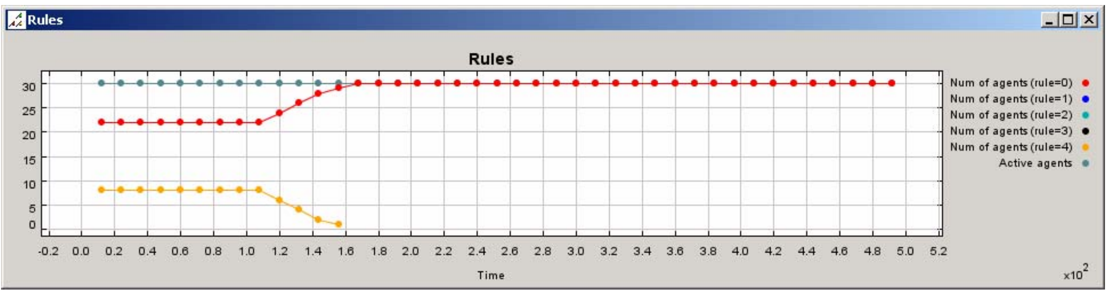

Figure 8: Distribution of rules between the agents over time (months). After 100 months agents start to change from rule 4 to rule 0 until all agents use rule 0.

The results from this scenario mainly show that if the economic risk from climate shocks is completely neutralised by compensation payments the agents quickly switch to the passive rule. This corresponds to the results described earlier for a constantly normal weather sequence without any extreme weather events, when agents get “lazy”. It also corresponds to the behavioural predictions one would make from the collective action structure of the LRS.

Scenario 2: Compensation payment of 50%.

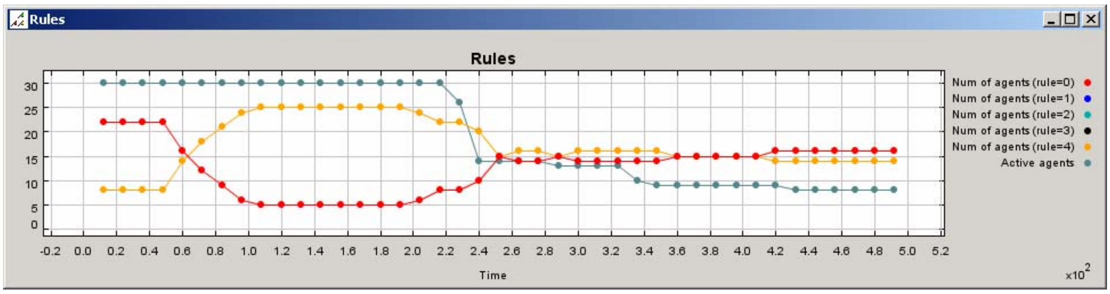

Figure 9: After 60 months agents start to change from rule 0 to rule 4 until 25 agents use rule 4. Due to the lack of economic success (see diagram below) the number of active agents decreases significantly, i.e. they exit the simulation.

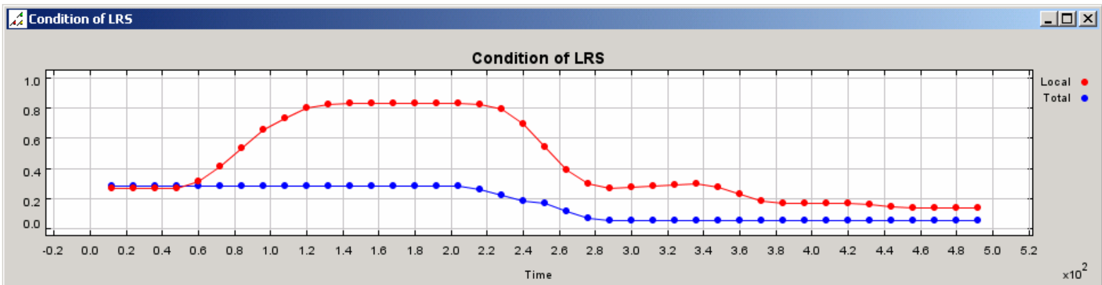

Figure 10: As agents start to change from rule 0 to rule 4 the average local LRS condition (red) in-creases significantly. Because of the local dependencies between the agents the average effective total LRS condition however stays low, probably caused by a free rider upstream. Thus, the total protection against weather shocks stays low and even decreases, as the number of active agents goes down.

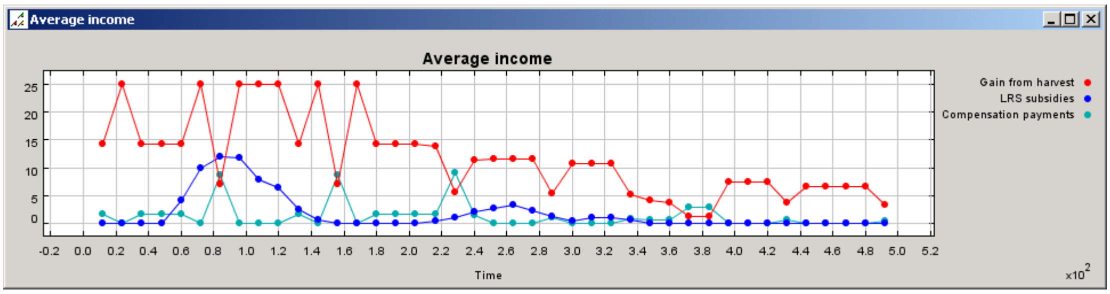

Figure 11: Compensation payments are much less than in scenario 1, but around step 60 there is an increased payment of LRS allowances. The average gain from harvest finally decreases as agents become inactive. Fewer allowances and compensations are paid to fewer agents.

This scenario shows that with a moderate amount of compensation payments all but 5 agents switch to rule 4 and maintain their LRS. Caused by the fixed defined local dependencies of the agents these 5 agents are enough to spoil the possible success of rule 4. Therefore, agents become inactive due to a bad economic balance or they switch back to rule 0.

Scenario 3: Compensation payment of 40%.

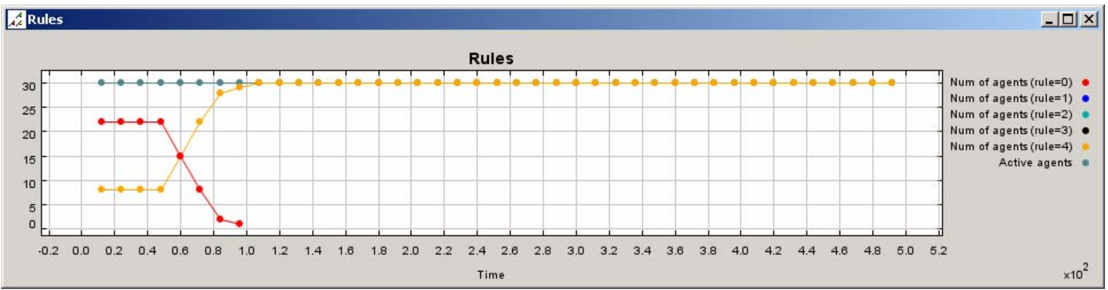

Figure 12: Radically different from the other scenarios all agents change to rule 4.

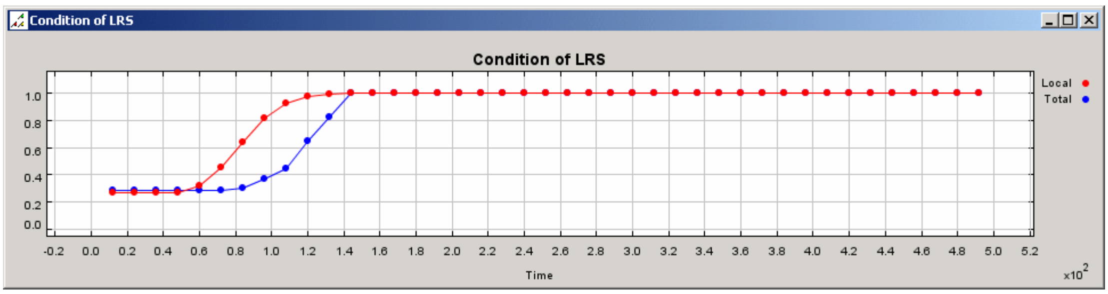

Figure 13: The effective total LRS condition (blue) improves significantly to 100%.

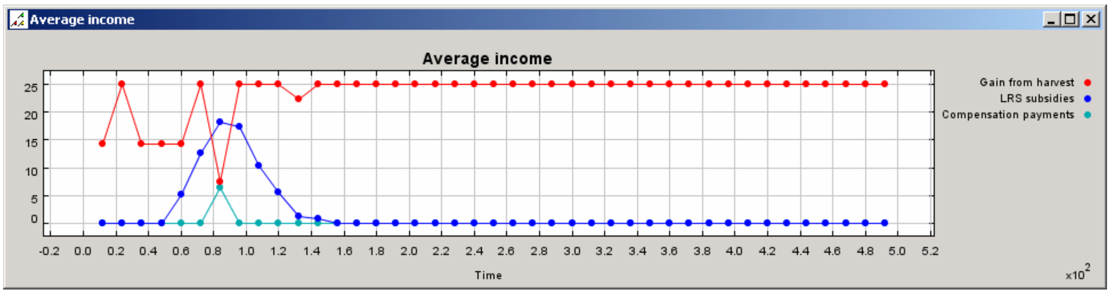

Figure 14: After approximately 90 steps the gain from harvest reaches the maximum level and compensation payment become 0. Also the allowances paid for LRS setup are clearly shown. They serve to build up a coherent LRS. After 160 steps the system stabilises.

For the parameterisations used in the three scenarios there is a critical value for the compensation payments somewhere between 40% and 50%. As may be seen from scenario 3, 40% compensation is not enough of an economic incentive for the agents to continue applying rule 0 (under the given climate, network and geographical conditions) and thus stay “lazy”. For that reason, the system stabilises quickly in a state where all agents collectively maintain their LRS. This occurs after a minimum investment of allowances and compensation payments.

Discussion

The SoNARe model has been abstracted to represent the most important and first key features of the Odra case study. It shows the impact of different interacting concepts (such as social networks, weather conditions, and spatial distribution of agents’ strategies). More important, it captures features like the role of free riders in the social dilemma and the impact of different policies, i.e. compensation payment for losses from drought and flooding vs. allowance to start the LRS. Based on the SoNARe model architecture, future model developments will be carried through.

A possible interpretation (within the bounds of the model’s validity) of the scenarios shown here is related to the beneficial use of a weighted combination of both compensations payments (for not loosing too much farmers in the region) and more targeted allowances to bring functioning LRS in place. This will be investigated further.

3.2.10 Outlook

As a next step it is planned to couple the biophysical model developed by WUT to the SoNARe model and hopefully to produce results that are comparable to the ones shown here. Some attention will be given to the localisation of cooperation and free-riding in the social networks and their impact on building functioning LRSs. Necessary sensitivity analyses will be carried through.

The storylines provided by Wroclaw University will serve as the basis to make the model more realistic by writing more differentiated behavioural rules for the agents.

Besides the social influence already implemented in the model, more differentiated psychological key variables, e.g. trust, will be considered.

It is planned to submit a paper on the SoNARe model to JASSS in the next months.

3.2.11 References

Albert, R. and Barabási, A.L. (2002): Statistical mechanics of complex networks. Rev Mod Phys 2002, 74, 47–97.

Barreteau, O. and Daré, W. (2003): A role-playing game in irrigated system negotiation: between play and reality. JASSS 2003, 6(3).

de Nooy, W., Mrvar, A. and Batagelj, V. (2004): Exploratory Social Network Analysis with Pajek. Cambridge University Press (Cambridge).

Jin, E., Girvan, M. and Newman, M.E.J. (2001): The structure of growing social networks, Phys. Rev. E 64, 046132 .

Moss, S., Edmonds, B., and Wallis, S. (2000): The Power Law and Critical Density in Large Multi Agent Systems. Centre for Policy Modelling Discussion Papers, Report No. CPM-00-71, <http://cfpm.org/cpmrep71.html>

Newman, M.E.J. (2003): The structure and function of complex networks. SIAM Review, 45(2):167–256.

Pujol, J. et al. (2005): How can Social Networks Ever Become Complex? Modelling the Emergence of Complex Networks from Local Social Exchanges. Journal of Artificial Societies and Social Simulation, 8 (4).

Wassermann, S. and Faust K. (1994): Social Network Analysis. Cambridge University Press, (Cambridge).

Watts, D. J. (1999): Small Worlds. Princeton University Press (Princeton).

3.3 Macaulay Institute

Project Month 1–6

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.

Project Month 7–12

3.3.5 Overview

Work by the Macaulay team can be divided into enhancements of FEARLUS-0-8 to prepare it for use with the prototype fine-grained model based on the Grampian case study, work on the possible uses of ontologies in relation to FEARLUS and more generally within CAVES, and design and coding work for the ontology-based version of FEARLUS (FEARLUS-1-0) to be used in the final Grampian case study models.

3.3.6 Enhancement of FEARLUS-0.8

Implementation has commenced of the designs for generalising the biophysical component of FEARLUS using look-up tables. As mentioned in the previous report, these look-up tables will enable FEARLUS to integrate with existing biophysical models where the latter are able to provide yield information for all combinations of situations in which crops might be grown.

In the first step, a new version of FEARLUS (0-8-2) has been created. This version features enhancements that enable more realistic representation of the environment, facilitating import of data from a GIS. The environment is divided into two layers: the cell layer, which stores data from the GIS; and the land parcel layer, which represents the division of the land into decision-making units. A land parcel consists of one or more cells, meaning that land parcels of different areas can now be simulated. (In the real world, land parcels are distinguished from their neighbours by physical boundaries (road, fence, hedge, stream, ditch), or by a history of different land use/management: fertiliser and pesticide use, previous tree-cover, laying of drains etc.) It is also possible to create blank cells that do not belong to any land parcel, enabling arbitrarily-shaped environments to be specified, as opposed to the restriction to rectangles in earlier versions. Figure 15 illustrates.

