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

CAVES Progress Report No. 3

March 2007 • Project 012816: CAVES –
Complexity, Agents, Volatility, Evidence and Scale

Duration: 2005-2008

Funded under the EU 6FP NEST programme.

Table of Contents

1 Introduction

Two shining beacons of success in the CAVES Project are the volume of evidence gathered for the three case studies and the harmonious and ongoing working relationships between the case study teams responsible for acquiring the evidence and the modelling teams responsible for using the evidence and specifying further evidential requirements. Whilst there is considerable interaction amongst the modelling teams, each has taken a slightly different approach to evidence-based modelling. The teams at the Centre for Policy Modelling and the Macaulay Institute are constraining their models directly and explicitly by detailed evidence while the team at CESR in Kassel have started from a higher level of abstraction (though still constrained by evidence to a degree that is unusual in complexity science). The CPM and CESR modellers have continued to develop models with a strong declarative element whilst the Macaulay modellers maintain their strictly procedural modelling approach. The combination of these differences and the close and collegial relations amongst the modelling teams is providing an important benefit for the complexity science community since we will be in a uniquely strong position to assess the difficulties and benefits of each approach and to form a view – or perhaps several views – concerning approaches to be taken in different circumstances and for different purposes.

There are similarly close working relationships underpinning the development and analysis of the evidence used to constrain model design and implementation. As Macaulay is unique in relation to the CAVES Project in providing both, evidence and modelling, the task of model validation is proceeding as smoothly as we anticipated. Relations also remain close between the separate modelling and case study teams working on the Odra and South African case studies. There have been visits between these pairs of teams and a field trip to South Africa was planned during the period covered by this report and has since taken place with very useful results.

The FEARLUS framework has been developed appropriately for procedural modelling of emerging and evolving social complexity and CPM and CESR have been developing their declarative modelling environments with useful exchanges of code and advice.

In all cases, the generation and analysis of social networks has been proceeding smoothly and effectively.

In summary, the project remains in very good heart and the work is progressing in line with our expectations at the start of the project. Delays at the start of the project while staff were hired and integrated into the partners' teams was fully anticipated during contract negotiations and incorporated into our plan to extend the project by two months subject, of course, to approval from the European Commission.

2 Modelling issues

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

2.1 Centre for Policy Modelling

Work at the CPM has concentrated on developing the declarative model for the South African case study. This included fathoming the limits of applying Jess for declarative modelling in a complex agent-based simulation model. Other foci of work have been further researching measures for social network analysis and commencing the validation process in collaboration with the case study team from SEI Oxford. In addition, two academic papers have been written and submitted to an international conference and a scientific journal (Alam et al. 2007, Alam et al. forthcoming).

2.1.1 Declarative Model

One focus of work has been the development of a declarative model for the South African case study. This model represents the agents' cognition and decision processes on a much finer scale than the procedural model.

Model description

The model distinguishes between several levels of social interaction: the individual, the household and the village level. The number of households and villages are model parameters. Households and their members are created according to the evidence from the RADAR IMAGE study data, concerning family structures and age distributions. These have been elicited through an extensive data analysis as described in the previous report.

After creating households the model assigns kinship links between them. Since empirical data about these is missing, we assume that kinship between households follows a small-world network. We then perform a cluster analysis on the links to decide which households belong to which village.

The model so far abstracts from real spatial data. Villages occupy a certain space on a grid of square cells. Households are assigned a random square inside their village. Judging from the Google Earth map of Sekhukhune villages are not densely "packed", i.e. we assume that not all squares inside a village are occupied by houses. The "density" of a village can be controlled via a model parameter.

Neighbourhood is established between households based on their position on the grid. All households that fall within a certain radius are regarded as neighbours. The radius is a model parameter.

80% of people are members of a church. All members of a household belong to the same church. There are churches of different denominations in the villages, one church per denomination. So far, the model applies a random distribution of 1-4 denominations per village and assigns households randomly.

In addition to the kinship network and neighbourhood links between households, people form friendship networks on an individual basis. We use the following assumptions here: friends are of the same gender, similar age (+/- 3 years for children, +/- 8 years for adults) and generally of similar characteristics. This means, that they have a number of personal attributes in common (modelled as abstract tags) and that they come from a similar background, e.g. attend the same church, are neighbours or know each other as being reliable and trustworthy (modelled as endorsements).

Stokvels are formed between groups of friends. Only household heads belong to a stokvel and they only belong to one stokvel at a time. If there is enough money to pay for the (food) expenses in a month, household heads express the desire to form a stokvel and ask the household heads amongst their friends. If enough of them (4-8) agree, they form a stokvel.

Household income depends on state grants for children up to 7 years old and pensions for seniors. As of now, every eligible child receives the grant while only about half of the eligible seniors receive the state pension. From the RADAR data we know that 58% of females aged 60 years and over and 50% of males aged 65 years and over receive pensions. These percentages are used in the model and assigned randomly to seniors.

Apart from the state grants the income of households relies on remittances and employment. A certain number of people are assigned migrant status at model initialisation. Number, gender and age distribution of migrants are taken from the RADAR data. Also, a certain proportion of adults (15-60 years old) are randomly assigned employment. 10% of those are assumed to have a steady government-paid job (e.g. teacher, hospital staff). We further assume that these – plus the ones with seniors receiving a state pension -- are the households that can afford to employ other people in the village, e.g. for piece jobs. At initialisation, people are randomly assigned to such jobs. Later on, "rich" households advertise job offers whenever they can afford it.

Friendship network issues

During verification of the model part creating the friendship network we discovered the following issue: A high proportion of links between friends is one-directional, meaning that e.g. person A regards person B as a friend but not vice versa. This raises the question of how friendship is defined. Is it always reciprocal?

We identified the model mechanism responsible for the current outcome. Agents (individual people) rely on endorsements to determine their friends. Endorsements can be thought of as labels used by an agent to describe certain aspects of other agents[1]. The model currently incorporates the following: is-kin, is-neighbour, same-church, same-denomination, is-acquaintance, is-friend, similar, most-similar, reliable, honest, capable, trustworthy, recommended (positive labels) and unreliable, dishonest, untrustworthy, incapable (negative labels). All agents use the same list of endorsements but differ in how they assess them.

To do so, agents rely on a so-called endorsement scheme which associates each label with a weight to express how much store an agent sets by this particular aspect of a person. Weights are modelled as integer numbers between 1 and *n* for positive labels and -1 and -*n* for negative labels, respectively. This allows for computing an overall endorsement value for a person, applying the following formula:

$$\sum_{i=1}^{numLabels} signum(weight_i) \cdot b^{|weight_i|}$$

Agents are assigned random endorsements schemes at creation, which differ not only in the weights they assign to the labels but also in the values used for *n* and *b*.

To form the friendship network, agents first determine which other agents are similar to themselves. This is based on abstract tags (to model character traits), which are assigned randomly to agents at creation. These tags are used to compute a similarity index (the number of similar tags), which in turn is used to generate "similar" and "most-similar" endorsements. Agents then compute the endorsement value for all known other agents and choose the ones with the highest values as their friends. Due to the different endorsement schemes agents with a high similarity index do not automatically end up as mutual friends. Since the weights are assigned randomly, agent A might rate similarity high, while agent B rates it low, thus making agent A pick B as a friend but not vice versa.

This is of course a technical artefact and we explored different ways to remedy it:

- Devising different endorsement schemes for different purposes. In this case, applying a special friendship endorsement scheme, which always ranks similarity (amongst the) highest.
- Increasing the maximal number of friends a person can have (a model parameter).

Interestingly, it turned out that the maximal number of friends is the key parameter here. Increasing this from the original 4-6 to a value of 6-12 resulted in a friendship network with a high

[1] Endorsements were first devised by Paul Cohen (1985) as a device for resolving conflicts in rule-based expert systems. Scott Moss (1995) modified and extended their use within a model of learning by social agents. This latter version of endorsements has been adapted for the declarative model.

proportion of reciprocal links.

2.1.2 Modelling Techniques

While still implemented in a Java/Repast frame, the declarative model is making extensive use of Jess, both in terms of the rule base and the working memory (number of facts). At quite an early stage in the development process we encountered a considerable slowing-down in the model execution even for small numbers of households. This made it unfeasible to run for the necessary larger numbers of households (> 100).

To find the performance bottleneck we applied the profiling tool JProbe[2], a trial version of which is available for free. This pointed us to Jess being the culprit. Jess uses the so-called Rete network, a special data structure to match facts to the left-hand side of rules, to allow for a fast execution of rules. To keep this data structure up-to-date it has to be re-computed whenever a fact is entered into or removed from the working memory. Since our model produced a large number of new facts at every single time step and comprised a fairly large number of rules, this re-computation took a long time.

After going through the rule base rule by rule to determine which of the rules affected the execution time the most, we decided on a two-fold approach to achieve a speed-up:

- *Reduce the number of rules*. Several rules did not deal explicitly with decision processes but were more or less procedural in nature; e.g. updating the households' cash or evolving tags. These were all ported to Java. Since some of them need to have access to the currently existing facts we implemented a Java class that browses through the fact base once per time step and delivers the necessary facts to the respective methods.

- *Reduce the number of facts*. Whenever suitable, we replaced explicit facts with corresponding fields in Java classes acting as shadow facts in Jess. For example, facts keeping record of household economy (`monthly-food-cost`, `expected-income`, `cash-in-hand`) were replaced by fields in the `Household` class (`monthlyFoodCost`, `expectedIncome`, `cash`) and updated from Java. Rules in Jess can access these fields as slots in the shadow fact.

  This works not only for single-valued fields but for lists of values, too. The `known-person` facts keeping track of which persons a particular individual is acquainted with could thus be replaced with a `knownPerson` field in the `Person` class, which acts as a multi-slot from Jess. Instead of looking for the existence of a `known-person` fact, the respective rules now test if the person in question is a member of the `knownPersons` multi-slot. This has proven to be much faster.

For the ongoing development of the declarative model we will have to find a balance between the expressiveness of declarative modelling in Jess and the faster execution of Java code.

2.1.3 Dynamic Network Analysis

Another focus of work has been to investigate measures for evolving social networks. Both the procedural and the declarative model take into account several types of networks such as those determined by the households' social neighbourhood, extended family ties and the agents' friends. There are also savings clubs that are, in fact, fragments of networks and less persistent. The resulting social networks are dynamic and changing over time as well as constrained by the underlying social processes. Traditional social network analysis measures, e.g. degree centrality, do not work for a variety of agent-based models (Edmonds/Chattoe 2005). Analyzing the topology of dynamic overlapping social networks is therefore a non-trivial task.