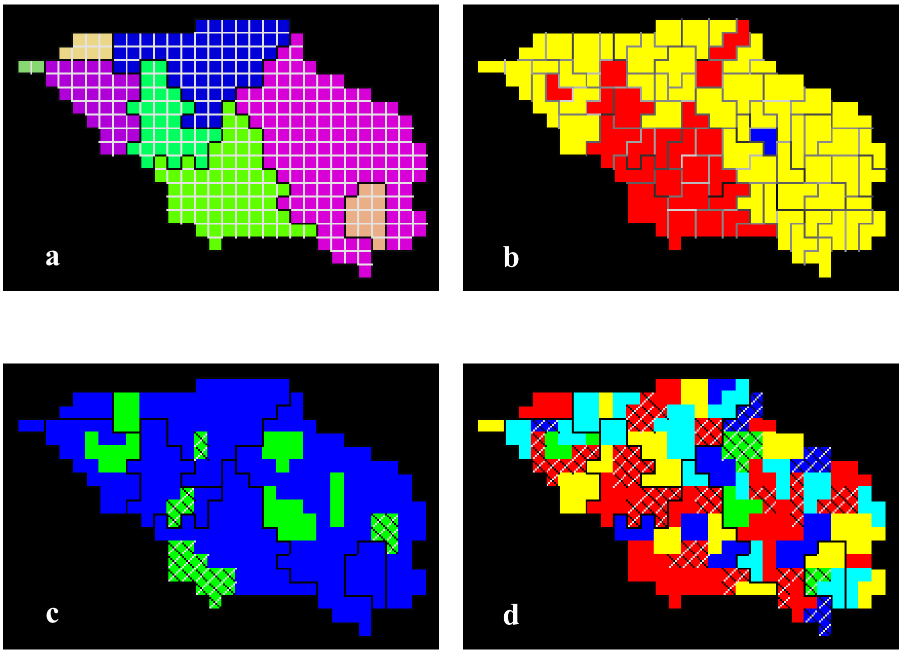

Figure 15: Example output from model 0-8-2. (a) Land cells coloured by the farmer who owns them. (b) Land parcels coloured by land use. (c) and (d) show farms with cells coloured by the most profitable and the most suitable land use respectively, with hashing used to indicate where there is more than one.

3.3.7 Ontology Literature Survey

The Macaulay modelling team have continued their work on identifying ways of using ontologies within agent-based modelling. A major part of this has been a survey of the most relevant parts of the extensive literature on ontologies. This has informed decisions about how ontologies will be used in future work using FEARLUS, and also proposals for their use within CAVES as a whole. Here, the main findings of the survey are briefly summarised, under headings reflecting the main relevant areas of the ontologies literature. In order to motivate the later subsections, we begin with a brief account of the main applications of ontologies thus far.

Application areas

The majority of work on applying ontologies has been outside the scientific arena, in areas such as knowledge management for businesses and e-commerce; or in developments which have both scientific and extra-scientific uses, such as e-learning and recommender systems. Within knowledge management (Abecker and van Elst 2004) ontologies are used to support knowledge visualization, search and retrieval; and as a basis for information gathering and integration. In e-commerce, the heterogeneity of product and requirement information used by sellers and buyers is a serious obstacle, which ontologies can help to surmount (Ying, Fensel et al 2004). E-learning applications can use ontologies to organize and classify resources such as scientific publications and software (Brase and Nejdl 2004); similarly, recommender systems can direct users to specific online papers (Middleton, De Roure and Shadbolt 2004).

Scientific disciplines for which domain ontologies have been built include biomedical science, ecology and bioinformatics, particularly as applied to molecular biology. McCray (2003) describes how a biomedical ontology has been built as part of a US National Library of Medicine project to provide integrated access to biomedical resources. Keet (2005) describes creating an ontology mapped to the STELLA ecological modelling tool, finding that the formalization of knowledge can itself suggest directions for further research in ecology. Stevens, Wroe et al (2004) review a range of bioinformatics applications.

Within multi-agent systems (MAS), ontologies have been applied to underpin agent problem-solving and inter-agent communication in financial applications (Sycara and Paolucci 2004). However, despite the fact that CAVES is using agent-based modelling, this line of work is not directly relevant to what we propose below, which does not involve the model agents themselves using ontologies. So far as we can ascertain, our proposed uses for ontologies in relation to agent-based modelling are novel.

Two "super-domains" of particular interest to CAVES are that of geospatial entities and relationships (Agarwal 2005), and that of processes (Grüninger 2004). The main geospatial applications are in relation to GIS. Issues include the relationship between purely spatial and spatio-temporal approaches, problems of scale and granularity, boundaries, and spatial vagueness. The representation of processes within ontologies has been explored in relation to domains including workflow management, Web services and manufacturing systems. More general process ontologies have also been attempted. Issues include sequence and repetition, causality, agency and the distinction between continuous processes and discrete events.

3.3.8 Ontology-Related Languages and Formalisms

There is a huge literature on ontology-related languages and formalisms. The most relevant to CAVES concerns the language OWL (Antoniou and van Harmelen 2004) and related formalisms. OWL is supported by the semantic web community[16], appears to be the most widely used formalism, is compatible with some of the most useful ontology-related software available – notably Protégé[17] and has a sound logical basis in the description logic SHIQ.

3.3.9 Ontology Learning, Particularly From Free Text

In a previous report, the Macaulay modelling team outlined a 7-stage procedure for deriving ontologies from textual evidence. The first three stages were:

1. Assembling evidence.

[16] http://www.w3.org/
[17] http://protege.stanford.edu/

2. Preliminary examination of evidence 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 used to suggest subclasses, properties, and supporting classes.

At that time, we were not aware of software systems developed to partially automate step 3. Ding and Foo (2002) and Gómez-Pérez and Manzano-Macho (2005) survey "ontology generation" or "ontology learning" systems, some of which take free text as initial input (perhaps along with an initial ontology, or a set of key terms), while others require structured or semi-structured sources. From free text, natural language processing software is used to extract candidate terms for the ontology's concepts, along with information about their taxonomic and other relationships. We plan to examine the KAON Text-To-Onto system (Maedche and Staab 2002), the language processing architecture GATE (Cunningham, Maynard et al 2005), and possibly other systems.

3.3.10 Upper-Level Ontologies, and Ontology Design Principles

Upper-level ontologies are designed to specify the key concepts and relations within some very broad domain. They cover such high-level distinctions as abstract/physical and object/state/process; and relations such as part/whole (mereology), connected or separate (topology), and spatial, temporal and causal relations.

Their intended function is as "a foundation for more specific domain ontologies" (Niles and Pease 2001). These authors describe the "Suggested Merged Upper Ontology" (http://suo.ieee.org). Other upper ontologies are described by Degen, Heller et al (2001) and Masolo, Borgo et al (2003). These last do not aim to provide a single monolithic top-level ontology, but to help people and computers understand one another by isolating the "fundamental ontological options", and providing a range of "foundational ontologies" as possible starting points for domain ontologies, each explicitly based on specific "ontological commitments". A closely related line of work is exemplified by Guarino and Welty (2004), who outline the "OntoClean" methodology for "validating the ontological adequacy of taxonomic relationships". This is based on highly general notions drawn from philosophical ontology, which are used to characterise "metaproperties" of terms in an ontology. Awareness of these helps ensure that a domain ontology avoids inconsistency.

3.3.11 CAVES Upper Ontology

Work is underway on a CAVES upper ontology, constructed using the DOLCE upper ontology of Masolo, Borgo et al (2003), and the OntoClean methodology of Guarino and Welty (2004), along with the existing ontologies of the Grampian and South African case studies.

3.3.12 Design of FEARLUS Successor

A successor to the current model 0 family of FEARLUS models has been designed.

Dubbed fearlus1-0, the design is founded on the use of ontologies to act as a bridge between evidence and case study data. Four ontologies are used, as described below.

First, the *domain ontology* consists of a description, using an ontology, of the real-world concepts that fearlus1-0 implements. The purpose of this ontology is to describe the conceptual context in which fearlus1-0 sits. It will therefore contain the description of concepts that do not appear in fearlus1-0 necessarily, but are related to it in some way.

Secondly, the *framework ontology* is intended to represent the concepts that are implemented by fearlus1-0. The framework ontology contains a description encompassing all the particular models it might be used to implement, including descriptions of implementation variants on particular concepts. The framework ontology imports the domain ontology, adding subclasses and subproperties to it that correspond to each of the implementation variants of those concepts in the domain ontology that fearlus1-0 provides implementations for.

Thirdly, a *scenario ontology* describes the concepts applying to a particular scenario or case study. The idea is that this ontology could be developed independently of the domain ontology, fearlus1-0 or from any particular model one might have in mind. This would be desirable for methodological reasons if, for example, one wished to show strictly that the model was derived from evidence rather than the other way round. Where there are no such constraints, the scenario ontology could of course be influenced by the domain ontology and/or design for a model.

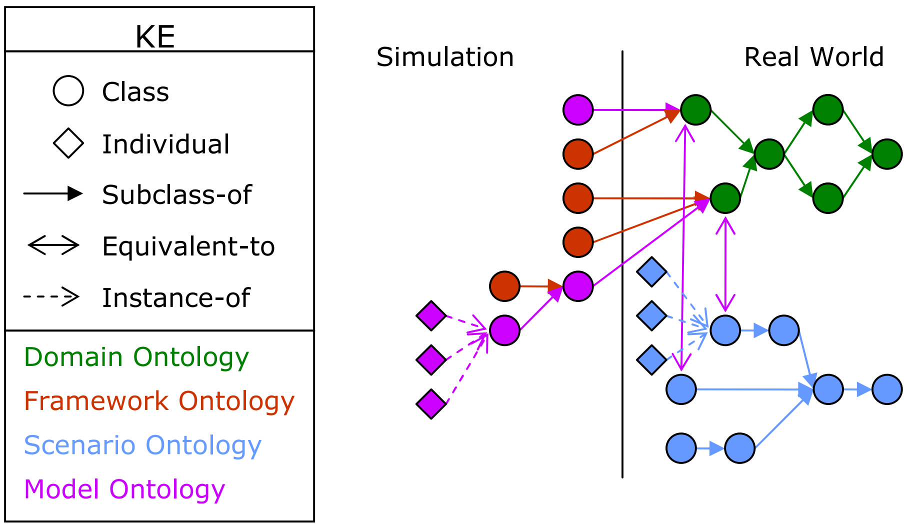

Figure 16: The relationship between the various ontologies. Assertions are colour-coded according to the ontology they appear in, with the domain ontology in green, the framework ontology in red, the scenario ontology in blue, and the model ontology in purple. The model ontology subclass-of and class assertions are also part of the framework ontology.

Finally, the *model ontology* brings fearlus1-0 and the scenario together, linking the evidence to the model through importing both the framework ontology and the scenario ontology. To reflect the fact that the model ontology is a specific instantiation of the modelling framework, the model ontology contains a subset of those concepts in the framework ontology that do not appear in the domain ontology, i.e. a particular choice of implementation variant concepts. The model is presumably intended to reflect some specific aspect of the scenario, so the model ontology needs also to specify how the relevant concepts in the scenario ontology are related to concepts in the domain ontology. The model ontology should explicitly state which classes in the scenario ontology are deemed equivalent to concepts in the domain ontology. Thus, an explicit, transparent link is created from entities in the scenario to their particular implementation in the model: all classes in the model are subclasses of concepts in the domain ontology that have been declared to be equivalent to concepts in the scenario ontology. The relationship between the various ontologies is shown in Figure 16.

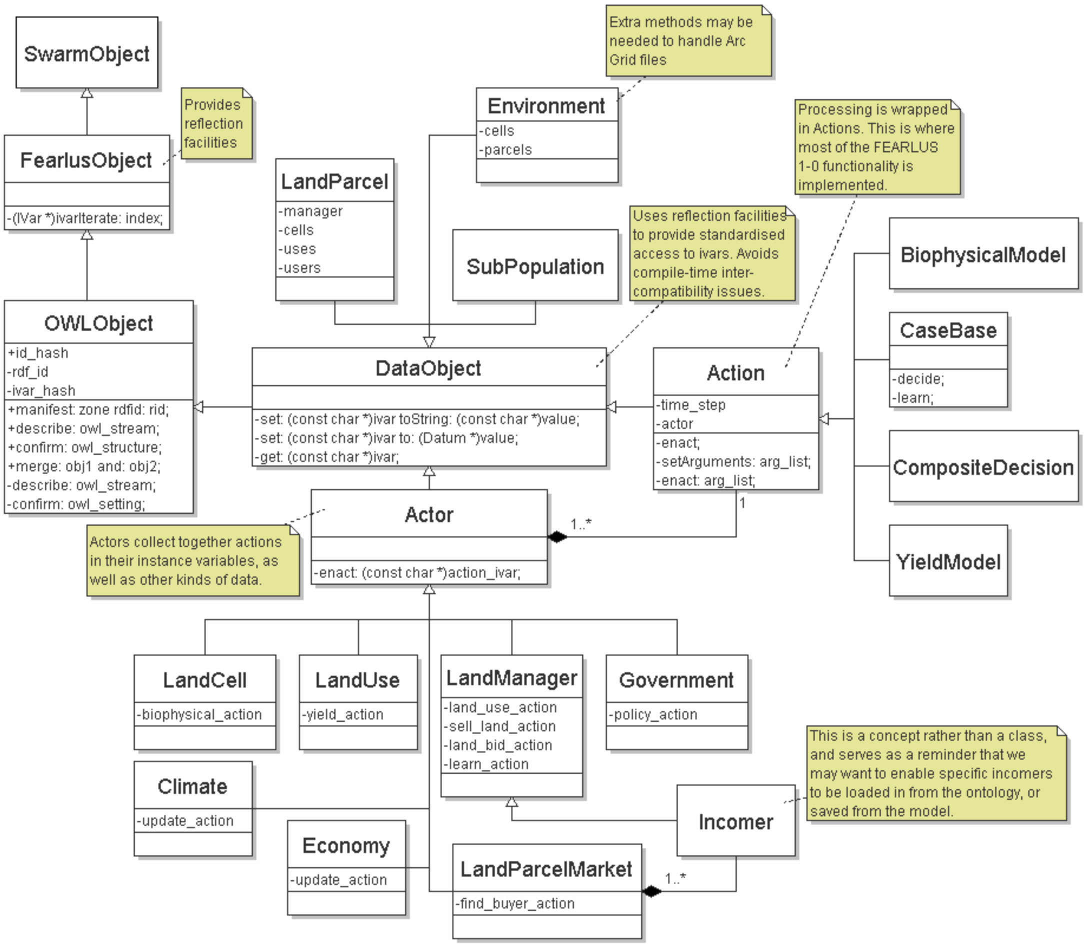

Figure 17: UML class diagram for FEARLUS-1-0. Note that a more Obj-C-like syntax is used for the instance variables and methods than in standard UML.

Figure 17 shows a UML class diagram outlining this aspect of the design of fearlus1-0. The OWLObject class is used to read in the model ontology and configure those classes and instance variables in subclasses of DataObject that will be used. Many classes act simply as data repositories (LandParcel, Environment and SubPopulation), not responsible for any change to the state of the system during a simulation, but containing information that will be used as a basis for change.

Actors are entities that are responsible for changing the state of the system. They are distinguished from the data repositories by merit of containing instance variables that store Actions rather than other kinds of information. A method is provided to cause the Action to run. Finally, subclasses of Action store the algorithms that will be used to create the changes that occur in the model.

This design has been submitted as a paper entitled "A new approach to modelling frameworks" to WCSS 2006. This is the first submitted CAVES publication from the Macaulay modelling group.

Project Month 13–18

3.3.13 Overview

Work by the Macaulay team can be divided into work on the prototype model based on the Grampian case study, and work preparing for the final Grampian case study model to be constructed during 2007. However, due to developments in our ideas for both these stages, a change of terminology has been decided on for our current and future software. The modelling system supporting the prototype model is now designated FEARLUS1-0; this is intended to be a stable platform, modified incrementally and minimally in future, with careful attention to backward compatibility. The ontology-based software to be used in the final Grampian case study model will now consist of two layers: a modelling platform, AMEBON (Approaches to Model-Evidence Bridging with Ontology Networks), and a modelling system CARLESS (Climate and Rural Land Use Simulation System).

3.3.14 Work on FEARLUS1-0 and Grampian Case Study Prototype Models

We decided there was a need for a small number of narrative scenarios to act as starting points for the Grampian case study prototype model, each telling a "story" which would guide us in the selection of parameter sets, and the assessment of model outputs. The development of FEARLUS1-0 has been undertaken in conjunction with the development of these scenarios.

Scenarios for the Grampian Case Study Prototype Model

The following scenarios have been selected as the first to be developed in detail and modelled in FEARLUS1-0:

Future Climate Change

Gradual increase in both, mean temperature (perhaps making some crops viable that are not so at present), and climate variability (increasing risks); at the same time, either:

- Increasing prices for bio fuel crops (probably oilseed rape, with which many farmers already have experience); and/or
- A considerable increase in the price of fuel, (it could well make sense to combine this with increasing prices for bio fuel crops: whether the increase in fuel prices was down to changes in the market price of oil, or to the withdrawal of tax concessions on agricultural fuel, it would be likely to be accompanied by increasing bio fuel prices); and/or
- Increasing environmental awareness generating increased premium on (local) organic produce, but with continuing social pressure from other farmers against going organic (there is a strong feeling evident in interviews that this would be "going backwards").

Future Epidemic

This could be foot-and-mouth (probably the most likely in the real world, since avian flu would not have a significant effect in the Grampian area, where poultry farming is economically insignificant). Alternatively, it could be a notional new disease (BSE, which was such a disease, was a "shock" that has been mentioned by a number of interviewees), in which case we could run multiple variants of the scenario, depending on the features assigned to the new disease. We could, for example, say that it makes it impossible to keep cattle (or sheep, or to grow barley…), for ten years – or poses a possible threat to human health, as BSE did, causing a price crash. Clearly, an epidemic scenario could be combined with the Future Climate Change scenario without much additional work, and the recent appearance of bluetongue disease far north of its usual range indicates that such a connection is plausible.

The Recent Past (1987-2003)

This will require the most work, because it will need to be based on the best available data about prices and climate. Within this period there are at least two "shocks" we can represent in FEARLUS1-0: "Black Wednesday" (sterling leaving the European ERM in 1992, leading to a sharp fall in the exchange rate and a rise in the sterling prices for agricultural exports), and BSE. Foot-and-mouth might be a third, but has been little mentioned so far: the actual disease never reached Grampian, but restrictions on movement of animals must have had some effect.

Progress Towards FEARLUS1-0

In the progress report of March 2006, we described FEARLUS0-8-2, which incorporated enhancements to enable more realistic representation of the environment, separating two layers: the cell layer, which stores biophysical data and can be loaded from a GIS; and the land parcel layer, which represents the division of the land into decision-making units.

The next step was to integrate look-up table code with version 0-8-2 to create

FEARLUS0-8-3 that uses look-up tables. The bit strings that formed the basis of the biophysical model in earlier versions of FEARLUS are now replaced with look-up tables, which allow the yield for a particular situation (in terms of climate, local biophysical characteristics, crop and land management options) to be looked up in a table. These changes allowed more sophisticated descriptions of land uses to be made, and included the facility for land managers to have social pressure for or against particular aspects of land management practice (e.g. choice of crop or management regime).

Concurrently, experiments were conducted with an earlier version of FEARLUS that uses an endogenised land market model (ELMM). Results indicate that the land market model does have an effect on the overall behaviour of the model in simulations involving innovating and imitating subpopulations of land managers. Integration of the endogenous land market model with FEARLUS0-8-3 produced FEARLUS0-8-4, considerably increasing the realism and flexibility of the model. Our investigations with ELMM took rather longer than anticipated, which caused a slight delay to the release of FEARLUS0-8-4 from July to mid-August.

FEARLUS1-0 is to contain the following extra functionality from FEARLUS0-8-4:

- Measures to implement economies of scale. Two options to implement this are to be provided. One, coded at the time of writing, enables farm-scale fixed costs to be represented, and allows these to change over time by providing the facility to load them from a file in each successive Year in the model. With this setting, Land Managers in the model with larger Farms will have lower costs per unit area. The other, to be coded shortly, will allow a bounded linear function to represent the cost savings due to economies of scale for particular land uses.
- Representation of off-farm income. Land Managers can take off-farm income each Year from a truncated normal distribution with a mean and variance each from a uniform distribution specified at Subpopulation level. This has been coded at the time of writing.
- Augmentation of Cased Based Reasoning agents with the facility to use a particular strategy when experimenting. Currently CBR agents simply choose a new Land Use at random when the reasoner cannot find a suitably matching case. This modification will enable the use of imitative and other experimentation strategies to be employed in this situation. This modification is to be coded shortly.

3.3.15 Preparatory Work on AMEBON, CARLESS and the Grampian Case Study Final Model

Work has continued on ways in which ontologies, and related formalisms and procedures, can be used in support of agent-based simulation, and specifically in linking models and evidence more transparently.

Roles of Ontologies in Agent-Based Simulation

Background methodological work on the possible roles of ontologies, and related formalisms, in supporting agent-based simulation has continued through the past six months, resulting in two conference paper. Gotts and Polhill (2006) argues that:

- To assist scientific understanding of the world's most complex systems – the mass societies in which most people live, and which have most effect on the natural environment – simulation model need to incorporate the social, technological and ecological aspects of human systems.
- Simulating systems containing such a broad range of entities and processes requires a highly modular, interdisciplinary approach, combining fundamentally different kinds of model.
- Ontologies, together with a range of associated technical and institutional developments, can provide the basis for such an approach.

The paper proposes the following possible roles for ontologies in agent-based modelling:

- To provide detailed, formal characterisations of simulation models, independent of specific programming languages and detailed algorithms, and which specify what aspects of the model are intended to be theoretically significant. If other researchers can show that any model built to the same specification gives significantly different results, they will have undermined any conclusions drawn from the original.
- To provide support for the development and validation of simulation models using sources including the existing literature, questionnaire responses, and expert/stakeholder interviews.
- To show what has been considered as potentially important, but not included in the simulation model – this is vital in the case of systems as complex and multi-faceted as the land use systems with which CAVES is concerned.
- To assist in formally modelling the workflow involved in designing and using a model or set of related models in simulating experiments (Christley et al 2004).
- To allow simulation experiments to be carried out in distributed fashion, and to allow simulation models to be made available for use by researchers other than their developers through the deployment of Grid technology.
- Finally, and of greatest significance here, to enable the processes of comparing and combining models, which is central to the approach to agent-based modelling we advocate. There is current research on ontology "mapping" (concerned with reusing existing ontologies, expanding and combining them – Ding and Foo 2002). If simulation models were routinely accompanied by ontologies, the process of comparing models in the same domain (identifying both similarities and key differences) could be greatly facilitated.

The second conference paper, Gotts and Polhill (forthcoming), argues that the kinds of system studied using agent-based simulation are intuitively, and to a considerable extent scientifically, understood through natural language narrative scenarios, and that finding systematic and well-founded ways to relate such scenarios to simulation models, and in particular to their outputs, is important in both scientific and policy-related applications of agent-based simulation. A projected approach to the constellation of problems this raises – which derive from the gulf between the semantics of natural and programming languages – is outlined, involving the use of mediating formalisms: ontologies and specialised formalisms for qualitative representation and reasoning such as Allen's temporal interval calculus (Allen and Kautz 1985), the region-based qualitative spatial representation and reasoning system RCC (Cohn et al 1997), and the QDE formalism (Kuipers 2001) for representing qualitative relationships between changing numerical variables.

Design of AMEBON and CARLESS

The design of AMEBON was presented at the first World Congress on Social Simulation (Polhill and Gotts, 2006). A UML diagram for the design is presented in the diagram below.

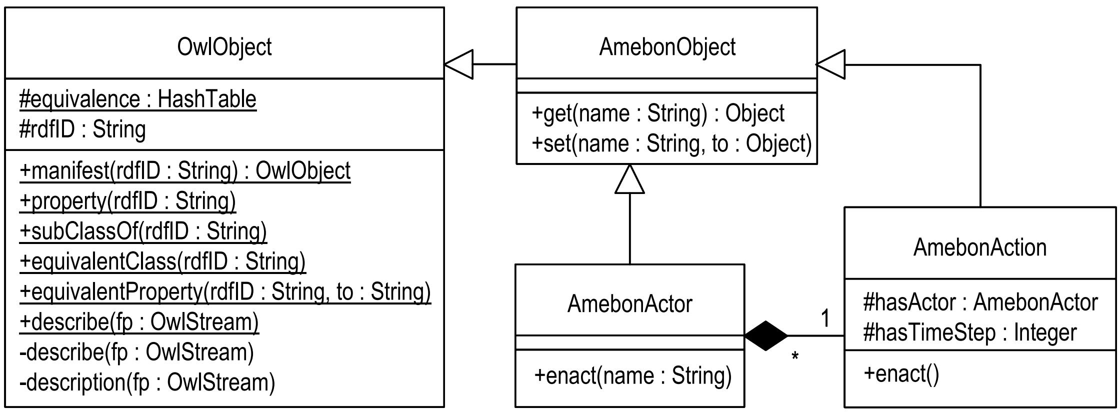

Figure 18: UML diagram of the AMEBON design.

AMEBON is a modelling platform design to read in an OWL ontology describing the objects that appear in the model, create them, run a simulation with them, and then output an OWL ontology describing the state of the system at the end of the simulation. The AMEBON design is based on a distinction between actors/actants (entities in the simulation that perform actions and/or have actions performed on them that change their state) and actions (processes that are performed by actors that cause the change). The model ontology will describe with actors, actants and actions are to appear in a particular simulation. Objects in AMEBON are designed to have a stricter sense of encapsulation than in standard OO programming languages, more akin to the principles of Agent-Oriented Programming (see e.g. Wooldridge, Jennings & Kinny, 2000), though not quite taking things that far. Instance variables (the "properties") of actors and actants are accessed through the same get and set methods, and methods (the actions) performed through calling an enact method with the name of the action to perform. This enables greater modularity in the architecture, in which new objects, properties and actions can be added without affecting the existing architecture.

CARLESS will be built on top of AMEBON, providing subclasses of AmebonActor and AmebonObject for the actors and actants that are to appear in CARLESS, and subclasses of AmebonAction for the actions. A CARLESS model ontology will be used to select those components of CARLESS that are to appear in a particular simulation.

From Interviews to Ontologies

The major input to the ontologies underlying the final Grampian case study model will be the interviews with farmers, and people with close social and professional links with farmers (family members, suppliers, accountants, officials of farmers' associations). In the last progress report, we described an experiment in producing an ontology from text sources (survey results and existing literature), which was presented at the International Conference on eSocial Science (Polhill and Ziervogel, 2006), and noted that we had not at the time of that experiment time been aware of possible relevant software systems, specifically the KAON Text-To-Onto system (Maedche and Staab 2002), and the language processing architecture GATE (Cunningham, Maynard et al 2005).

During the last six months we have investigated these, as well as searching for other software systems that could be relevant. Of these others, the only system that is both relevant and available for general use is the commercial NVivo system[18], used by social scientists to annotate interview and similar material for the purposes of qualitative research. Text-To-Onto[19] and its successor, Text2Onto[20], which is currently under development, are specialised for use with large text corpora in a business context, and are also sparsely documented. GATE[21], by contrast, is non-specialised, and copiously documented.

A provisional approach has been developed for constructing a scenario ontology for the Grampian case study from a set of interview texts, and a pre-existing (land use) domain structure ontology (which is imported to the scenario ontology):

1. Assemble source materials. In the Grampian case, this means some subset of the interview transcripts, as anonymised and annotated by the interviewer (Lee-Ann Small). This annotation involves assigning pieces of interview text to a set of "nodes", corresponding to topics either selected in advance, or emerging from the interviews themselves (e.g. "biggest-changes" (in farming over recent years), "BSE", "attitudes-to-borrowing"). These nodes are being chosen by Dr. Small using normal qualitative research criteria, without direct input from the rest of the team (although the latter were of course involved in setting the general themes of the interviews). One piece of text may be assigned to multiple nodes.

2. Decide on key (not necessarily high-level) concepts.

    - Go through the annotated interviews, deciding whether any nodes need to be merged, added or deleted for the purpose of constructing the scenario ontology.

[18] http://www.qsrinternational.com/
[19] http://sourceforge.net/projects/texttoonto
[20] http://ontoware.org/projects/text2onto/
[21] http://www.gate. ac.uk

- o For each of the resulting NVivo “nodes” (topics), select 1-5 (preferably not more than 3) key concepts, along with a small number of relations between them, and key attributes. So far as possible, the key concepts should be “rigid” (meaning they cannot cease to apply to an individual entity while it retains its identity), and have clear identity criteria (see the Guarino and Welty 2004 discussion of the “OntoClean” methodology for ontology construction).
- o Attach these concepts and “properties” (relations + attributes) to the imported domain-structure ontology, expanding it if necessary.