[2] http://www.quest.com/jprobe/

One possible approach is the use of the nonparametric statistic Kolmogorov-Smirnov (K-S) test (Neave/McConwa 1987) to compare the snapshots of a simulation run at regular intervals. This test gives a probability (p-score) that two sets of data (here, node frequencies) might come from the same distribution; it does not depend upon assumptions as to the type of distribution. A low p-score is thus a good indication that the two sets of data have different distributions.

In a first study, we evaluated simulation runs with the procedural model, considering only the households' neighbourhood links which are mapped into a social space. We wanted to know if the choice of the initial characteristic network affects the subsequent course of the simulation. Alternatively, if the network structure is governed by the local social processes, then a choice of initial configuration should be 'significant'. The K-S test is applied to compare the network composition at different time steps. A couple of linking configurations have been considered. Since no data was available about how the households are socially linked we referred to standard network structures. Simulations were set up for two initial network configurations: the Watts-Strogatz (WS) and Random Network with 10% density (RN).

Firstly, we compared snapshots (from the 200^{th} tick onwards) at every 50th tick with the initial households' social network configuration (the 0^{th} tick). Simulations were run for both WS and RN setups. The p-value calculated for each snapshot compared to the 0^{th} tick is plotted in Figure 1. In both cases, the statistic indicated that the structure has changed from that at the start at around 500-600 ticks.

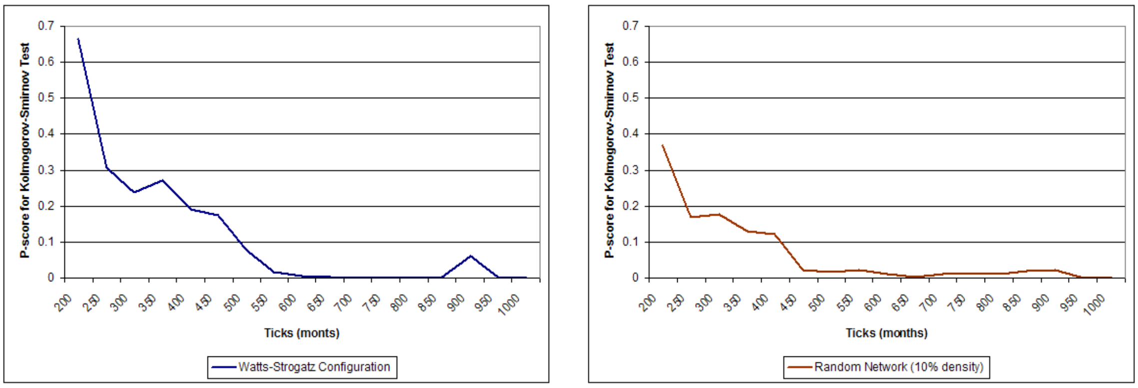


Figure 1: P-score of the K-S test; the degree distribution of the social network taken at every 50th tick, compared with that at tick 0 for WS (left) and RN (right).

In the next step, a set of snapshots were chosen $P= \{P_0, P_{0+\Delta}, P_{0+2\Delta} \ldots\}$ where each member represents a distinct population of agents (households), P_0 being the population at t=0. Here we perform the K-S test for each consecutive pair $(P_i, P_{i+\Delta})$. For Δ =50, the distribution did not change successively (figure 2) until near to the end of the simulation runs. This indicates that the network did change significantly then within 50 ticks. It happened when there were no longer adults

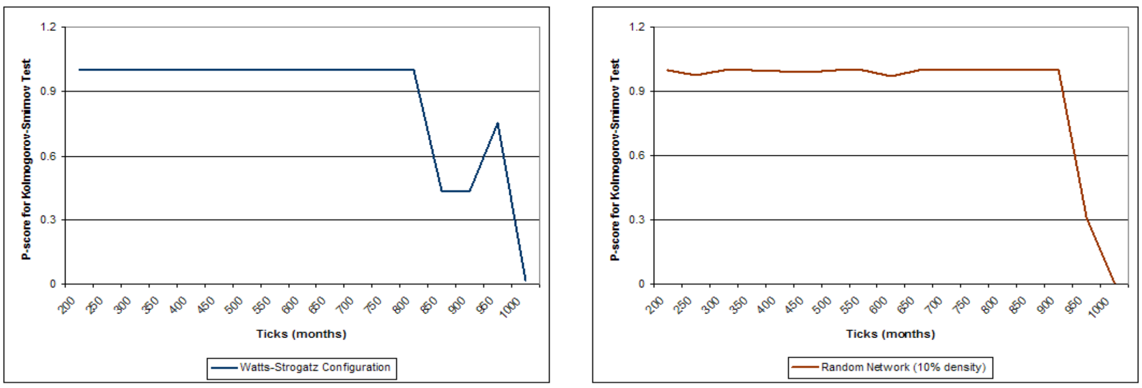


Figure 2: P-score of K-S test; network snapshot at every 50th tick, compared with the previously taken snapshot for WS (left) and RN (right) initial settings

Assuming the size of delta to be small, we performed the test on the degree distribution of the households' social network again on the two different configurations. The choice of lag size was 200 this time, and as figure 3 shows, a different pattern is seen. Unsurprisingly, one cannot expect to see a radical change in the distribution at a much finer scale. However, changes do become evident for larger time periods.

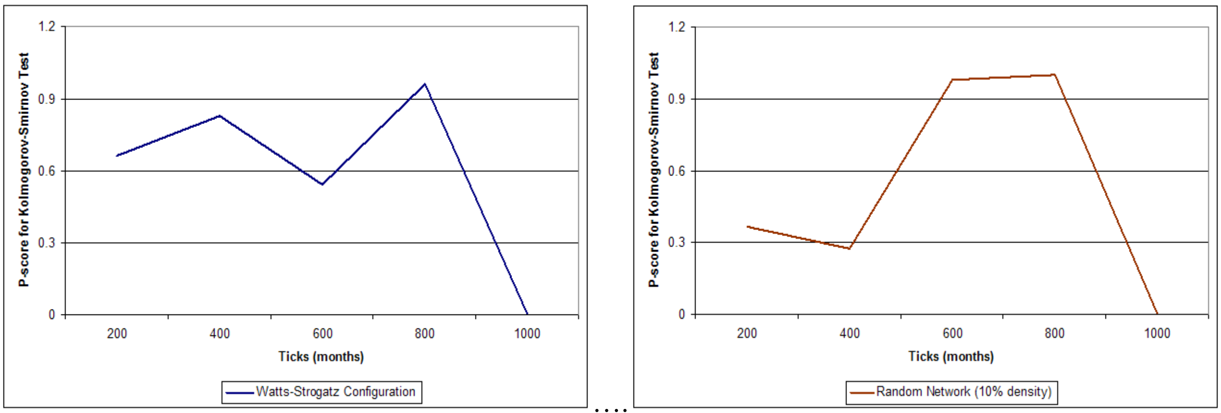


Figure 3: P-score of K-S test; network snapshot at every 200th tick, compared with the previously taken snapshot for WS (left) and RN (right) initial settings.

In comparing two snapshots as two populations, the choice of the lags seems to be significant. Obviously, the choice of an 'appropriate' Δ varies with respect to the social phenomenon, which an agent-based model attempts to capture. The analysis is in earlier stages, but may be useful for identifying statistical signatures (Moss/Edmonds 2005) for agent-based social networks, where a rigorous treatment is currently lacking. The method applied here could well be one possible solution. A paper about these findings was accepted for the international conference AAMAS (Alam et al. 2007). We are also planning to submit a paper to the Fourth Conference of the European Social Simulation Association (ESSA 2007).

2.1.4 Validation

Together with our case study partners at SEI Oxford we have also started the validation process for both models (see section 3.1). The first step has been to compile a list of all assumptions that have been applied in building the models. This list was then discussed by the domain experts and returned to the modellers with comments and suggestions of what to change.

The next iteration will be to realise the refined assumptions in the models and discuss the results with the domain experts again. We expect that this iterative process of model development will support identifying gaps in the available empirical data and thus guide the knowledge elicitation process in general.

2.1.5 Internet Portal

In the course of integrating the Wiki component for the knowledge base and glossary it was decided to move the entire internet portal to the Wiki. This has the advantage that all partners can not only contribute to glossary and knowledge base but work collectively on all parts of the portal, making it an even better internal communication tool.

To achieve this, a new Wiki-Layout was designed to match the established layout of the old version. Moreover, the portal was extended to incorporate links to project-related documents and an area to discuss work in progress. The internet portal can be found at `http://caves.cfpm.org`.

2.1.6 Future Work

- Extend the declarative model to incorporate further social processes like burial societies, marriages and the spread of HIV/AIDS. Include a genetic algorithm with cross-over and mutation for tags so that a part of them is inherited from the parents.
- Continue work on social networks and statistical signatures.
- 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.

2.1.7 References

Alam, S; Meyer, R.; Norling, E. (2007): Agent-based Model of Impact of Socioeconomic Stressors: A Dynamic Network Perspective. Accepted for AAMAS 2007 (International Conference on Autonomous Agents and Multi-Agent Systems), Hawaii, May 14-18, 2007

Alam, S.; Meyer, R.; Moss, S.; Ziervogel, G. (forthcoming): Impact of HIV/AIDS in the Context of Socioeconomic Stressors: An Evidence-driven Approach. Submitted to JASSS (Journal of Artificial Societies and Social Simulation), currently under review.

Cohen, P. (1985): Heuristic Reasoning – An Artificial Intelligence Approach. Boston, Pitman Advanced Publishing Program.

Edmonds, B.; Chattoe, E. (2005): When Simple Measures Fail – Characterising Social Networks Using Simulation. Presented at the Social Network Analysis: Advances and Empirical Applications Forum, Oxford, July 16-17 2005. Available online as CPM Report No. CPM-05-158 at http://cfpm.org/cpmrep158.html

Moss, S. (1995): Control metaphors in the modelling of decision-making behaviour. Computational Economics, v. 8, pp 283-301

Moss, S.; Edmonds, B. (2005): Sociology and Simulation – Statistical and Qualitative Cross-Validation. In: American Journal of Sociology, 110(4), pages 1095-1131.

Neave, H.; McConwa, K. (1987): Distribution free methods. Open University Press

2.2 Universität Kassel

2.2.1 Organisational issues

Michael Elbers was recruited and started work on December 1, 2006.

2.2.2 Additional case study meeting

At the end of November 2006 (23-24), Dr. Karolina Krolikowska visited the Kassel group in order to discuss the results of the interviews she conducted with stakeholders in the case study region. Abstracting the most pervasive aspects from these interviews a set of rules was compiled to reflect the decision making of farmers in the Odra region. A number of farmer types (unwilling/willing part-time farmers, big farmers without fish pond, fish pond owners and water partnership initiators) were identified along several dimensions including land ownership, activity related to LRS maintenance and agricultural knowledge. For each type of farmer the extracted rules relate to farming, LRS maintenance and social network activity.