3. Detailed analysis. This involves examining the nouns, verbs and adjectives in the interview scripts to fill in additional concepts, and properties/relations. Again, the properties/relations should be considered in the light of the OntoClean methodology. Working through the list of nodes, but where necessary cross-checking semantic relationships using whole interviews, GATE will be used to:
   - o Find all examples of particular semantic classes of nouns and noun-phrases – e.g. spatial and temporal locations, people, money, occupations/professions, animals, crops, personal relationships… Each class should correspond to a concept in the scenario ontology, with new concepts being added to it if necessary.
   - o For each such class, add the sub-concepts and instances to the scenario ontology.
   - o Find verbs and nouns associated with the concepts and sub-concepts added to the scenario ontology, and use them to add relations and attributes.
4. Construction of “defined” classes. (Those with necessary and sufficient membership conditions).
5. Checking that the scenario and domain structure ontology provide all the required concepts, and that there are no superfluous concepts or properties/relations. There may well be concepts that are not explicit in the interviews, but which common sense, or knowledge of the land use domain, indicate are required. Conversely, the overall structure of the ontology may suggest that some concepts are of negligible importance. However, such additions and removals, and the reasons for them, should be documented.
6. Testing. Using the FaCT++ reasoner that has become available for use with Protégé, test the consistency of the scenario ontology, and the satisfiability of all concepts, as far as is feasible.
7. Refinement. Using additional interview texts held back for the purpose.

3.3.16 Future Work

As over the past six months, work on the prototype model and on the software for the final model is planned to proceed in parallel. Completion of the software for the prototype model, and first simulation runs using that software, are expected during September: the scenarios outlined above, and possibly others, will be investigated over the next six months. Simultaneously, the new Grampian case study ontology, based on the evidence from the interviews carried out by Dr. Small, will be constructed, ready for input to the final Grampian model. The next six months are also expected to see the implementation of the first version of AMEBON, which is currently planned to be built using Repast Simphony.

3.3.17 References

Abecker, A. and van Elst, L. (2004): Ontologies for Knowledge Management. Ch.22 in Staab, S. and Studer, R. (eds) Handbook on Ontologies, Springer

Agarwal, P. (2005): Ontological considerations in GIScience. International Journal of Geographical Information Systems 19(5): 501-536

Allen, J.F. and Kautz, H. (1985) A model of naive temporal reasoning, pp.251-268 in J. Hobbs, J. and Moore, R. (eds), Formal Theories of the Commonsense World, Norwood, NJ., 1985, pp. 251-268

Antoniou, G. and van Harmelen, F. (2004): Web Ontology Language: OWL. Ch.4 in Staab, S. and Studer, R. (eds) Handbook on Ontologies, Springer

Brase, J. and Nejdl, W. (2004): Ontologies and Metadata for eLearning. Ch.28 in Staab, S. and Studer, R. (eds) Handbook on Ontologies, Springer

Christley, S., Xiang, X., & Madey, G. (2004): Ontology for agent-based modelling and simulation. In Macal, C. M., Sallach, D., & North, M. J. (eds.) Proceedings of the Agent 2004 Conference on Social Dynamics: Interaction, Reflexivity and Emergence, co-sponsored by Argonne National Laboratory and The University of Chicago, October 7-9. http://www.agent2005.anl.gov/Agent2004.pdf.

Cohn, A.G., Bennett, B., Gooday, J. and Gotts, N. M. (1997) RCC: a calculus for region based qualitative spatial reasoning. Geoinformatica 1, 275-316

Cunningham, H., Maynard, D., Bontcheva, K. and Tablan, V. (2002): GATE: A Framework and Graphical Development Environment for Robust NLP Tools and Applications. Proceedings of the 40th Anniversary Meeting of the Association for Computational Linguistics (ACL'02). Philadelphia, July 2002

Degen, W., Heller, B., Herre, H. and Smith, B. (2001): GOL: Towards an Axiomatized upper-Level Ontology. Proceedings, FOIS-01: International Conference on Formal Ontology in Information Systems, October 17-19 2001, Ogunquit, Maine, USA

Ding, Y. and Foo, S. (2002): Ontology Research and Development Part 1 – A Review of Ontology Generation. Journal of Information Science 28(2): 123-136

Gómez-Pérez, A. and Manzano-Macho, D. (2005): An overview of methods and tools for ontology learning from texts. The Knowledge Engineering Review 19(3), 187-212

Gotts, N.M. and Polhill, J.G. (2006): Simulating Socio-Techno-Ecosystems. First World Congress on Social Simulation, 21-25 August 2006, Kyoto, Japan.

Gotts, N.M. and Polhill, J.G. (forthcoming): Narrative Scenarios, Mediating Formalisms, and the Agent-Based Simulation of Land Use Change. EPOS 2006: Epistemological Perspectives on Simulation, II Edition, University of Brescia, Italy, October 5-6, 2006

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

Grüninger, M. (2004): Ontology of the Process Specification Language. Ch.29 in Staab, S. and Studer, R. (eds) Handbook on Ontologies, Springer

Guarino, N. and Welty, C.A. (2004): An Overview of Ontoclean. Ch. 8 in Staab, S. and Studer, R. (eds) Handbook on Ontologies, Springer

Keet,C.M. (2005): Factors affecting ontology development in ecology. Presentation given on 1-7-2005 at the KRDB Research Centre, Free university of Bozen-Bolzano, Italy <http://www.meteck.org/files/OntoDevelEco.pdf>

Kuipers, B. (2001) Qualitative Simulation, in Meyers, R.A. (editor-in-chief) Encyclopedia of Physical Science and Technology, Third Edition, pp.287–300.

Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A. and Schneider, L. (2003): The WonderWeb Library of Foundational Ontologies Preliminary Report. WonderWeb Deliverable D17, version 2.1, IST Programme of the Commission of the European Communities project number IST-2001-33052

McCray, A.T. (2003): An upper-level ontology for the biomedical domain. Comparative and Functional Genomics 4(1): 80-84

Middleton, S.E., De Roure, D. and Shadbolt, N.R. (2004): Ontology-based Recommender Systems. Ch.24 in Staab, S. and Studer, R. (eds) Handbook on Ontologies, Springer

Niles,I. and Pease, A. (2001): Towards a Standard Upper Ontology. Proceedings, FOIS-01: International Conference on Formal Ontology in Information Systems, October 17-19 2001, Ogunquit, Maine, USA

Polhill, J. G. and Gotts, N. M. (2006): A new approach to modelling frameworks. First World Congress on Social Simulation, 21-25 August 2006, Kyoto, Japan.

Polhill, J. G. and Ziervogel, G. (2006) Using ontologies with case studies: an end-user perspective on OWL. Second International Conference on e-Social Science, 28-30 June 2006, Manchester, UK. http://www.ncess.ac.uk/events/conference/2006/papers/abstracts/PolhillEndUserPerspectiveOnOWL.shtml

Stevens, R., Wroe, C., Lord, P. and Goble, C. (2004): Ontologies in Bioinformatics. Ch. 32 in Staab, S. and Studer, R. (eds) Handbook on Ontologies, Springer

Sycara, K. and Paolucci, M. (2004): Ontologies in Agent Architectures. Ch. 17 in Staab, S. and Studer, R. (eds) Handbook on Ontologies, Springer

Wooldridge, M., Jennings, N. R. and Kinny, D. (2000): The Gaia methodology for agent-oriented analysis and design. Autonomous Agents and Multi-Agent Systems 3 (3), 285-312.

Ying, D., Fensel, D., Klein, M., Omelayenko, B. and Schulten, E. (2004): The Role of Ontologies in eCommerce. Ch.30 in Staab, S. and Studer, R. (eds) Handbook on Ontologies, Springer

3.4 Politechnika Wroclawska

Project Month 1–6

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 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.

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 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.

Project Month 7–12

3.4.3 Biophysical Model for the Odra Case Study

The model for the Odra Case Study, developed by the Kassel team, requires a biophysical model that would be responsible for hydrology simulation, weather generation and crop prediction. This model is being developed by the WUT-IP team. At this point the model is capable of simulating such processes as soil water movement, channel routing and crop growth. It also contains a weather generator. As we expected, the hydrological part of the model was the most problematic one. At this point, soil water movement is based on Darcy's equation, while channel routing algorithms use Manning's equations for calculating the volumetric flow rate. The weather generator and the crop sub-model are based on SWAT (Soil and Water Assessment Tool). The weather generator produces daily values of rainfall, maximum, minimum and average temperatures, solar radiation and potential evapotranspiration. Rainfall is generated using a Markov chain and an exponential distribution. The temperature and solar radiation values are generated from a standard distribution, while the potential evapotranspiration is calculated using the Hargeaves method. The crop sub-model implements the heat unit theory. Plants' growth is stimulated by temperature and can be reduced by a shortage of water. The crop sub-model also takes into account the destructive effect of flooding on crops. For simplicity's sake, we decided to ignore the influence of nutrients and fertilizers on plant growth. However it is technically possible to include these two factors in the crop sub-model.

The biophysical model is fully spatially explicit i.e. for every time step it produces 2D grid spaces with such data as ground water level and biomass. The model requires a digital elevation model, currently in ASCII grid format, and a channel network specified in an ESRI shapefile. We are planning to add support for GeoTIFF files in near future.

Because of the complexity of this type of models and their intrinsic composite structure, we decided to build the biophysical model around the design pattern called “delegation”. The object which represents the whole model doesn't actually do any modeling on its own. It delegates tasks to specialized objects, called delegates, which perform required operations and return the results. At this point there are seven delegates. Each of them is responsible for one of the following tasks: weather generation. soil water movement, channel routing, rainwater distribution, evapotranspiration, crop growth, crop management. Each of these tasks can be performed in various ways, using various algorithms. The strong point of the “delegation” design pattern is that it makes it possible to easily replace one delegate without making any changes to other delegates. This in turn means that we can easily experiment with, for example, different channel routing algorithms, or various weather generators.

The biophysical model does not contain any data visualization. However, we have developed a graphical front-end for the model, which enables viewing of spatial and non-spatial data. In order to visualize spatial data, we have developed a special GUI component, that is capable of displaying multiple layers of both raster and vector data. Values of non-spatial data can be plotted on a 2D graph or displayed in a text field.

3.4.4 Tools for Visual Model Assembling and Editing

We are developing a set of tools, that would allow both programmers and non-programmers to build, or more precisely, to assemble models. The process of model assembly would resemble the creation of a GUI using a visual editor and the so called “pick and plop” technique. The modeler would select components from a palette and place them on a diagram representing the structure of the model. The tools would enable connecting these components together with various links to create a model. Some of the components in the palette would be visual components that enable data visualization or simulation controls, while other would be non-visual components of the model's logic. We are expecting that the majority of non-visual components will simply represent various classes from the Repast library.

The development works on such a set of tools are well on their way. The first POC (Proof Of Concept) project has been recently completed. In its current form, the tool enables placing both visual and non-visual components on a boundless plane and connecting them together.

3.4.5 Discrete Choice (Opinion Dynamics) Models

We have investigated the properties of two classes of discrete choice (opinion) models. One of them, more often used by economists, is formulated using generalized utility function, where utility may include subjective components. The evolution of the second class is formulated using “supporting” and “persuading” influences on individuals states. We have found that these two approaches can be formulated in a uniform way and each of them can be expressed in terms of the other. We have formulated a generalized model, which covers models of Brock-Durlauf, linear variant of Nowak-Latane and Holyst-Kacperski. Multistability of this model was shown within mean-field approach, i.e. within some range of model parameters and strength of external influences two stable stationary states are possible, one of them being more favorable then the other. This phenomenon can be used as a representation of “social traps” – overall utility of the system is smaller than it could be, because the system is stuck in a “trap". In our simulations this phenol-menon is manifested as existence of hysteresis, a phenomena similar to the occurrence of hysteresis in physical systems undergoing phase transitions. Work on formulation of general model covering the widest class of existing discrete-choice models is in progress.

Based on this generalized model, the analysis of influence of “self supportiveness” was performed. It occurs that in the presence of individuals’ tendency to persist in the state they are in, the phenomenon of “social traps” is more pronounced and stronger external influence is needed to convert the system to the more favorable state. In our simulations this effect was manifested as widening of the hysteresis.

We have analyzed the similarities and differences between this approach to the social systems and physical methods used in exploring physical systems, (particularly, thermodynamics of spin systems). In literature the parallel between social models and physical models was pointed out; in spite of this, many difficulties in this parallel may arise when analyzing the background of physical and social approaches. The status of so-called “social temperature” and “stationary states” in social systems is still questionable.

The preliminary investigation of spatial cluster formation in the model of binary choice was performed. The results known from literature were verified. Investigation on the role of non-uniformity of individuals, existence of a strong leader, interaction between individuals vanishing with mutual distance and topology of a network, is in progress.

We have formulated the differential equation for the “mean choice” within the mean-field approximation. Preliminary analysis of dynamics of such systems was performed. We have found, that these simulations confirm earlier results obtained within stationary state approach, but they allow much more detailed analysis, including obtaining a dynamical potential. We are still working on some simple models that would give intuitive insight in the meaning of inflows and outflows in the model.

We have investigated properties of the Brock-Durlauf model on small world networks. It occurs that increasing probability of "rewiring" (i.e. creation of long-distance connections among individuals) causes increase in the width of hysteresis. Similar effect is obtained by increasing the number of neighbours. This may lead to interesting considerations on definitions of "dimensionality" of networks regarding their topology. It also gives assumption to investigations of one more effect of the small world – only "a few" long distance connections modify structural properties of the system making it multistable.

3.4.6 Evolution of Social Networks

We have reviewed the literature about the structures of social networks and implemented algorithms for the calculation of their generic parameters. We have also reviewed the literature regarding evolving networks.

For some evolving networks the phenomenon of Self Organized Criticality can be observed (e.g. the power law in distribution of avalanches of changing links). Such networks, having also a sociological interpretation, may be a promising material for modelling social phenomena. We have created an initial model of an evolving social network, basing on recent research on social network evolution. Currently, we are investigating mechanisms leading to high clustering coefficient in social networks. We have also started working on the implementation of opinion dynamics models on evolving social networks.

Project Month 13–18

3.4.7 Biophysical Model for the Odra Case Study

During the last six months, the biophysical model for the Odra Case Study model has been under constant development and is now ready for integration with the agent based model and subsequent testing. The biophysical model is spatially explicit. It models such processes like water movement in the saturated and aeration zones of the soil, surface runoff, crop growth, interception of rainfall by canopies, evapotranspiration, channel water routing, maintenance and natural deterioration of land reclamation system etc. In order to fulfil all those tasks, the model requires a wide array of input data, most of which are spatial. These include the following: DEM of the modelled region, a vector map of the land reclamation system, a vector map of land parcels divided according to property and soil class, a crop database, a weather data file. The parcel map should contain identification numbers of agents that are responsible for given land parcels. These id numbers can later be used to integrate the biophysical model with an agent based model of farmers' behaviour.