2.2.3 Modelling

The SoNARe (Social Networks of Agents' Reclamation of land) model was further extended and re-fined. First and foremost, the decision making of the agents was enhanced on the basis of the rule set described above and the biophysical model developed by the Wroclaw University of Technology group (WUT) was coupled with the agent-based model.

The agents in 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 production rules implemented in Jess and Jess's reasoning engine to simulate the cognitive control structure and decision making of farmers and other relevant actors. Moreover, large parts of the code are written in Java and using the Repast library functions 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*. Simulated flooding or drought lead to a loss of crop yield.
- *Maintenance of the land reclamation system (LRS)*. Following assumptions already made in earlier models, 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*. The difficulties concerning land and water use in the Odra case study region are mainly due to the fact that the conditions encountered on individual land parcels depend highly on the amount of effort devoted to the maintenance of the LRS on every land parcel downstream (especially in cases of flooding) and on the sluice gate operation of upstream neighbours in cases of drought. This asymmetrical dependency entails a social dilemma structure, in turn giving incentives to free riding, for example (i.e. not providing LRS maintenance on one's own parcel when being the first after a row of cooperators). 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 actors.

Environment (the biophysical model)

The biophysical model simulates the effects of different weather conditions, LRS maintenance and LRS neglect as well as sluice gate operations on the water levels and thus on the crop yield of individual land parcels along a channel . It offers a number of parameters to be varied across simulation runs, most importantly:

- *Weather sequence*. This is a repeated sequence of normal, wet and/or dry years, which differ in the mean water levels in the channel per month.
- *Land parcels*. The number of land parcels per channel and the type of land use of each land parcel can be set. Currently, the biophysical model allows for two types of land use, namely arable and fish-pond, of which only the former is used in the simulation runs described in the next section. Moreover, the initial condition of the LRS section on each individual parcel can be specified.

The model is run at monthly steps with crops for arable land parcels being planted in month 5 and harvested in month 10. The condition of the local LRS slowly degrades when it is not maintained and (at present) it fully recovers within a month of an agent first maintaining it. By instantiating the biophysical model multiple times it is possible to simulate any number of channels in parallel without any interrelations among them.

Agents

Currently, two types of agents are modelled: the farmer and the water partnership initiator (WPI). Farmers are embedded in a dependency network according to the location of their land parcel along their respective channel and in an overall acquaintance network which is randomly superimposed on the dependency networks thus spanning all channels. The acquaintance network not only includes all farmer agents, but also a WPI, which is linked to all farmers in a star-like manner. Presently, the WPI is not a farmer itself.

In order to capture the main characteristics of the evidence obtained from the stakeholder interviews and to make a feasible abstraction of the compiled decision rule set, the agent-based model was enhanced along several lines. The focus of these enhancements lies on explicitly contrasting social and economic influences on the decision making of farmers with regard to LRS maintenance. To this end two agent internal dimensions of perception were introduced which drive a farmer agent's decision making: economic success and social support.

- The perception of *economic success* is shaped by a yield perception threshold used for appraising yields, the yield success memory capacity, i.e. the maximum number of years for which experienced "good" or "bad" yields are memorised, and an economic sensitivity which determines the extent to which agents attach importance to economic factors
- The perception of *social support* is a function of the social influence level of a farmer's acquaintances (farmers and WPI) and the agreement/disagreement between farmer and acquaintance concerning LRS maintenance

On the basis of these two factors farmer agents decide on their LRS maintenance strategy changing it if their perceived economic success and their perceived social support fall below a certain level. Thus, in addition to the impact of the actual yields the model also reflects general opinion dynamics amongst farmers (cf. Latané, B. (1981), Friedkin, N. (1998)). This exertion of social influence is strictly symmetrical in the sense that each farmer supports a farmer who uses the same strategy as he himself by the same amount as he imposes pressure on a farmer using the opposite strategy.

The opinion dynamics were further extended to account for the social influence of WPIs on the pervasiveness of the LRS maintenance strategy and thus on the formation of a working water partnership (WP). A WPI agent exerts its social influence in favour of LRS maintenance once it perceives at least three farmers who have big losses; it does not exert any influence otherwise. The WPI's level of social influence is independent from that of the farmers. The institution of the water partnership is active as long as at least three farmers maintain their respective LRS; it is inactive otherwise. Farmers are automatically members of the WP as long as they maintain their LRS. However, at present, WP membership does not have any effect and there are no costs modelled for LRS maintenance.

Model execution cycle, sequence of events

For each year the model executes the following sequence of events: In month five the biophysical model simulates the planting of crops for each land parcel and in month ten it simulates the harvesting of these crops. Finally, at the end of every year the agent-based model is run performing the following sequence of steps: all agents perceive and memorise their individual yields, exert their social influence, perceive the social influence exerted on them, re-rate their economic success and then make their decisions for the next year. It is important to note that agents are synchronised at every step. Above all, this means that agents take their decisions simultaneously.

2.2.4 Simulation results

This section provides an overview of the first simulation results that were produced with the described model.

All three scenarios start off with no LRS maintainers and the LRS fully intact. The same weather sequence was used throughout: Two years with normal weather conditions are followed by one year with wet weather. This pattern is then repeated for the whole run. One important and general observation regarding the effect of weather on the crop yield as modelled by the biophysical model is that years are independent from each other in the sense that a wet year, for example, always has the same effect regardless of when it occurs.

Scenario 1 – Baseline scenario

The baseline scenario simulates ten land parcels along a single channel. Simulation runs exclude both effects of social influence between agents and agents' rating of their economic success. Thus, agents do not change their LRS maintenance strategy (they never maintain). Given these settings the model not only shows that the mean yield loss under wet weather conditions is significantly higher than under normal weather conditions, but also that the mean deviation is much greater. In cases of flooding and all farmers neglecting their LRS yields of farmers at the top of the canal are considerably worse than the yields of those further at the bottom. Moreover, the farmers at the bottom experience only minor differences between normal years and wet years in this scenario.

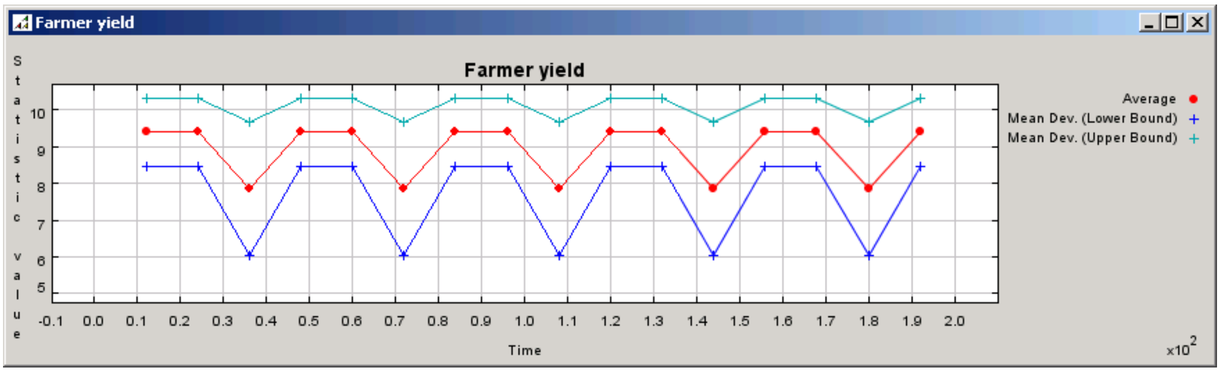


Figure 1: Average yields of farmers bounded by the mean deviation. Each dot marks the average yield for one year, i.e. sixteen years in total are shown.

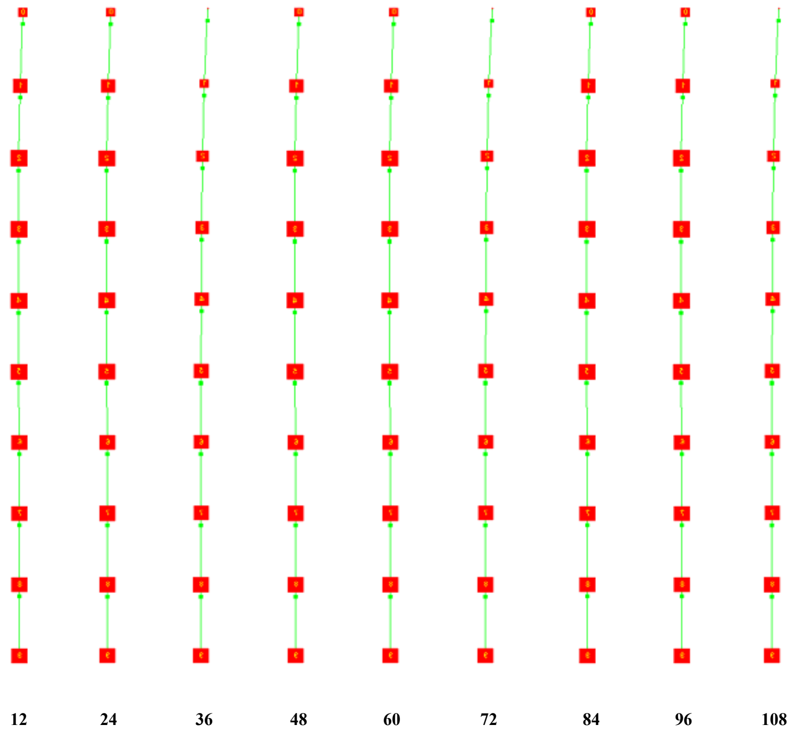


Figure 2: Yields and LRS strategies on individual land parcels along a single channel over time (years 1 to 9, i.e. months 12 to 108). Land parcels are represented as squares connected by green lines indicating the dependency relation (flow direction from top to bottom). The size of each square is proportional to the yield for the respective land parcel. The red colour indicates LRS neglect.

Scenario 2 – The influence of economic success

The second scenario differs from the baseline scenario only in that the farmer agents now appraise their economic success in the way described in the model setup. Again, social influencing is excluded and one channel with ten farmers is used.

In the simulation, the topmost farmer(s) first start maintaining LRS after a wet year. This is probably due to the combination of the positional disadvantage even in normal years (see scenario A) and bad yield in a wet year. As may be seen in figure 6, the perceived economic success of maintainers increases substantially and then settles on a slightly higher level than the corresponding values for non-maintainers. It can be observed that the number of maintainers converges to a stable state at 80% (8 farmers).

Figure 3: Strategy adjustments over time as an indicator for volatility. Each dot marks the proportion of farmers who have switched their LRS strategy in one year (30 years in total).

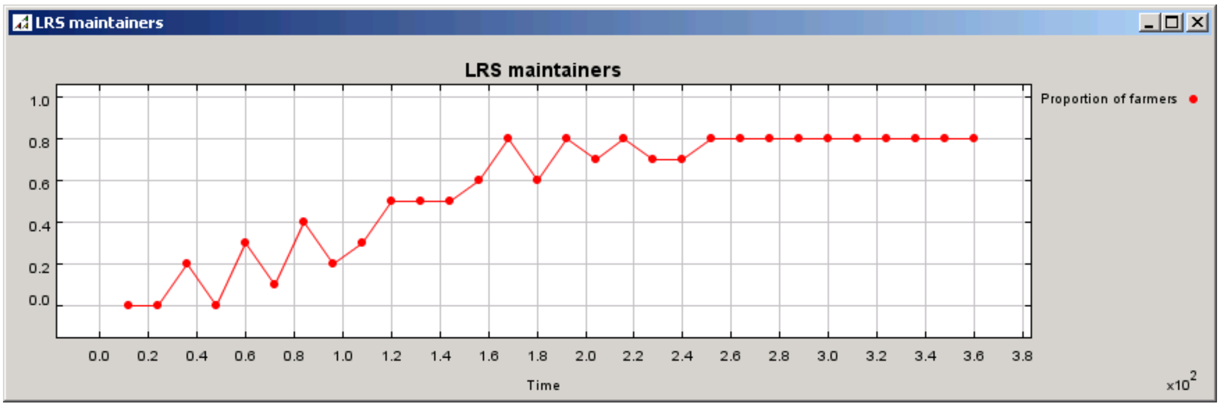


Figure 4: Proportion of LRS maintaining farmers over time.

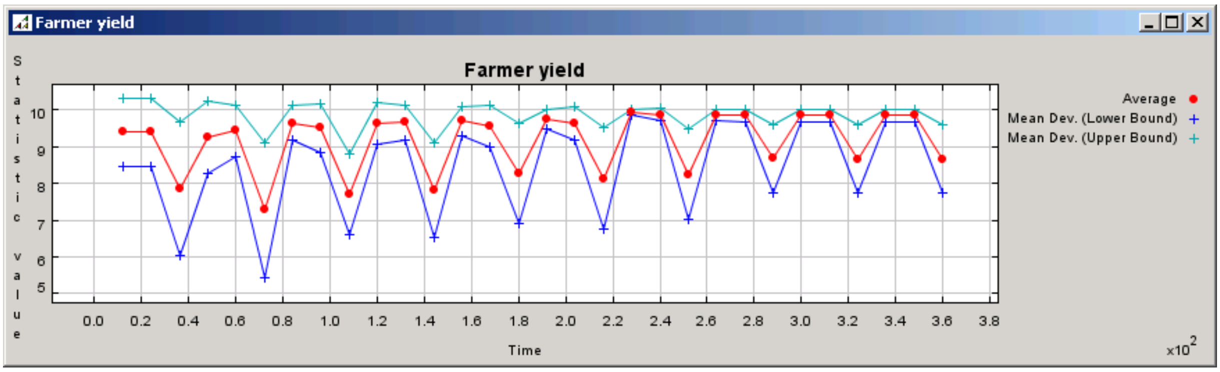


Figure 5: Average yields of farmers bounded by the mean deviation. Each dot marks the average yield for one year (30 years in total).

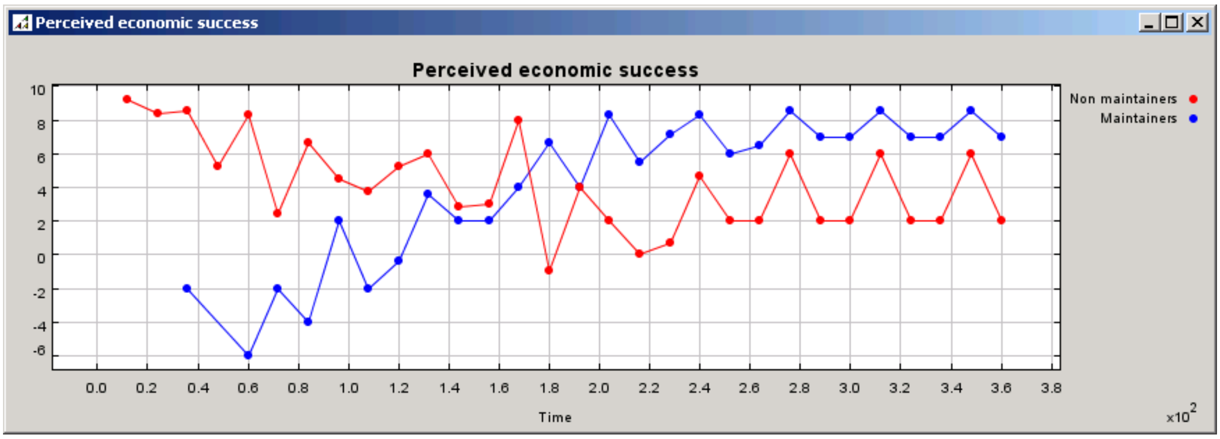


Figure 6: Mean perceived economic success of maintainers and non-maintainers over time with the yield threshold set to 9.0, an economic sensitivity of 2.0 and a yield memory capacity of 5 years, i.e. the perceived economic success has a lower bound of -10 and an upper bound of 10.

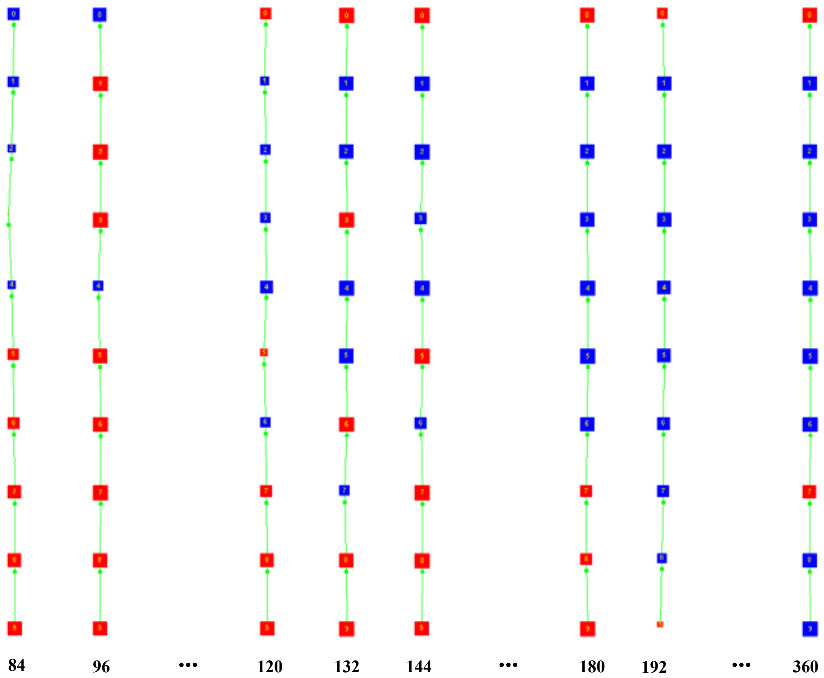


Figure 7: Yields and LRS strategies on individual land parcels along a single channel over time (years 7-30; numbers refer to months). Land parcels are represented as squares connected by green lines indicating the dependency relation (flow direction from top to bottom). The size of each square is proportional to the yield for the respective land parcel. Red squares indicate LRS neglect, blue squares LRS maintenance.

Scenario 3 – The combined influence of economic success and social support

In the third scenario agents use both their past economic success and social influence of other agents as a basis of their decision making. Farmers keep a limited memory of their past economic success. Five years are memorised, yields above 9.0 are recalled as "good" yields / years, and other yields are recalled as "bad". Good years memorised increase the perceived economic success, bad years decrease it. To investigate a reasonable number of farmers, in the simulation ten independent channels with each ten land parcels are considered. This amounts to 100 farmers in the simulation.

All farmers and one WPI are embedded in an acquaintances network that spans all ten channels. They exert their social influence as described in the previous section. In the simulation a quasi scale-free network topology with an average node degree of 10 is used. The assumption of a scale-free topology is supported by Odra case study narrative storylines. At least scale-free networks are more realistic than Watts-Strogatz Small-World topologies that were used in earlier experiments (see e.g. Barabási (2002)). The scale-free network used in the simulations is generated by an algorithm described by Ebel et al (2002). This algorithm allows generating a sufficient proxy of a scale-free network for 100 nodes and an average node degree of 10.

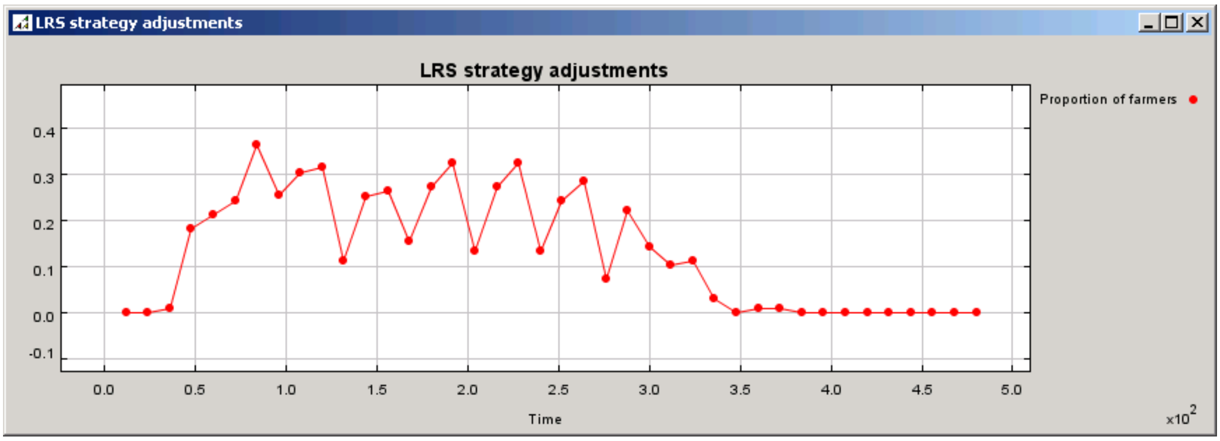


Figure 8: Strategy adjustments over time as an indicator for volatility. Each dot marks the proportion of farmers who have switched their LRS strategy in one year (40 years in total).