The integration is possible thanks to a thin abstraction layer, that provides the ABM with the information that it requires and allows a feedback from agent's to the biophysical model. The layer defines two classes that are responsible for this communication, the agent's perceptions class and the agent's behaviour class. The former contains data that a farmer would usually get from his environment, like what parcels does he own, what is currently growing on those parcels, what can grow on those parcels, what is the condition of the land reclamation system and what were the recent crop yields. Agent's behaviour contains information about what a specified agent wants to do on their parcels. Thus, it contains such data like the kind of crops to be planted on every parcel and weather or not to maintain the land reclamation system. This information can be fed to the model at any time during the simulation and the model will take care of performing specific operation in proper time. This applies mainly to such operations like planting and harvesting crops, which have to be performed on specified dates for specified types of crops.

3.4.8 Binary Choice (Opinion Dynamics) Models

We proposed a general framework for individual-based binary choice, social interaction models. We show that under the assumption that an agent's choice depends on her/his previous choice, the economic models using the utility function and social psychology models using the impact function are mathematically equivalent. This equivalence allows us to use insights from both economics and social psychology to develop models applicable to a wider range of situations. Economic models with social interactions already expanded the limited homo economicus assumptions of neoclassical theory. Utility of agents is not only based on economic benefits and costs but also includes gains and losses related to conforming to others' choices. Psychological perspectives add inertia to agents' decisions, a certain tendency to follow last choice. Furthermore Nowak and Latane proposed to disaggregate the social interaction term into the two parts: supporting (sustaining the last choice) and persuading (pushing for change from the last choice). In general these social interaction terms can be nonlinear functions of the number of supporting or persuading agents. Nonlinearity of these functions allows us to model more realistic assumptions about agents' decision rules and can result in a richer equilibrium structure.

We introduced the so-called mean field approximation, where any individual choice depends only on the average choice of all other agents. Unlike other authors we do not explore the link with statistical physics directly, which allows us to obtain results for a larger class of models. Although mean-field approximation does not allow us to include the geometry of social relationships (social network), it becomes a significant reference model. It allows us to distinguish specific aspects of social space geometry which can generate different results. Within a mean-field approach both utility function models and social impact models can be formulated as 'threshold' models. Such conceptualization, first introduced by Granovetter, was recently developed by Levy who showed its equivalence with mean field utility function models. However both Levy and Granovetter considered transitions of individuals' choices only in one direction (adoption), such that reversing a decision is impossible. It is very useful approach for problems like innovation diffusion. We considered a more general framework where transitions can occur in both directions. In such cases simple graphical analysis used by Granovetter and Levy to obtain systems equilibria does not hold. We proposed a general procedure for analyzing equilibria in such models.

We proposed a generalized model for the binary choice problem. In this model one takes the utility function in a certain form: it is assumed to be an additive form of individual preferences, inter-personal influences, randomness and ‘self-supportiveness’ (individuals’ inertia). Individual preferences can result from personal characteristics or can be a result of external influence. In physics of phase transitions this component is called the ‘external field’. With these assumptions we obtain a more specific condition for systems equilibria. Also we discuss relation between some forms of randomness in individuals’ choices and heterogeneity of agents’ characteristics. As we relax some of the assumptions of the existing social interaction models (e.g. the form of the individual utility random term, heterogeneity conditions on some parameters) our results contribute to the synthesis of the field and expand possible applications.

We provided some examples of how, using specific assumptions, many binary choice models known from literature can be obtained from our generalized model. We also showed how different assumptions about the form of social interactions can produce different equilibrium structures. Specifically, variety different equilibria (we have presented an example with five different possibilities) can be obtained in one of the versions of Nowak-Latane model.

These results are presented in the article “Integrating economic and psychological insights in binary choice models with social interactions” which will be soon submitted to “Journal of Economic Behavior and Organization”.

3.4.9 Measures of Clustering of Agents’ Binary Characteristics

Measures of clustering in the binary choice models have been reviewed and tested. The analysis and comparison of the following different clustering coefficients have been performed:

- quantitative:
  - clustering probability (CP) and generalized CP,
  - analogy of clustering (transitivity) coefficient,
  - group-level and individual-level index of clustering
- qualitative - the likelihood of the difference from random configuration:
  - spatial autocorrelation analysis: Morgan’s I
  - join-count analysis
- cluster size: density and average cluster size

Description: The first two classes of clustering measures base on the idea of counting the pairs of neighbours that share the same attitude. The clustering reads the degree to which spatial neighbourhood and the decisions are correlated. In the random configuration of a given proportion of the binary decisions (say +1 and 1) the total number of pairs of nodes that are of the same sign should be significantly smaller than that in the configuration that displays clustering.

The third class bases on the intuitive definition of the cluster and points out the cluster-size distribution in the given network structure. It deals with the clusters physically present (in the sense that they are to be shown one by one) in the system. The clustering is described by means of clustering density and the notion of the average cluster size.

The aim was the comparison in order to and chose for further analysis a 'good' measure of clustering, i.e. such that indicates whether and/or how far from randomness the observed spatial pattern is. This condition is satisfied by the measures that use the statistics methods of hypotheses testing (join-count, spatial autocorrelation) and by the group-level index of clustering. However, the latter is not (so far) easily determined in non-regular network case.

3.4.10 Evolution of Social Networks

The difficulties of gathering evidence on exact structure of social networks in a case study directed as into an attempt to recreate a social network structure based on the rules of network evolution. Below we present preliminary results for very simple rules.

Modelling the network evolution we have made the following assumptions:

- the number of vertices is constant
- each vertex can join restricted number of another vertices, which is much smaller than the whole number of vertices – in our case it is about 10 – this number we call critical value
- building and evolution of the network consist of adding new connections (edges) and breaking some existing ones
- the probability of establishing new edge for two chosen unconnected vertices is bigger if these two vertices have common neighbour (a vertex to which both are connected)
- we have three groups of vertices, what can be thought of as three villages and each vertex can belong to only one group – the whole network consist of N=300 vertices, 100 vertices in each group

We analyzed our model in order to find parameters, which give us a network with social features: large cluster coefficient C (larger than that for random graph), positive degree correlation coefficient r, and a rather narrow than flat degree distribution with given average value.

The simulations were started with two kinds of initial structures: (i) an empty network, i.e. a network without edges, and (ii) a network with special initial configuration of edges where every vertex (except the first and last) has two neighbours with nearest indices in numerical representation (we can call it a regular graph). In the first part of network evolution – i.e. the construction of the network – we were working with the following probabilities values:

- for adding a new edge – 0.01
- for adding a new edge between vertices from different groups – 0.01
- for adding a new edge between vertices which have a common neighbour– 0.4 and 0.9 when they are from the same group

- for breaking the edge – 0.01 and 0.001
- for breaking the edge connecting vertices from the different groups – 0.001 and 0.0001
- for breaking the edge connecting vertices having common neighbour – 0.001 and 0.0001

After the average vertex degree has achieved its critical value, the values of adding probabilities were multiplied by 0.0001 and the evolution became slower and more stable. We noticed that in the rapid initial part of building a network the values of C, r and <k> were first increasing and then falling. We needed to find appropriate moment in iteration to slow down the evolution (or in other words – to stop building stage). The transition from fast to slow evolution was done in two ways: A – sharp, when the average vertex degree of the whole system achieved the critical value; B – broad, when the evolution is slower only for these vertices, which degree achieved the critical value. In the case A we got networks with wide distributions of vertex degree, while in the case B – with very narrow ones. In social network the distribution is supposed to be narrow. In the case B the degree correlation coefficient r behaves more chaotically than in the case A during the evolution and sometimes is negative, which means that the network is not social. In the case A the values of r were in the interval 0.14-0.35. The values on cluster coefficient C were in the intervals: 0.19-0.27 (B) and 0.25-0.33 (A), and were slightly decreasing during the evolution. In comparison, for the corresponding random graph, cluster coefficient took values 0.03-0.06.

The structure of the initial network had no essential effect on values of C and r. But we found that it had an influence on the structure of the evolving network, even after a large number of iterations. In the evolving structures, the ratio of number of connections between pairs of vertices from different groups to the whole number of connections was from the interval 0.4-0.6 – in the case 1 and 0.1-0.2 – in the case 2.

Our simulations were performed within not very long time – the number of iteration can be compared with a period from a few to several dozen years. Now we are studying the stability of achieved networks and this purpose needs much longer simulations.

3.4.11 How to Proceed

The development and integration of the Odra biophysical model with the agent-based model developed by the Kassel team will continue to be an important task for the next six months.

The results established for binary choice models in the mean field case will be compared in case of diverse social geometries (network structures). Clustering of agents subgroups will be investigated for such models.

Modelling of network evolution will be continued with an attempt to incorporate evidence from case studies.

Initial analysis of models produced by other modelling teams will be performed in order to proceed with appropriate generalizations.

4 Case Studies

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

4.1 South African Case Study (SEI Oxford)

Project Month 1–6

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 19 illustrates the Role class hierarchy used in the model. The reader is referred to section 3.3.1 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.

Figure 19: 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 20. 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.

4.1.5 Links with Other Projects

CAMP project: Catchment Management and Poverty[22]

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

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

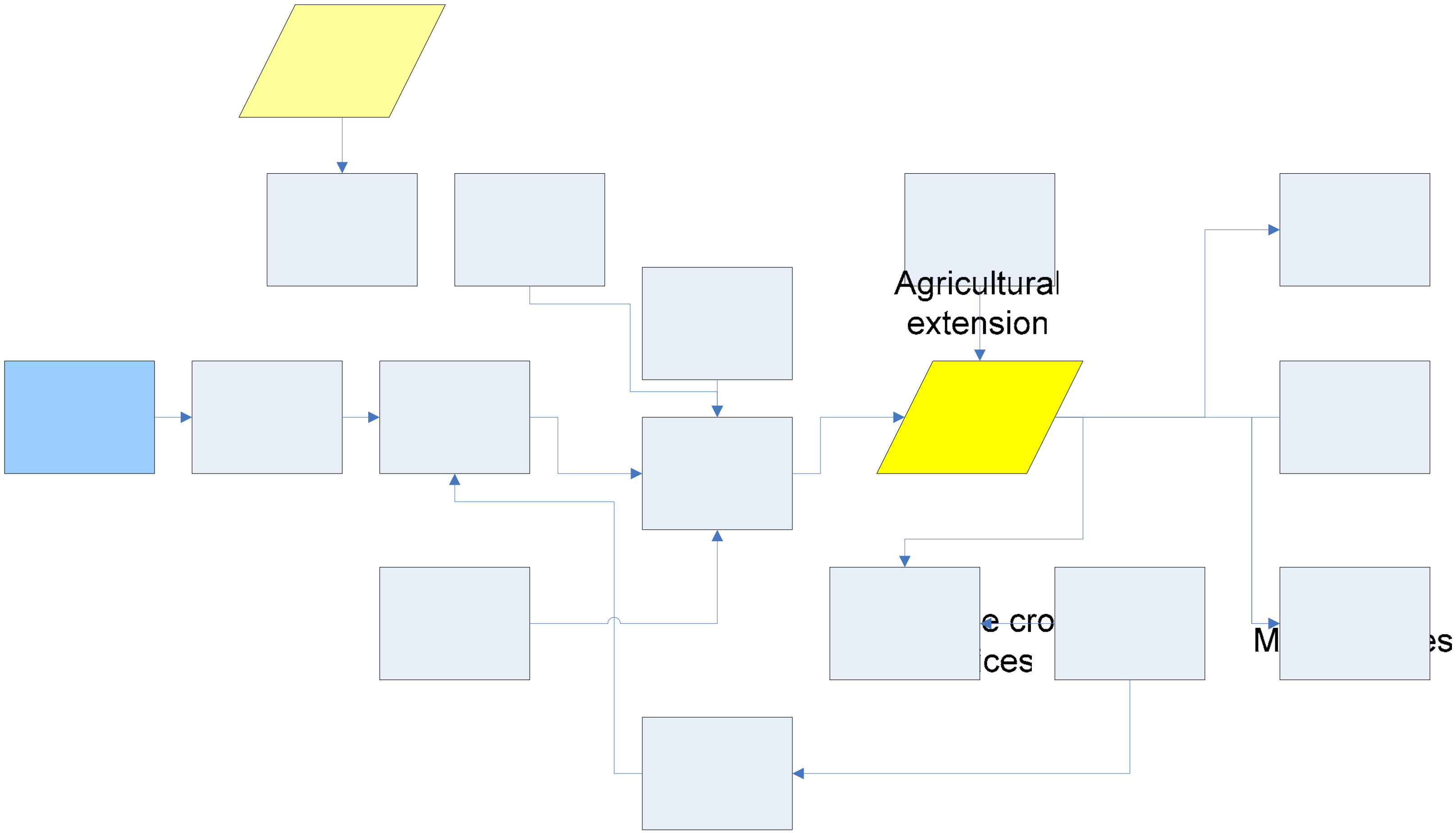

Figure 20: A conceptual diagram of the impact of drought on small scale farmers in Mangondi.

- 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

Project Month 7–12

4.1.6 Introduction

Due to the evidence-based nature of this project, a fieldtrip was carried out in February 2006 to ascertain the key drivers and stresses in the chosen case study area of Sekhukhuneland. Much data and background work has been done in this area which has helped to inform the initial pilot model. Collaboration with the previous FIVIMS (Food insecurity and vulnerability information mapping systems) project was established and with the RADAR project to explore collaboration using their data that relates to HIV/AIDS and social networks.

The fieldtrip highlighted some key issues which now need to be integrated with the next round of modelling. There is also the need to focus the future fieldwork in order to provide more detailed information to feed into the model and in order to communicate with stakeholders as to their needs.

The proposed outline will be presented to stakeholders to establish whether this approach might be of interest and use to them in their planning.