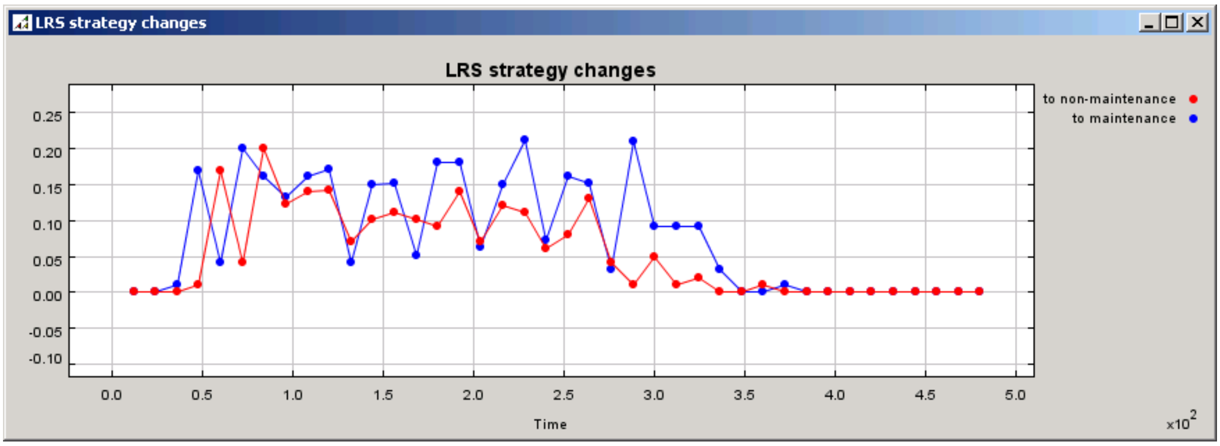


Figure 9: Proportion of strategy changes to maintenance and to non-maintenance strategy over time.

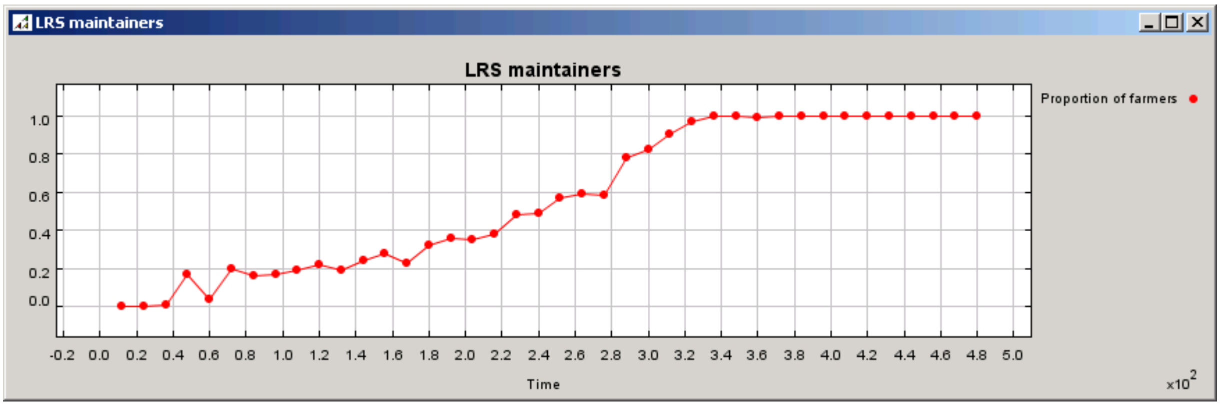


Figure 10: Proportion of LRS maintaining farmers over time.

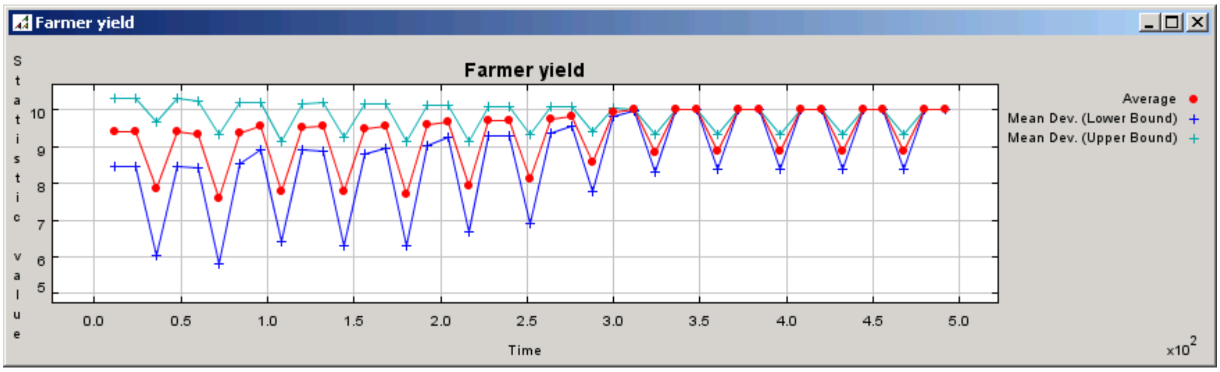


Figure 11: Average yields of farmers bounded by the mean deviation. Each dot marks the average yield for one year (40 years in total).

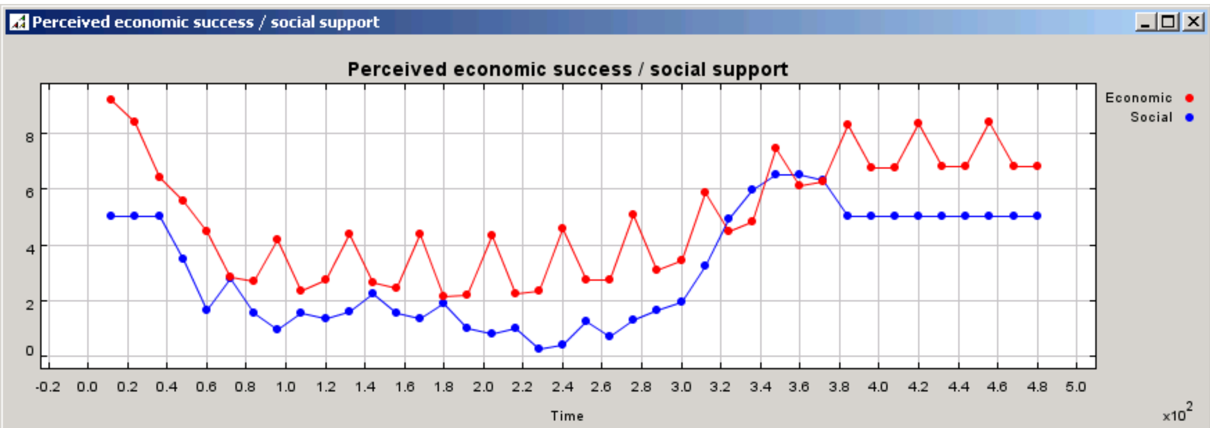


Figure 12: Mean perceived economic success and mean perceived social support over time.

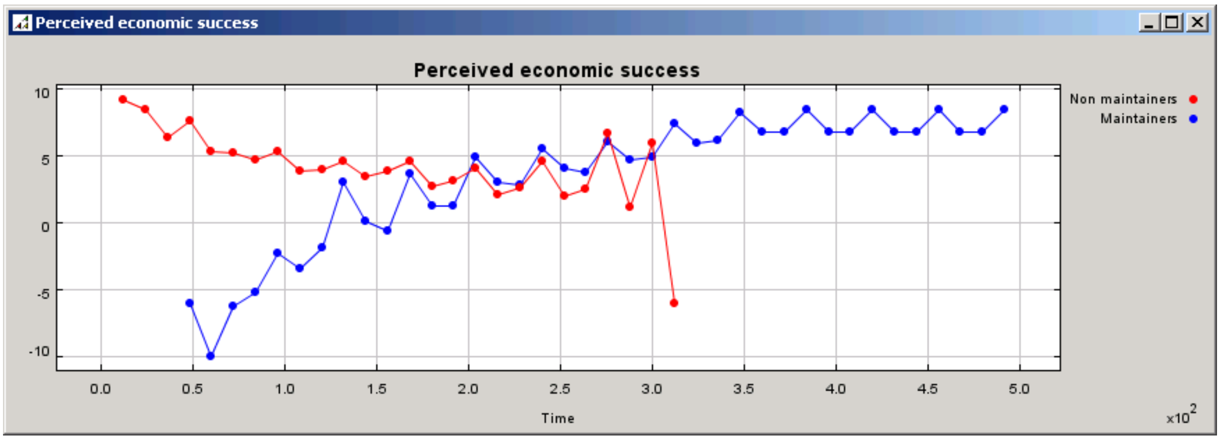


Figure 13: Mean perceived economic success of maintainers and non-maintainers over time with an economic sensitivity of 2.0 and a yield memory capacity of 5 years, i.e. the perceived economic success has a lower bound of -10 and an upper bound of 10.

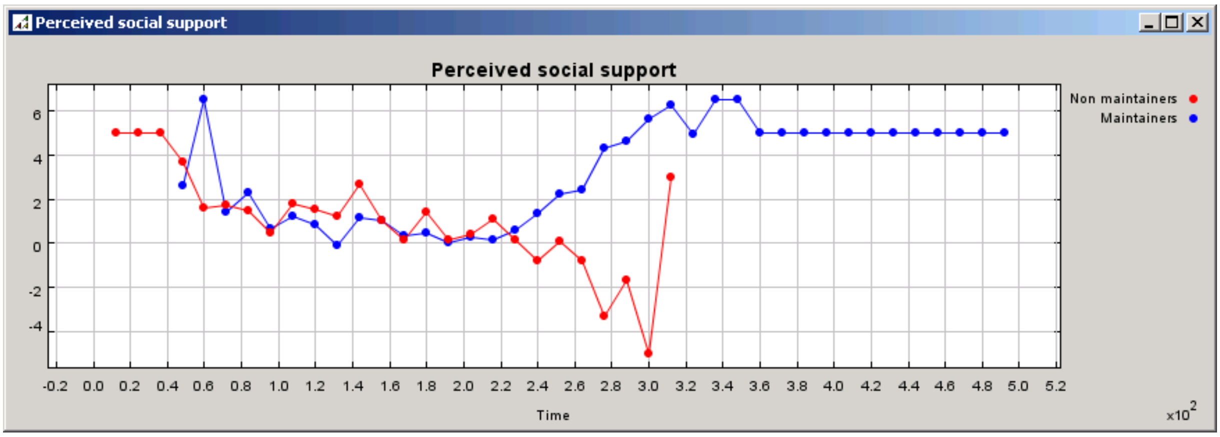
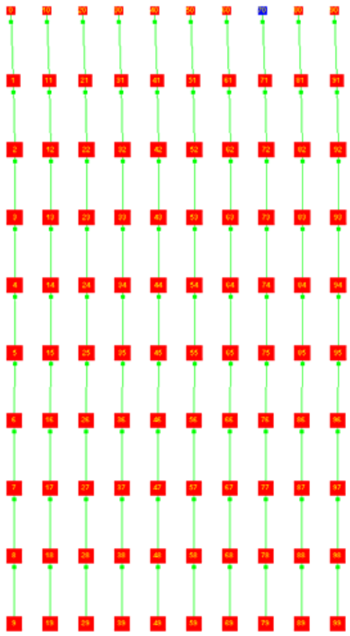
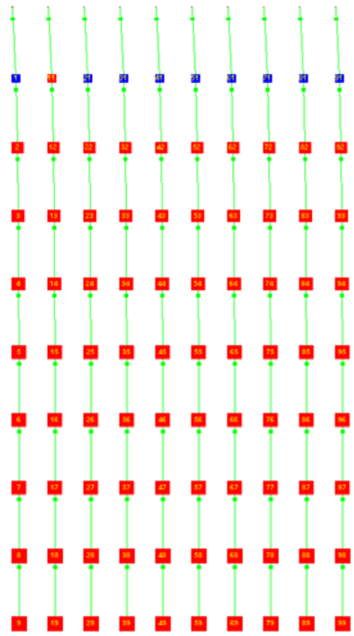