4.1.7 Case Study Focus

There were a number of issues that emerged as important and relevant to development and livelihoods in the Sekhukhuneland area. Water is a key stress at the local village level, as it constrains daily activity and development such as field-based agriculture and other activities such as home gardens, projects that require water such as brickmaking and livestock. Water is also a key stress at the district level. There are separate schemes for agricultural and bulk water and both of these are strained. There is talk of decreasing agricultural water quotas to supply more water to the bulk schemes. Development is being undertaken to improve the supply of bulk water as the district and regional hospitals both had to have water tankered in last year. There are a number of agricultural schemes that are aimed at poorer farmers that require water. Mining is an important part of the Sekhukhuneland economy and it is rapidly growing. This has implications for water use. A large dam has been planned and construction is expected to start this year yet there are environmental concerns about it.

Another key stress in the area is employment. Over 90% of the population is rural and there are very few jobs in the villages. People therefore have to migrate to find employment. It is seldom that households migrate but rather it is the individuals who migrate. This migration appears to be based primarily on social networks as people migrate to areas where they have friends or family they can stay with while they are looking for work. When they do find a job they send remittances home and visit their village a minimum of twice a year. Although due to urban living cost their remittances are sometimes infrequent. The mines also form a part of the employment opportunity although have associated costs. Many people from the district move outside the district to find work as that is where their connections are based. Many of the people securing work on the mines are from outside the district as they have the necessary experience. There is a move to support training of local people so that they can become more qualified and their can be increased local employment. Associated with the influx of migrants from outside the district, is an increase in HIV/AIDS. This has implications for local social networks. The water stress also has an impact on health and so health becomes a key concern in this nexus. At the same time there is the suggestion from villagers that health has become worse because of the food people eat. They are unable to grow enough of a range of vegetables to remain healthy as their grandparents used to in the past.

The water and employment issue related closely to food security as it appears that this district is not able to feed itself. Households therefore need to secure income to secure their food. Water availability could help to supplement food supply.

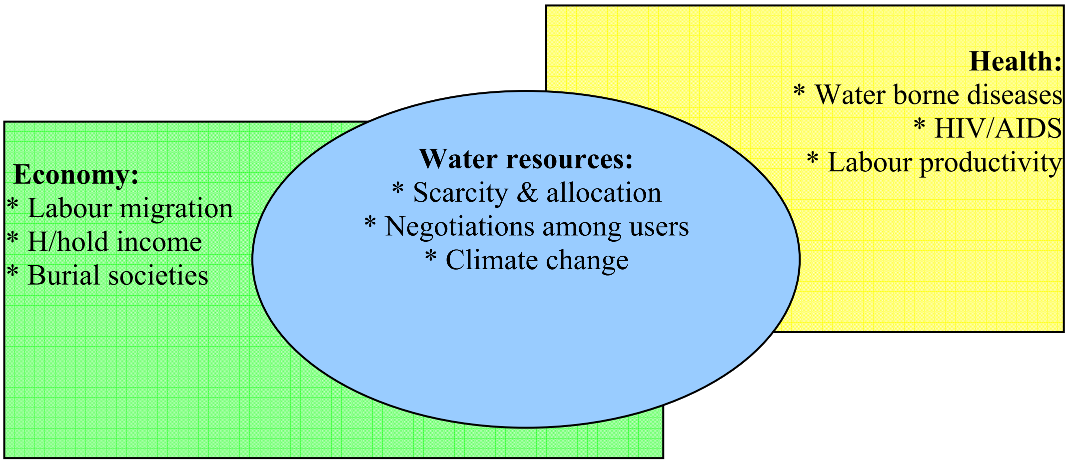

Figure 21: Schematic of areas of interest that emerged out of fieldwork and could be modelled.

4.1.8 Social Networks

Social networks are of interest to the CAVES project. Certain areas of interest with regards to social networks emerged during the fieldwork and are presented below.

Traditional institutions

Marriage

People marry both from within and from outside the village. If they marry someone from another village, they decide together which village they will live in. Traditionally they would have lived in the husband's village. Marriage appears to be a significant investment as money or other goods known as *lebola*, have to be paid to the bride's family. It is also necessary to have a big celebration where the whole village is usually invited.

Burial societies

Most people in rural villages belong to burial societies. These societies have different forms, with some being more formal than others. The banks run funeral policies but most people were involved in village based schemes that often seem to cover a number of villages in the area and village members serve on the committees.

Burial societies also have an important social function. The society meets once a month usually and matters relating to the society are discussed. After this, a social gathering is held which strengthens networks. Enrolment in these societies also seems to be the one activity that people do as an investment. It is interesting that death is the one certainty and so this is perhaps seen as insurance so that the family can manage the future after the expenses of funerals and feasts.

Food

People tend to borrow food from neighbours if they do not have sufficient. People might go to their neighbour's house and eat there or 'borrow' food. This food is usually not taken back but means that if the other household is in need of food they can expect reciprocity.

Activity/employment

Social networks seem to be key in accessing employment. People mentioned that they look for work outside the village in areas where they can stay with people they know. Even though there are mines opening up in Steelpoort, an hour and a half away, people from Mohlotsi are more likely to travel to Gauteng, which is further, as that is where they know people. When they are looking for work, the person they are staying with will host them and they do not have to pay for accommodation but will be expected to return the favour when they do find employment. Information about jobs reaches people where there are strong social networks. People said they do not spread the word widely about new jobs, but someone might send information about job prospects to their close friends and family back in the village.

Community projects

People are keen to get involved in community projects. The youth seem to be primarily interested in community projects if they can access an income through them. Older people would like to get an income but also seem to want to be involved if the project can provide produce. Communal gardens were therefore of interest to the youth only if they would be able to produce for market. In the villages we visited, water is a key constraint and so although they were interested in establishing gardens, they were battling to secure water. Many community-based projects seem to focus on supporting the youth or women. These projects seem to hinge on key individuals and at times, those individuals' links to higher levels. For example, if the headman supported a project and had connections to the municipal level, the project was more likely to get going.

Family networks

There are strong family networks. Families who are related might meet once monthly to discuss issues and socialise. Often children live with their parents even when they are married and have children. The grants received by individuals are seldom used only for individuals but tend to be used to purchase food for the whole household. Household members often seem to relocate in order to cope with stress. For example, in one household the children had been sent to live with the woman's sister who was able to care for them. In another case, the household members who were living together were split up when the daughter received an RDP (Reconstruction and development programme) house.

Linkages across scales

All villages have a traditional head. In most villages there is a headman, usually established through family networks and birth rights. These headmen are under a chief, who will oversee his own village as well as a number of surrounding villages that will have headmen, although one might have a foreman who has more status than the other headmen and helps to support the chief. These chiefs report to the ward counsellors. Ward counsellors are supported by the municipalities. The ward usually consists of a number of villages. This system enables the municipalities to get feedback from the villages without having to communicate with each village separately as the ward counsellor is responsible for that. The ward counsellors report to the municipality. There are a number of municipalities in a district and a handful of districts in the province and 9 provinces in South Africa.

4.1.9 Conclusion

In order to use the RADAR data it is necessary that the modellers meet with the RADAR project coordinator to ensure that there is a common understanding of the data and how it will be used. This will also enable to modellers to understand the context of their model and to engage with potential users of the model. This will be undertaken in a fieldtrip in June.

A masters student will be supervised by the SEI. Her dissertation in Sekhukhuneland will focus on water and the associated allocation rights and governance issues. She will use the WEAP model to frame her work and feed this information back to the modellers.

The fieldwork has enabled a more focused pathway to be identified. This needs to be developed further with the modellers and case study experts to ensure the modelling addresses the methodological challenges it set out to explore at the same time as representing the case study material in a way that integrates multiple stresses and social networks and can be validated by stakeholders.

Project Month 13–18

4.1.10 Recent Activities

Over the last 6 months, there have been two foci in the South African case study work. Firstly, working with the MMU team and RADAR to establish a focus using the RADAR data to inform a model of food insecurity in Limpopo province, South Africa. Secondly, writing up and analyzing the data and information from the February fieldtrip.

Linking to RADAR

A meeting was held in Oxford in June 2006 that brought together the MMU modellers, the South African case study researchers, a RADAR staff representative (linked to University of Witwatersrand and London School of Hygiene and Tropical Medicine) and another colleague from South Africa with experience in food security policy.

Arising from the meeting was a verbal agreement that the CAVES team could use the RADAR data if they focused on areas that had not been explored in the current RADAR analysis. The focus of CAVES on understanding complexity through an analysis of food insecurity, social networks and environmental factors was therefore of interest to the RADAR team. The RADAR team has excellent data on social networks, coping strategies and other factors that supports investigation of these themes. The text below outlines how the CAVES project will collaborate with RADAR.

Rationale for using RADAR data for South Africa case study

The CAVES project seeks to model interaction between agents and the environment using agent-based modelling approaches. This approach allows for the development of interaction among agents based on evidence, facilitating emergent behaviour. This emergent behaviour can identify new patterns and characteristics of the system and potentially guide further investigation. The model would draw on the RADAR study that addresses issues around HIV/AIDS, social capital and gender-based violence. Through their work, the RADAR team have gathered large amounts of data around numerous areas related to individual and village-level characteristics. This data would drive the design of an agent-based model.

The model would focus on issues around food insecurity. Project members have been involved in issues related to food insecurity and vulnerability to climate, water and health stress in previous work in Sekhukhune district (Poverty and vulnerability project: Adapting to multiple stresses and FIVIMS-ZA: Food insecurity and vulnerability information mapping system). Models designed with the RADAR data, being both formal and couched in the qualitative language of stakeholders would lend increased precision and clarity to investigations of the dynamics between social capital, livelihood assets and food security. The models would be validated in relation to data on income, wealth, HIV/AIDS status and household structure. Further issues around migration and food security are also of interest and the model could draw on the RADAR and FIVIMS-ZA data to develop this component. The simulation models will incorporate physical models to determine water availability and land use as core elements of the environment in which social vulnerability dynamics emerge.

The CAVES modelling process is intended to contribute to policy-relevant debates. The above focus is pertinent to current policy and intervention activities in South Africa. It is hoped that the CAVES model output will provide lessons and insight into food insecurity dynamics that can be fed into policy networks such as Renewal, the South African Government Social Cluster Food Security Group and FIVIMS-ZA.

Analysis of fieldwork data

A comprehensive report is available that analyses the qualitative data, from village and district level interviews and focus groups as well as quantitative analysis from the Discrete choice analysis (DCA) and Stated preference (SP) technique.

The conclusion from the report, Adapting to climate, water and health stresses: insights from Sekhukhune, South Africa, reads as follows:

‘Focusing on water, climate, health and food security this project has investigated how vulnerability is configured and what decisions people make regarding adaptation strategies in response to a multitude of stresses, which interact in complex and dynamic ways. The findings indicate that water scarcity and limited economic opportunities are two major constraints to development at both the village and district scale that undermines adaptive capacity. People to a large extent relate both climate stress and food insecurity back to these two dominant stressors. This highlights the need for integrated responses to support local adaptation that departs from hegemonic sectoral approaches. It also requires improved and increased communication between government and local communities to facilitate the integration of strategies being implemented at different scales and better align expectations. Municipal government needs to carefully assess the likely environmental, social and economic impacts of investing in different sectors, incorporating a view on climate change and prioritizing water saving and wealth distributing options with the aim of minimizing future vulnerability.

This study has shown the benefits associated with combining qualitative and quantitative methods in adaptation research. In addition, enormous value and insight has been gained by linking this work with other ongoing research programmes in the area and extending research activities to include individuals at the village level as well as local and district government officials, thereby addressing important scale issues associated with vulnerability and adaptation.’

4.1.11 Next Steps

The MMU pilot model using the RADAR data will be evaluated to establish which areas will be investigated further. The completed report from the February fieldtrip will also provide additional information to steer the direction of the model. There are a number of potential areas of interest to explore further that depend on the initial results. Some of these areas include links between social capital, livelihood assets and food security and how they link to income, wealth and HIV/AIDS; household structure; role of water related to tribal and civic politics; migration and food security and developing links to policy. Further fieldwork will be discussed at the September CAVES meeting.

4.1.12 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 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)

Project Month 1–6

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.

Project Month 7–12

4.2.5 Progress on Literature Review

Research Questions

1. How has (agricultural) land use at a study site in North East Scotland changed over the past 20 years?
2. Why does agricultural land use change?
3. What is the role of land users' social and informational networks in this process?

Question One

Dr. Small has completed a preliminary review of background literature on data sources on land use change in Grampian, including data from interview-based surveys of farmers in the region. She utilised these findings primarily to identify a field research area, and as context for the development of an interview guide. Ongoing contextualisation of study findings, through synthesis with literature on land use in North East Scotland, will occur throughout the research.

Questions Two and Three

Dr Small has completed a preliminary literature review of social networks as studied in the social sciences. She began by reviewing the history of social network studies, before identifying four major approaches to the study of social networks: social network analysis, actor network theory, social capital, and social networks as studied in the rural sociology literature. She identified the primary assumptions of each of these approaches and compared them to the assumptions of the CAVES project, and the Grampian case study in particular. Recognising that there is no 'perfect fit' within the sociology literature for the CAVES project, she determined that the concept of 'social capital' would be most useful as a theoretical approach for structuring the research in the Grampian case study.

Social capital is an interdisciplinary concept widely accepted in the social sciences. Social capital has varying definitions, dependent on the major theorist utilised, but in general is considered to be "the features of social organization … that can improve the efficiency of society by facilitating coordinated actions" (Putnam, 1993, p. 167). Social capital may include trust, social norms, social credentials, information channels, family relationships, voluntarism, group membership and community engagement. While typically recognising that social capital is held by groups or within the context of relationships, rather than by individuals, studies of social capital do not usually address specific types of network structure. Studies of social capital vary widely, ranging from quantitative analysis of standardised survey data to qualitative description of social norms.

For the purposes of the Grampian case study, the flexibility in definition and use of the social capital concept is beneficial, allowing integration of the diverse ideas embedded in the project documentation. The concept fits well with the proposed qualitative field research, and theoretical linkages to economic decision-making. Use of the social capital concept allows the focus of field research to remain on the role of social networks within land use change processes, without the identification of whole or ego-centred networks. As such, social capital enables the development of a more feasible field research approach (under the conditions of the Grampian case study) than a more formal 'social network analysis' of nodes and ties. The literature on social capital is sufficiently wide ranging as to include all of the issues identified in the CAVES project documentation, but care will have to be taken to ensure theoretical consistency when bringing these ideas together to define social capital in the project. Some work will also need to be done in order to integrate social capital with complexity theory.