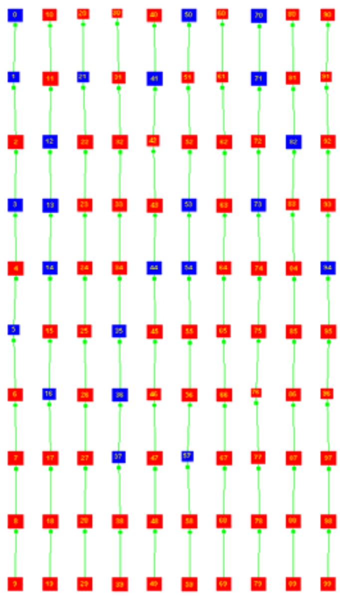
Figure 14: Mean perceived social support of maintainers and non-maintainers over time with a social influence level of 0.5 for farmers and 1.5 for the WPI and an average acquaintance degree of 10.

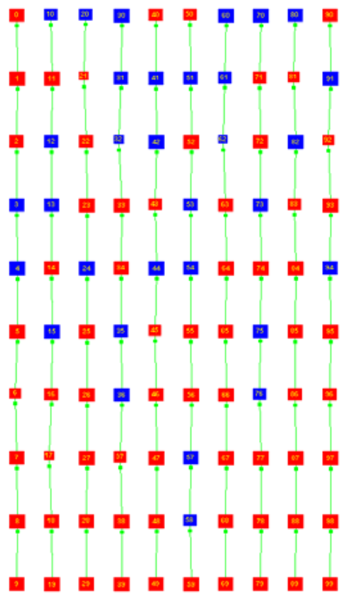




36

48





168

180

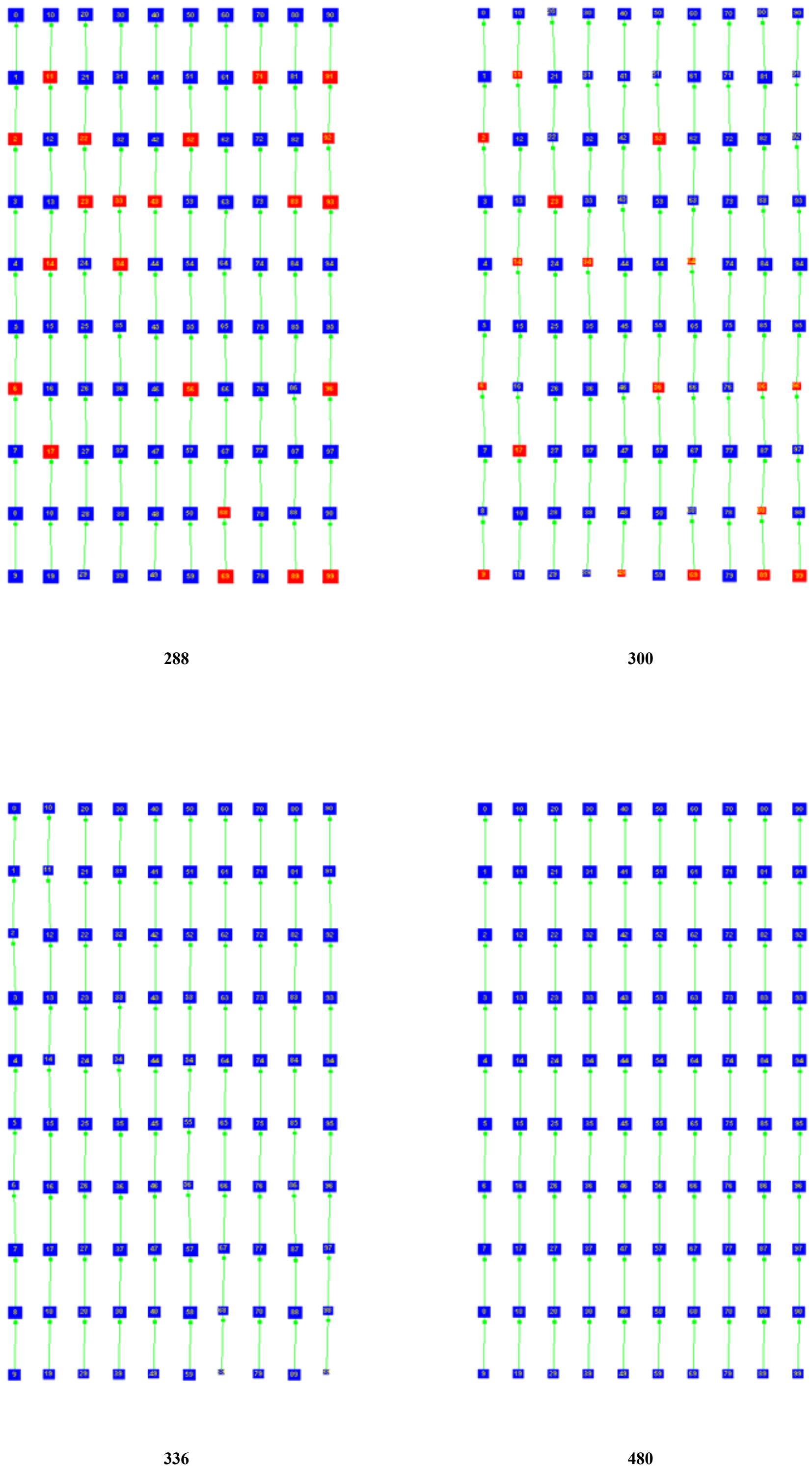


Figure 15: Yields and LRS strategies on individual land parcels along ten channels (left to right) over time (numbers refer to months, e.g. months 36 and 48, see above). Land parcels are represented as squares connected by green lines indicating the dependency relation (flow direction from top to bottom). The size of each square is proportional to the yield for the respective land parcel. Red squares indicate LRS neglect, blue squares LRS maintenance.

2.2.5 Discussion

In the work of the CAVES group situated at CESR in Kassel, it has been a guiding assumption that it is fruitful to model on a medium level of abstraction, i.e. to keep just between too much detail and too much abstraction. On one hand, there has to be a certain amount of recognisable empirical characteristics of the domain being modelled. On the other it has to make sure the model is interpretable in terms of complexity indicators, network characteristics, and basic theoretic or phenomenological structures, like social dilemmas.

The SoNARe model has made a step in this direction. It comprises a social dilemma structure derived from a geographical structure, a complex geo-physical model, stylised behavioural rules represented in actor types and well-founded psychological assumptions about social influence, memory span and social networks. While the model is being tested with only a smaller number of actors, it is easily scalable to several hundreds of actors without losing the basic environmental or social structure.

CESR has chosen a hybrid modelling approach using rules as a means to represent the core cognitive aspects of decision making while using a numerical way of blending different categories of reasons underlying the decisions, e.g. economic and social influences. There is clear evidence from neuropsychology for such an unconscious continuously changing pattern of neuronal activity, shifting towards a local minimum of activation – the decision (Spivey & Dale, 2006).

The SoNARe model produces, besides other behavioural indices, a measure of volatility: the amount of strategy changes by the actors. This measure will be investigated further, especially, how phases of volatility come about, e.g. through environmental (weather) or social conditions.

2.2.6 Outlook

First of all, the current model is to be calibrated and extensively tested by performing sensitivity analyses. Following this, it is intended to further approximate the situation in the case study region and press ahead with the investigation of the covariance of network properties and collective behaviour as well as possible phases of volatility. To these ends, the work group in Kassel plans to enhance and study the model with respect to:

- introducing allowances and compensation payments as economic factors,
- including additional land use types (fish ponds, in particular),
- modelling different farmer types in terms of economic and social perception and different sets of decision rules,
- distributing farmer types heterogeneously, and
- modelling sluice gate operation strategies of farmers on individual land parcels.

The group also intends to submit papers to the ESSA conference and the JASSS journal.

2.2.7 Literature

Barabási, A.-L. (2002). Linked: The new science of networks. Perseus Publishing, Cambridge.

Ebel, H., J. Davidsen, and S. Bornholdt (2002). Dynamics of social net-works. Complexity 8(2):24–27.

Friedkin, N. (1998). A Structural Theory of Social Influence. Cambridge University Press, Cambridge.

Latané, B. (1981). The Psychology of Social Impact. American Psychologist 36: 343–56.

Spivey, M.J. & Dale, R. (2006). Continuous dynamics in real-time cognition. Current Directions in

Psychological Science, 15, 5: 207–211.

2.3 Macaulay Institute

2.3.1 Introduction

Our intention is to construct models of the development of rural land use in Upper Deeside and surrounding areas of the Grampian region, between the mid 1980s and 2030. Models of the period up to the present will be aimed at determining how far our modelling approach can replicate the course of land use change (and farm size change, so far as we are able to obtain data on this), given the major external shocks and stresses impinging on the land use system over that period. Models of the period from the present to 2030 will be linked to scenarios related to climate change; and if there is time, to scenarios of new epidemics with an impact comparable to BSE.