4.2.6 Research Plan

Dr. Small has developed a research plan to address the research questions. She will undertake qualitative interviews with 50 – 60 land managers, including 10 – 12 'successors' (individuals intending to become the primary land use decision maker in the future), plus approximately 20 key informants (agricultural business-people and agricultural stakeholders). This range of interview respondents will give depth to the study, through the identification of different perspectives on the process of land use change. The interviews of both primary land managers and their successors are expected to result in the identification of different networks and social influences on land use change, which will assist in identifying 'networks of networks', as well as identifying changing patterns of land use in both the past and anticipated future.

Upper Deeside has been chosen as the study site location. The area stretches from Kincardine O'Neil to Braemar, with half of the territory within the boundaries of the recently established Cairngorm National park. The area is fairly large, but this was considered necessary to include a broad range of land use holding types and patterns. In addition, holdings within the study area are often quite large, due to the large estates in the area and extensive agricultural production characteristic of hilly and semi-mountainous land. The large study site is also more conducive to the 'snowball' sampling technique, whereby initial respondents identified members of their network for further research. Pilot testing demonstrated that these connections are often somewhat distant. The initial sample is taken from two key informants – one with access to large scale farms and estates, and another with access to small-scale farms. The combination should ensure a wide range of study respondents.

The interviews will be structured utilising an 'interview guide' of issues relevant to the research questions, rather than a formal questionnaire. This will allow respondents to speak openly about what they view as the most important issues surrounding land use change, and the types of social networks in which they are involved. The interview guide, currently being pilot tested, addresses the following topics:

- Land holding characteristics (scale, commodities produced, management structure)
- History of land use change, and new land acquisition on the holding
- Largest change in land use during the current land manager's tenure
- External shock of greatest significance during the current land manager's tenure
- Process of land use decision-making
- Response to the Single Farm Payment (major policy shift significantly impacting farm incomes)
- Group membership and community participation
- Sharing behaviours – information, labour and equipment
- Anticipated future changes to land use
- Respondent demographic characteristics

The interview guide is expected to evolve over the course of the research, as other important areas become identified. Study results will be utilised to answer the study questions, and identify decision rules for land use change and transactions. These decision rules will in turn be utilised to guide model development.

4.2.7 Field Research

The initial stage of field research has begun, with interviews of seven farmers, three estate managers and two key informants completed to date. These interviews form the 'pilot study' – preliminary work to evaluate the utility of the interview guide and general research approach. Some minor changes have been made to the interview guide, but the general approach appears to have been successful. Interviews have ranged between 45 minutes and 2 hours in length. Analysis of these interviews is underway. Potential respondents have been identified for a second wave of interviews, to occur in March.

4.2.8 Stakeholder Validation Protocol

The 'triangulation' of research findings through identification of three different data sources (in this case: land managers, land manager successors and key informants) is a well-established means of academic validation of study findings. In addition, we are planning a workshop for winter 2006, in which study respondents will be invited to respond to the decision-rules identified by Dr. Small as arising from the field research. This workshop will occur before data analysis is complete, to ensure that the study is truly responsive to respondent feedback.

As part of the research process, several additional stakeholders have been identified. These include staff from the Cairngorm National Park Land Use Division, the Aberdeenshire Council Rural Development, and the National Farmer's Union. Dr. Small will interact with these individuals on a bi-monthly basis to discuss the research process and findings.

Project Month 13–18

During the March – August time period, Dr. Small completed the pilot study data collection, preliminary analysis of that data, and the bulk of primary field research. A summary of pilot study findings was circulated to stakeholders and the CAVES research team for feedback in late May. An introduction to the CAVES project and preliminary results of the pilot study have been presented in poster format to two conferences; a paper has been written for presentation to a third conference in late August.

4.2.9 Pilot Study

Field research for the pilot study was completed in April. For the pilot study, interviews were conducted on 12 farms and 3 estates in the Upper Deeside Region (Finzean to Braemar). These were supplemented by interviews with four key informants (professionals working in the agricultural industry). Analysis of pilot study interviews is intended to identify possible outcomes of the broader study, and the extent to which research questions are being adequately addressed.

Research Questions

1. How has (agricultural) land use at a study site in North East Scotland changed over the past 20 years?
2. Why does agricultural land use change?
3. What is the role of land users' social and informational networks in this process?

Land Use Change

From the pilot study, it was clear that although the Grampian region has experienced less in terms of land use change than the Poland or South African study sites, there have been subtle changes in land use on Upper Deeside farming operations over the past 20 years. These include: increasing scale of operation and intensity of livestock production, reduction in number of commodities, decreased use of inputs to arable land and increasing participation in environmental programming. Estate managers reported actively encouraging their tenants to engage in environmental programs and to increase the scale of their operations. As a result, total numbers of tenancies have reduced and there has been limited development of hobby farming in highly tenanted areas.

Causes of Land Use Change

The primary reasons given by land managers for changes in land use were economic – the perceived necessity to respond more efficiently to market and subsidy trends, in order to maintain profit margins. Increasing mechanisation and reduced farm labour availability were also of importance. Due to the nature of agricultural production, however, changes in land use and farming operations in general do not respond immediately to changes in economic signals. Agriculture is highly based on seasonal and climactic factors – land managers perceive it to be unfeasible to make rapid changes to either livestock or crop production. Land managers also believe that commodity markets follow cycles, and that quick response is imprudent. Similarly, land is a scarce resource and therefore must be acquired when it becomes available, not according to a long term plan. Thus, land use change processes are both slow and complex.

Social Networks in Land Use Change

The role of social networks in land use change was initially considered on four levels: access to information, social norms, resource sharing and community engagement. From the interviews, it became clear that the reputation of the farmer is important for securing access to rented land. This appears less true with regard to buying land. Land managers reported accessing information from discussion groups, printed publications, SAC advisors, contract workers and input salespeople visiting the farm, informal farm visits, international farm visits and general observation of other farmers' activities. They also identified social norms about the meaning of being a good farmer. They reported that sharing of labour and machinery is limited, but has increased in recent years due to financial necessity. All of the interviewees were members of the National Farmers Union, and most had other active community involvement. Based on the pilot study research, it is difficult to say what impact social networks are having on land use change, as although the respondents varied in terms of social network engagement, they appeared to be making fairly uniform changes in their land use.

Outcomes

From the pilot study data analysis, it was clear that the research questions were largely being addressed. However, the lack of direct comment made by respondents about the relationship between their land use decision-making and social networks made necessary an increased focus on this linkage in subsequent work. Pilot study participants were also identified as predominantly large-scale, progressive style farmers; efforts were made to include a broader range of participants in the field research that followed.

4.2.10 Primary Field Research

Primary field research began in June. The number of interviews completed to date (including the pilot study) is 39 (44): 24 farmers (29 including the spouses or partners also participating), 4 identified successors (to farmers in the study), 3 estate managers, and 8 key informants. This number is somewhat lower than initially projected, due to low respondent availability in July; this is offset by the higher than anticipated number of joint interviews. The increased emphasis on social networks included participant observation of key farmer interaction settings: the local agricultural Mart, a Young Farmers Event, a regional agricultural show and a local agri-environmental initiative. To reflect the inclusion of these research activities in the research plan, the total number of anticipated interviews has been reduced to 45 land managers (approximately 30 farmers, 8 successors and 7 estate managers) (from 50 – 60 land managers), and 15 key informants (from 20 key informants). The validation protocol will be undertaken largely with individuals who have already been interviewed.

Data analysis of the primary field research is in progress. The increased emphasis on social networks has resulted in the identification of multiple types of network within the study site, largely based on commodity type, and attitudes towards change.

4.2.11 Conferences

Dr. Small presented a poster entitled ‘Introducing CAVES: Complexity, Volatility, Evidence and Scale’ at the Participatory Approaches in Science and Technology conference, held in Edinburgh, Scotland June 4 – 7, 2006 and at the Rural Sociological Society Annual Meeting in Louisville, Kentucky, August 10 – 13, 2006. Dr. Small will also present a paper entitled ‘Modelling Social Networks: An Application for Rural Sociology?’ based on methodological considerations in data analysis, at the XII International Summer School, "The Empire of Rules and the Ethics of Social Networks" organised by the Institute of International Sociology of Gorizia, in collaboration with European Society for Rural Sociology, August 28th – September 5.

4.2.12 Directions for Future Work

Primary activities for the next six months are as follows:

- Development of methodology for utilising field research in modelling. Dr. Small is in the process of working with other team members to develop a methodology for deriving model components from field research. Utilising qualitative field research data for agent based modelling is a relatively new area of academic work, which could benefit from the development of novel approaches. This protocol will include specific validation techniques for interacting with land use decision-makers.
- Ongoing data analysis of field research to date. This will occur in primarily in

September and October, while farmers are busy with harvest.

- Completion of primary field research. Further interviews will occur in October, to reach the initial target of 45 interviews. These are expected to include primarily successors and estate managers.
- Implementation of respondent validation protocol. This will occur in December/January.
- Initiation of the fourth wave of field research. This is expected to being in March, to reflect responses to the validation protocol.

4.2.13 References

Putnam, Robert D. with R. Leonardi and R. Nanetti (1993): Making Democracy Work – Civic Traditions in Modern Italy. Princeton University Press, Princeton, New Jersey.

4.3 Odra Valley case study (Uniwersytet Wroclawski)

Project Month 1–6

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[23].

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.

Knowledge elicitation – field trips, interviews and workshops

During field trips we watched the study area to make general reconnaissance and

[23] 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.

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 22) 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.

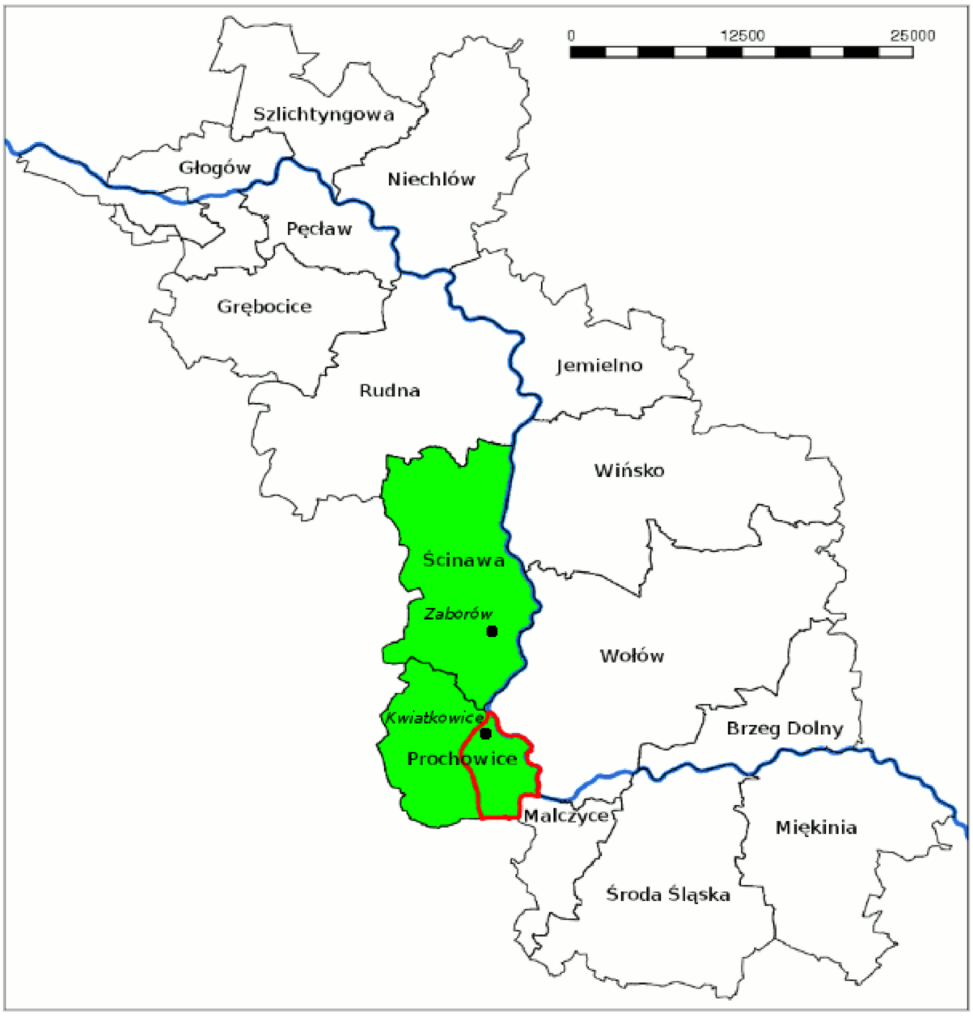

Figure 22: A map of the Odra Valley case study area, showing the administrative division. Communes of special interest are coloured green. The red polygon indicates the area of the social aspects of land reclamation model proposition. Villages where workshops with local community took place are indicated.

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 modelling (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 Description of the Work Undertaken

Knowledge elicitation

Decision rules – field research

In order to collect sufficient data for modeling team, the research tool – semi-structured interview questionnaire (in two versions: for experts and landowners) was created. Included questions concern following issues: land use changes in the past, decision rules regarding land use, land reclamation system maintenance and collective action, social networks, future scenarios. The questionnaire was tested and discussed with experts.

Interviewees (landowners) in Rogow Legnicki village were identified (names and addresses, phone numbers) according to landownership maps, data form local authorities and other sources of information. Following decisions made at the Wroclaw meeting in September 2005, we have chosen landowners owing the land along Kwiatkowicki and Rogowski Canals.

The covering letter was prepared and distributed among interviewees in order to get permission for interviews. Technical aspects of field research were fixed (preparing required documents, printing questionnaires, buying dictaphones etc.).

Interviews were conducted in February. All listed addresses were visited. We managed to conduct 14 interviews. Five persons refused to give an interview. The rest was unavailable; from their neighbours we do know that most of them moved away to other localities and do not cultivate land anymore. Additionally two long (up to 4 hours) and very detailed expert interviews were conducted.

Formal rules – legal acts analysis

Legal acts concerning land reclamation were collected and analyzed in order to find formal rules of land reclamation system maintenance and related collective action. We regard these rules as important, since all landowners and relevant institutions must follow them. These rules indicate the hierarchy of responsibility and the scope of decision reference of our potential agents. The results of this analysis were described in a document „Legal Framework for Land Reclamation and Collective Action in Water Partnerships".

KnETs

The work on knowledge elicitation tool (KnETs) for the Odra case study has been started (according to a template sent by Sukaina Bharwani).