The work of the CAVES Macaulay Modelling Team over the last six months can be divided into the following areas:

- Development, documentation and use of FEARLUS 0-8-2-1
- Development and documentation of FEARLUS 1-0
- Planning for scenario development and associated model construction and use
- Data collection
- Design and coding of AMEBON
- Discussion of various ways that ontologies could be used to link evidence to models

Details on each of these follow. In addition, a paper (Gotts, in press) drawing on Deliverable 6 (critical examination of resilience theories and measures), although not directly related to the central issues of CAVES, has been accepted by *Ecology and Society*.

2.3.2 FEARLUS 0-8-2-1

It was decided for reasons independent of CAVES to construct a general model (i.e., one not linked to any specific real-world case) concerning interactions of social networks and government policy in the control of agricultural pollution. These models were produced within FEARLUS 0-8-2-1, a revision of an existing version, in parallel with the design and development of FEARLUS 1-0. Only once these models were compared with models being developed for the Odra case study was it realised that they could serve as prototypes for CAVES models of future Grampian scenarios. FEARLUS 0-8-2-1, its user guide, and a range of parameter files will therefore be placed on the CAVES CVS server. A short paper describing the models is being prepared for ESSA 2007.

FEARLUS 0-8-2-1 has some of the features of FEARLUS 1-0 developed for use in CAVES, but not all. Most significantly here, its Land Manager agents, like those of FEARLUS 1-0, use case-based reasoning, and in general base decisions on a time-variable weighting of expected profit on the one hand, and expected approval or disapproval by neighbours on the other. However FEARLUS 0-8-2-1 does not incorporate the ELMM land sales model, using the simpler and less realistic land transfer algorithm from earlier versions of FEARLUS; and like those versions, uses bitstrings rather than the more flexible lookup tables in calculating land use outputs.

Controlling diffuse pollution of any kind raises monitoring and enforcement problems; we have begun to investigate the feasibility of mutual monitoring among groups of farmers, offered a collective reward if measured levels of the pollution they produce between them are satisfactory. For phosphate or other water pollution, this would need to involve the farmers in a river catchment or sub-catchment, pollution levels being measured just downstream of the furthest downstream farms. However, the model is sufficiently general to be applied to airborne pollutants or greenhouse gases, if levels of these can be measured only, or at lower cost, over an area containing multiple farms.

The model developed includes a Government agent, which in initial versions simply hands out a reward to all Land Manager agents if total pollution levels are below a threshold. The preliminary results obtained so far suggest that the presence of such a reward scheme, and social disapproval of pollution, each have some effect in lowering pollution levels, and that a combination of the two is more effective than either alone.

2.3.3 FEARLUS 1-0

FEARLUS 1-0 was released early in the last six month period. As documented in the last report, it implements various enhancements derived from the Grampian case study findings:

- Methods to implement economies of scale via a farm-scale fixed costs parameter, which can be made to vary over time, and parameters that reduce overheads in proportion to area of land assigned to a particular land use.
- Provision for land managers to have off-farm income sources.
- Augmentation of Case Based Reasoning decision making agents with the possibility to use imitative and other experimentation strategies when aspiration thresholds are not met.

FEARLUS 1-0 also contains a prototype feature to automatically output three kinds of ontology. The first kind is an ontology of the capability of the system as a whole, which contains all classes and properties implemented in the modelling framework. The second simply shows those classes and properties that are enabled by a particular application of FEARLUS as per the parameter files. An example class hierarchy is shown in Figure 1. The third provides data on individuals within the simulation at a particular time. An example showing the network of approval and disapproval in a particular year is shown in Figure 2.

There are a number of issues with the ontology generation prototype feature that mean it needs further development. First, it shows classes such as LTGroupState, which are not ontologically significant. Ideally, only ontologically significant classes would appear in the ontology.

Second, the object-oriented inheritance hierarchy does not have the same semantics as the OWL subClassOf relationship, with the result that in some cases, the subClassOf relationship can point in the wrong direction. In the OO case, the subclass relation, whilst not entirely unrelated to the underlying reality, will nevertheless be based at least in part on coding convenience (for example, minimising repetition of segments of code, or (related) maximising reuse), may be stipulated by particular approaches to system design (e.g. design patterns), and is further complicated (in those OO languages providing such mechanisms) by the availability of such things as interfaces (in Java), protocols (Obj-C) or templates (Java 5 and C++). These examples thus serve to illustrate that ontologies add value to model descriptions, and highlight the potential pitfalls in relying on UML diagrams (which would naturally tend to reflect the OO code rather than the ontology) for this purpose. It may be noted that in any case, authors such as Guarino and Welty (2004) suggest the ontological subclass relation tends to be over-used.

Finally, similar issues occur with the instance variables, and their translation to OWL properties. Here there is the added complication that the principle of encapsulation in OO classes is intended to hide the implementation, including non-public instance variables, whilst OWL properties enjoy a rather more independent existence from OWL concepts.

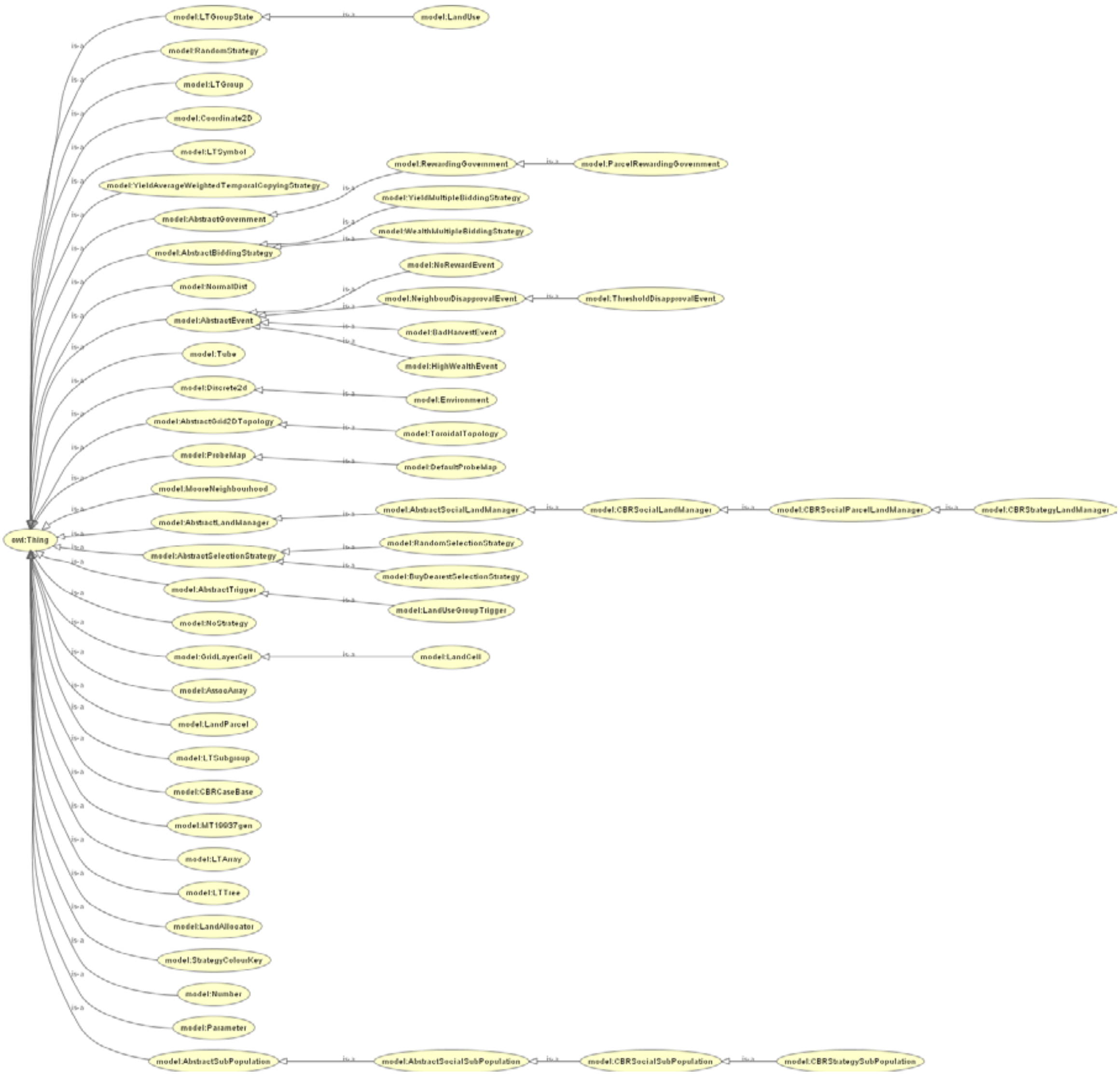


Figure 1. Example concept hierarchy from a model ontology.

It may be possible to deal with some of these issues by adopting coding conventions that highlight the ontologically significant features of the source code in a way that facilitates automatic ontology generation. This work will be continued outside the CAVES project.

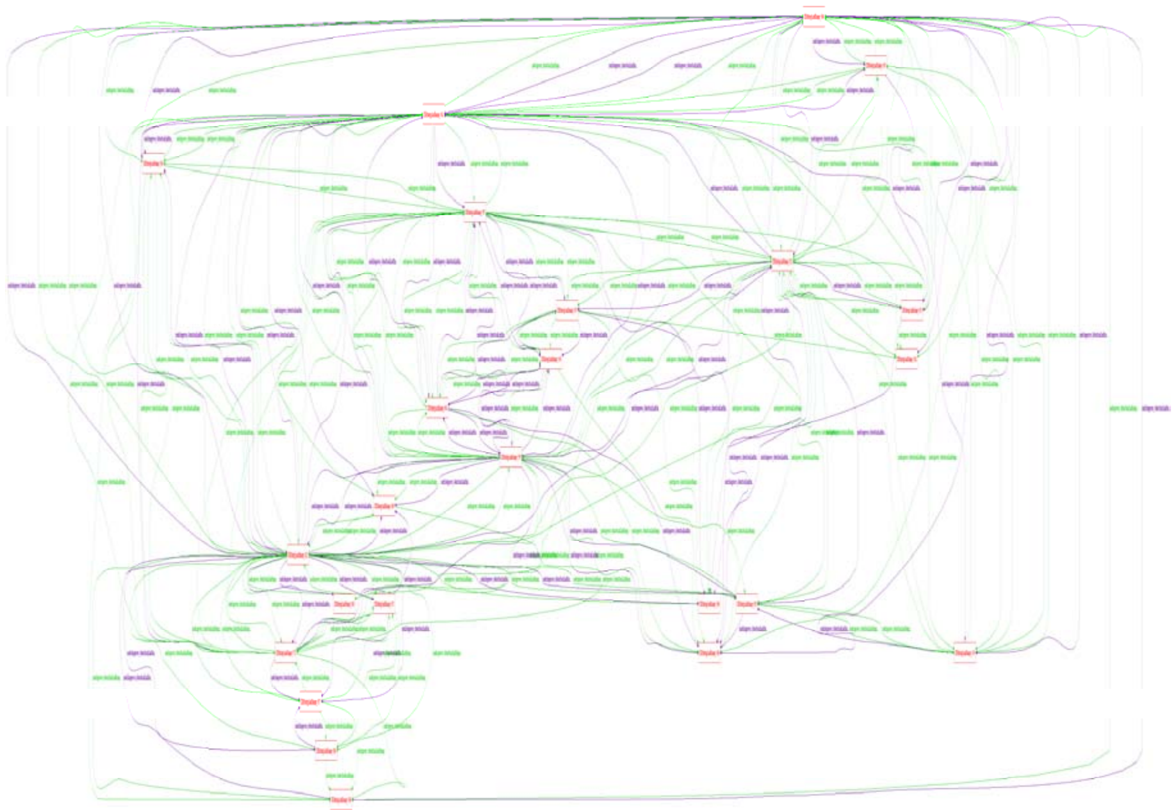


Figure 2. Example social network of approval (green links) and disapproval (purple links) among land managers (red boxes) derived from a model state ontology.