GIS

A 1:5000 map of the village Rogow Legnicki was scanned, geo-referenced and integrated into the Odra river case study GIS database as a backdrop map for digitization of following thematic vector maps, which will be the input for a development of the land reclamation model:

- land reclamation system (canals, ditches, water bodies, culverts),
- land ownership and land use along the Rogowski and Kwiatkowicki canals in Rogow Legnicki.

In order to facilitate understanding of land ownership and land use in the Odra case study area, consultations with local authorities and experts were held, as well as legislation and literature regarding GIS based ecological-economical modeling in the agricultural landscape were studied.

The methodology for field verification of LULC and land reclamation system maps, derived from cartographic and remote sensing sources, was developed.

In cooperation with Wroclaw University of Technology, the TOPMODEL hydrological model software was evaluated to asses the TOPMODEL usability as a basis for Odra river case study biophysical model.

Further sheets of the contemporary 1:10 000 topo maps where obtained, georeferenced and stored in the Odra river case study GIS database. The maps are being used as a backdrop for digitising elevation contour lines and points in the catchments encompassing Rogow Legnicki.

A preliminary digital terrain model (DTM) was interpolated using the above data, to be utilized in the biophysical model developed by Wroclaw University of Technology. Different interpolation algorithms were evaluated. Due to numerous spots of high curvature and prevailing low denivelations in the study area terrain, the natural neighbor interpolation has proved to yield best results.

Other issues

Meetings

- Organisation and coordination of CAVES project partners meeting held at the

Wroclaw University, Wroclaw 2005 Sept. 27-30, including the presentation "Evidence collected so far", as well as the field trip with accompanying field trip guide "Odra River Valley Case Study - Policy Relevant Issues for Modeling". Accomplishment of formalities connected with the CAVES Meeting set off.

- The whole UWr team participated in the Wroclaw CAVES Project Meeting 26-29.09.2005.
- The whole UWr team is going to participate in Aberdeen CAVES Project Meeting 13-17.03.2006.

Data for Kassel modelling team

A document "Information Package for Modeling Team No. 1" was prepared and sent in order to answer the questions from Kassel modelling team concerning floods/droughts, crops, knowledge/skills. Initial GIS data package for Kassel University modelling team was included.

Project management

Large amount of organisational and office work connected with project management is being done continuously (preparing documents – invoices, reports, agreements, project description for www site etc., contact with University offices, purchase – tickets, stationery etc.).

Technical maintenance

Our team learned CAVES CVS server usage for data exchange with CAVES project partners. Besides, the Odra case study staff computer network and workstations where maintained due to moving to another work place. GIS software (Grass 6.1, GDAL, PROJ.4, Quantum GIS) was updated for stability and performance improvement.

Project Month 13–18

4.3.5 Description of the Work Undertaken

Social research

Family and neighbourhood networks in the study area were identified based on documents provided by local authorities as well as personal visits in the Rogow Legnicki village. These are "one to one" networks. The links between all agents were found. These data were inserted into a database in Excel and together with explanatory note sent to the modelling team in Kassel as Info Package 3.

The team from WUT had a number of meetings with a new colleague from the Institute of Social Sciences in order to work out the concept of further research on networks and its connection with modelling.

The analysis of interviews performed during the first part of field research in winter is still in progress. The open coding is finished and the "codes book" is ready and was the subject of discussion during a joint workshop with the modelling team in

Kassel. Thirteen narrative storylines (2-3 pages for each agent) were prepared in two versions, Polish and English, and sent to the modelling team and experts.

The paper on knowledge elicitation methodology in the Odra case study was prepared and put into the knowledge base on the CAVES internet portal.

The KnEts tool for the Odra case study has been prepared in collaboration with Sukaina Bharwani from SEI, starting at the last project meeting in Aberdeen and continued by means of discussions held via email and Skype.

A paper on the integration of GIS analysis and social research in the Odra case study to be submitted to Landscape Ecology journal is currently being prepared. The same holds true for a new, modified version of the interview questionnaire.

Work with resilience issues regarding the Odra case study has started and is in good progress now. Following the guidelines in the report by Jan Sendzimir (cf. Deliverable No. 6), the conceptual model of the land reclamation system in Rogow was prepared.

GIS research

In the course of social and GIS fieldwork, the existing land ownership and land use data were revised and additional data were collected. The GIS database was updated accordingly.

The digital elevation model (DEM) for the biophysical model, derived from data digitised from 1:10000 topographic maps, is finished. Among other means for evaluating its quality, a sample photogrammetrically created DEM, covering a fragment of the case study area, was purchased from WODGiK (the region's geodesy and cartography agency) and integrated into the GIS database.

Final versions of the two script extensions (programs) for Grass 6.x GIS, mentioned in the previous report, developed to accomplish preparing data for the biophysical model, where published on the Grass WIKI site. These are:

- `r.surf.nnbathy`, see http://grass.gdf-hannover.de/wiki/GRASS_AddOns#Raster_add-ons
- `v.flip`, see http://grass.gdf-hannover.de/wiki/GRASS_AddOns#Vector_add-ons

Land cover dynamics research has been started. The first of the three available land cover time series has been digitised from the most current 1:25 000 topographic maps (1993/94).

4.3.6 Next Steps

Social Research

- Second part of field research (September)
- Field research results analysis (October)
- Field research full report (November)
- There is the proposition of validating the research results during the meeting with stakeholders approved by our expert, Andrzej Ruszlewicz.

GIS Research

- If the photogrammetrically created digital elevation model (DEM) proves more suitable for the biophysical model, creating the whole DEM by means of photogrammetry will be considered.
- The remaining two time series (1982/3 and 1930's) of land cover will be digitised. The time series will then be assessed for change detection using GIS.

5 Deliverables and Milestones

All Deliverables and Milestones for the first 18 month period of the CAVES project have been met. There have been some minor changes to a subset of the Deliverables (see section 5.1), including due dates (see section 5.2).

5.1 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:

No	Deliverable name	WP	Modified name	Re-allo-cated 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 Odra River valley	3	Prototype of the initial model for the Odra 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[24]. 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.

5.2 Current Deliverables

There are eight Deliverables due within the first half of the project. These are the initiation report (project month 6), the internet portal (project month 9), a knowledge base (month 12), a report on resilience theories and measures (month 12), prototype models for each of the case studies (month 12 and 18) and a working paper on case study structure, stakeholder/agents and validation data (month 18).

Since work on the Grampian prototype model was delayed due to the necessary enhancement of the underlying modelling software (FEARLUS), it was decided to swap the due dates of Deliverable No. 4 (Grampian prototype) and Deliverable No. 7 (Limpopo prototype model). Thus the prototype model for the Grampian case study is now due in project month 18 and the one for the Limpopo case study was due in project month 12.

Considering the current progress in the case studies and the respective model development, we anticipated a delay for Deliverable 8, the working paper on case study structure, stakeholder/agents and validation data, as reported in the progress report at month 12. This will enable us to base the working paper on experience instead of an ideal outline, which will enhance the value of the Deliverable significantly. At the last consortium meeting in Oxford it was therefore agreed to postpone the due date of Deliverable 8 to month 24 at the latest.

The following Deliverables are available now:

- Deliverable No. 1: Initiation Report. This report is publicly available from the CAVES internet portal (see below).
- Deliverable No. 2: Internet Portal. The CAVES internet portal went on-line in project month 3 and has since been enhanced considerably. It consists of a publicly accessible part and a part with restricted access for eligible users. The URL is `http://caves.cfpm.org/`.
- Deliverable No. 3: Knowledge Base. The knowledge base contains briefing notes on foundation topics for the CAVES project. These are publicly available as part of the internet portal. The first version of the knowledge base has now been replaced by a Wiki[25]-based component.
- Deliverable No. 4: Prototype of the initial model for the Grampian case study. The full source code of the Grampian prototype model is available on the CAVES CVS server, which provides access for all CAVES participants.
- Deliverable No. 5: Prototype of the initial model for the Odra River valley. The full source code of the Odra River prototype model is available on the CAVES CVS server, which provides access for all CAVES participants.

[24] 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.
[25] See e.g. http://en.wikipedia.org/wiki/Wiki

- Deliverable No. 6: Critical examination of resilience theories and measures. Since the dissemination level of this report is PP (restricted to other programme participants), it can be found in the restricted area of the CAVES internet portal.
- Deliverable No. 7: Prototype of the initial model for the Limpopo region. The full source code of the Limpopo prototype model is available on the CAVES CVS server, which provides access for all CAVES participants.

5.3 Future Deliverables

At the end of the next six month period (project month 24) another Deliverable is due, the prototypes of generalised statistical mechanics models (Deliverable No. 9), with the WUT team as the leading participant. The postponed Deliverable 8 will also be available by then.

5.4 Milestones

All of the six milestones for the first half of the CAVES project have been achieved. These are two milestones each for work packages 1 and 2 and one milestone each for work packages 3 and 4:

- Milestone 1.1 (month 1): Launch workshop at the beginning of the project.
- Milestone 1.2 (month 6, 12, 18): Consortium meetings every 6 months. These meetings review methodological developments and agree the forward work plan.
- Milestone 2.1 (month 6): Initiation of case study with stakeholders. By month 6 field work had successfully been started in all three case studies.
- Milestone 2.2 (month 12-18): Prototype models for each case study have been developed and are available on the CAVES CVS Server and .internet portal.
- Milestone 3.1 (month 12): Prototype model evaluation and interactive stakeholder workshop. This has been integrated into the consortium meetings at month 6 and 12.
- Milestone 4.1 (month 9): Review of analytical methods from first consortium workshop. These reviews form part of the knowledge base and are available via the CAVES internet portal.

6 Appendix

6.1 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 project meeting in Wroclaw at the end of September 2005 was therefore extended by one day.

Due to time constraints of several project partners the project meeting in Vienna will take place from 27th-30th March 2007 instead of the originally planned date in February 2007.

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

6.2 Plan for Using and Disseminating the Knowledge

6.2.1 Exploitable Knowledge and its Use

Due to the nature of our research, the CAVES project hasn't generated any exploitable results yet.

6.2.2 Disseminating the Knowledge

The CAVES project is dedicated to publish results in relevant scientific journals and conferences. During the first 18 months of the project, the following papers and posters have been written:

- S. Alam and R. Meyer, Finding Suitable Analysis Techniques for Agent-based

Networks Generated from Social Processes, in Proceedings of Social Network Analysis: Second Forum on Advances and Empirical Applications, Leeds, UK, June 2006

- G. Polhill and G. Ziervogel: Using ontologies with case studies: an end-user perspective on OWL. NCESS 2006, Second International Conference on e-Social Science, 28-30 June 2006, Manchester, UK
- S. Alam: Towards better socioeconomic policies in the Sub-Saharan region, Poster at the ESRC Research Methods Festival, Oxford, UK, July 2006
- S. Alam, R. Meyer and G. Ziervogel: Modelling the Socio-Economic Impact of HIV/AIDS in South Africa. WCSS 2006, First World Congress on Social Simulation, 21-25 August 2006, Kyoto, Japan
- A. Ernst, F. Krebs and C. Zehnpfund: Dynamics of task oriented agent behaviour in multiple layer social networks. WCSS 2006, First World Congress on Social Simulation, 21-25 August 2006, Kyoto, Japan
- N. Gotts and G. Polhill: Simulating Socio-Techno-Ecoystems. WCSS 2006, First World Congress on Social Simulation, 21-25 August 2006, Kyoto, Japan
- B. Werth, S. Moss, G. Ziervogel and T. Downing: Modelling Migration in the Sahel: An alternative to cost-benefit analysis. WCSS 2006, First World Congress on Social Simulation, 21-25 August 2006, Kyoto, Japan

  This paper won the student contest at WCSS'06 in Kyoto and was awarded the gold medal for best demonstration.
- E. Norling, B. Edmonds and S.J. Alam: Complexity in Socially-Inspired Simulation. Poster presented at the annual meeting of the EPSRC NANIA project, Ambleside, 29-31 August 2006
- L.A. Small: Modelling Social Networks: An Application for Rural Sociology? Paper presented at XII International Summer School, "The Empire of Rules and the Ethics of Social Networks", August/September 2006
- L.A. Small: Introducing CAVES: Complexity, Volatility, Evidence and Scale. Poster at the Participatory Approaches in Science and Technology conference, Edinburgh, June 2006; and the Rural Sociological Society Annual Meeting, Louisville, KY, USA, August 2006.
- S. Alam, A. Ernst, F. Krebs, R. Meyer, S. Moss and C. Zehnpfund: Evidence-Based Social Networks: Testing for Complexity. Submitted to MA4CS'06, a satellite workshop to ECCS'06, the European Conference on Complex Systems, Oxford, 25-29 September 2006.
- N. Gotts and G. Polhill: Narrative Scenarios, Mediating Formalisms, and the Agent-Based Simulation of Land Use Change. EPOS 2006: Epistemological Perspectives on Simulation, II Edition, University of Brescia, Italy, October 5-6, 2006
- K. Ostasiewicz, M. Tyc, P. Goliczewski, P. Magnuszewski, A. Radosz, and J. Sendzimir: Integrating economic and psychological insights in binary choice models with social interactions. To be submitted to Journal of Economic Behavior and Organization.

The CAVES project was also presented at the GIACS/STREPS Forum of Complexity Scientific Research Projects during the European Conference on Complex Systems 2006 (ECCS '06) in Oxford, 25-29 September 2006.

In addition, the CPM is involved in the Manchester Complexity Seminars[26] where members of the modelling team have so far contributed several talks about their work in progress for the CAVES project.

The CAVES project will continue to publish results in relevant scientific journals and conferences. The internet portal at http://caves.cpfm.org is another medium to disseminate the knowledge gained in the project. It is especially suited to share the developed software – in our case: simulation models – with a broader scientific audience. At the end of the project, the software will be offered to permanent open source sites such as sourceforge.org.

All software is and will be developed using open source languages and programming libraries and demonstrator and production software will be distributed under open source licenses. Under these licenses, commercial exploitation will be subject to royalties that will be distributed amongst the partners in proportion to their contribution as agreed amongst themselves at the end of the project.

The major single event for disseminating results will be an international workshop aimed at both academic experts and key officials involved in climate policy in general and land use issues in particular. We have set aside a substantial sum in the project budget (nearly €8000) for this purpose. The objective of the final workshop will be both to present our results and to engage the experts and officials in exercises with the procedures developed by the CAVES project.

6.2.3 Publishable Results

None yet, see 6.2.1.

[26] See: http://cfpm.org/nania/mancompl.html