Clearly, the prototype ontology generation feature of FEARLUS model 1-0 is far from ideal. However, to the extent that diagrams such as figure 2 detailing ontologically interesting features of the state of the simulation can easily be created, the prototype ontologies generated are useful. Further, the exercise has identified a number of ways in which the process can be improved. The fact that some of these may entail making semantically significant changes to the way the code is written is a double-edged sword. On the one hand it may be undesirable from the point of view of elegant OO coding to adopt these conventions, but on the other, it may facilitate a coding style that is more transparently related to the ontology (explicit or otherwise) on which it is based. The process has also enabled us to highlight some of the differences between ontological and OO representation of a model that at least raise the question of whether description frameworks that are too closely tied to OO (such as UML) are the most appropriate approaches to representing the ontologically significant features of a model.

A user guide for FEARLUS model 1-0 describing the available functionality in full is nearing completion at the time of writing.

2.3.4 Planning for scenario development and associated model construction and use

Modelling the period mid-1980s to the present

Since the actual course of events must be considered nondeterministic, no single run of the model could be expected to reproduce the actual course of events precisely; besides which, the data available is limited by what is available to us, and specifically by confidentiality restrictions. Therefore we will be assessing whether the model, once constructed, is able to reproduce the direction and qualitative magnitudes of the trends found in the data.

The general approach to be followed is as follows:

- Draw up a provisional ontology covering the aspects of the world considered important, and how they relate to each other. Work has begun on this, enriching the "upper level" ontology developed earlier with information from the Grampian ontology drawn up by Ruth Meyer, and from Lee-Ann Small's interviews of farmers in Upper Deeside and key informants in farming-related professions.
- Select those aspects of the world (identified in the ontology) that can be represented in some way by parameters in the model.
- For each of these aspects, determine what data are available.
- Where there is data, consider how it can best be encoded in the model parameters.
- Where there is no data good enough to be worth using, run models with a range of plausible combinations of parameter values, to discover which of them are important, hence determining the range of assumptions needed to confirm or disconfirm the model in relation to the data that are available.

Lee-Ann Small's interviews have enabled us to identify a number of key trends relevant to our modelling approach over the period to be modelled, notably tendencies for farms to increase in size, reduce the number of commodities produced (for example, pigs, potatoes and dairy cattle have largely disappeared from the study area), and shift toward less labour intensive land uses. Somewhat to our surprise, it has turned out that a high proportion of the more successful farms have a likely successor to the current farmer, suggesting that at least up to 2030, family farms are likely to persist in the study area.

Social and informational networks have altered over this period, probably becoming less geographically restricted, due to the closure of local marts (farmers now attend the larger mart at Thainstone), and the widespread adoption of local 'phones. However, the existence of distinct but overlapping networks of different kinds of farmer (with the more specialist being geographically broader), and the important role of representatives of firms supplying a range of inputs in passing on both technical information and gossip seem to have held throughout.

The major external shocks (note that a "shock" is a sudden change: it can have a positive effect rather than a negative one) and stresses during the period to be covered evident from the interviews are the following:

- During the 1980s there was a rapid fall in land and output prices, with banks foreclosing on a large number of loans to farmers around 1985.
- 1992 saw the UK leave the ERM, leading to devaluation of the £ sterling. In the same year, the McSharry reforms to the CAP occurred, and IACS arable aid payments were introduced. These developments caused a sharp land price rise to 1995, which continued more gradually as cereal prices dropped.
- In 1996, news that Bse could be transmitted to humans destroyed the export markets for beef and breeding stock.
- The 2001 foot and mouth epidemic did not reach the area, but control measures affected tourism and to a lesser extent farming.
- The approach of the CAP switch to the Single Farm Payment caused great uncertainty. Effects of the change itself have so far been less marked than expected, as they coincided with a rise in product prices.
- Finally, there is some evidence that milder, wetter winters are changing stock-keeping

practices, leading to stock being kept indoors for more of the time.

Modelling the period from now to around 2030

The plan for climate change scenarios is as follows:

- Draw up a provisional ontology covering the aspects of the world considered important, and how they relate to each other. (This will be largely the same ontology as for 1987-2003, but may require some additions.)
- Decide on a small set of main scenarios: determined by the extent of climate change over the period of the simulation, and the extent and type of policy/social pressure for mitigation (for each of the variations in this regard, draw up a list of policies, social pressures and price changes).
- Select those aspects of the world and main scenarios (identified in the ontology) that can be represented in some way by parameters in the model.
- For each of these, decide how to formulate parameters for main scenarios: these will (wholly or mostly) be defined in terms of statistical distributions, from which actual time series will be constructed.
- Within each main scenario, undertake a (fairly large) number of runs, using these statistical distributions, and also varying factors unrelated to climate (such as subsidy regimes) to test the robustness of the outcomes.

Planning to model epidemic scenarios has not yet commenced.

2.3.5 Data collection

Data sources have so far been selected for past and present climate data, and for past land use and land use potential. The only other essential types of data for the FEARLUS modelling planned, concern the effects of climate on production of the commodities produced (or likely to be produced in future), and the prices of those products over the period up to the present.

The best source for land use data for our purposes is the Agricultural Census, carried out annually in June. We have gained access to data from this source from EDINA (Edinburgh University Data Library). It employs highly detailed land use categories, which will require to be aggregated for our purposes. It is georeferenced using 10km, 5km and 2km squares, although created from parish-level data.

The same source, using the same squares, also provides a classification of land suitability for particular uses which appears suited to FEARLUS:

- Urban Built up areas - unsuitable for distribution of census items.
- Water Inland water including loch/lake, pond, reservoir – unsuitable for distribution of census items.
- Upland Including heathland, moorland – suitable for distribution of certain census items.
- Woodland Including native and non-native forestation excluding orchards - unsuitable for distribution of census items.
- Restricted Agricultural Land Parks, golf courses, airfields, industrial sites, some MOD land – unsuitable for distribution of census items.
- Agricultural Land (Scotland) Agricultural Land of various grades in Scotland. [What this means has still to be investigated.]

- Agricultural Land Grade 1&2 Best and most versatile agricultural land with no or minor limitations.
- Agricultural Land Grade 3 Agricultural land with moderate limitations due to soil, relief or climate (or combination thereof).
- Agricultural Land Grade 4 Agricultural land with severe limitations due to adverse soil, relief or climate (or combination thereof).

So far as climate data and scenarios are concerned, UKCIP is the chosen source. The UKCIP02 data archive is has datasets that comprise the UKCIP02 climate change scenarios:

- 50km resolution modelled baseline (1961-1990)
- 50km resolution future climate change (2020s, 2050s, 2080s)
- 5km resolution future climate change (2020s, 2050s, 2080s)
- 5km resolution future monthly temp. time-series (2011-2100)

Observed climate averages for the UK, including 5km resolution observed climate baseline datasets (in the same format as the 5km future climate change datasets) are available from the UK Met Office. These run from early in the twentieth century to 2000, and appear to be at the right level of detail for our work, enabling a relatively straightforward classification of years in terms of growing season, prolonged hot/cold/wet/dry periods, and total rainfall over the growing season.

We will need to classify years by how good they were for growing particular crops/raising particular types of stock – perhaps 4 levels of goodness for each.

2.3.6 AMEBON

Development of AMEBON is in progress. The development will take place in three stages:

1. Creation of prototype classes to rigorously enforce an ontology (i.e. presence/absence of class/properties), and separate actors, actants and actions.
2. Integration with OWL-API (or similar) library to load and process ontologies.
3. Integration with Repast (or similar) agent-based modelling library to create simulations.

AMEBON is not now likely to be used within the CAVES project, and staff time spent on developing it is not being counted toward matched funding. However, CAVES, by revealing the difficulties in applying an existing modelling system without a clear ontological basis to a new case study, and relating it to new types of evidence, has provided much of the impetus for AMEBON's development; the approach being taken to it can therefore be counted as among the outputs of CAVES.

2.3.7 Using ontologies to link evidence to models

Research is on-going on the various ways that ontologies could be used to link evidence from case studies to model structure. One mechanism was presented, as mentioned in the last report, at WCSS 2006 (Polhill & Gotts, 2006), whilst another at EPOS 2006 (Gotts & Polhill, 2006), covered a different approach. Various other approaches have been identified. It may be that the best approach will depend on the case study.

2.3.8 The next six months

The main modelling activities over the next six months will be:

- Continue work on and with the general models of collective reward policies, begun using

FEARLUS 0-8-2-1, within FEARLUS 1-0. A subset of pollution control policies will be selected for detailed investigation.

- Begin systematic comparisons between the statistical properties of output timeseries from the FEARLUS models of point 1, and output timeseries from other CAVES models.
- Complete collection of the data necessary to construct the 1980s – present scenarios and models.
- Follow the plans for 1980s – present scenarios and modelling outlined in section 2.3.4
- Undertake stakeholder validation of those aspects of the models for which this is appropriate (see the Grampian case study report).
- Write up the results of the preceding step for publication.
- Determine changes necessary to FEARLUS 1.0 to produce a modelling system (FEARLUS 1.1) suitable for implementing the final models of the Grampian case study.
- Code FEARLUS 1.1.

2.3.9 References

Gotts, N.M (In press) Resilience, panarchy and world-systems analysis. *Ecology and Society*.

Gotts, N.M. and Polhill, J.G. (2006) Narrative scenarios, mediating formalisms, and the agent-based simulation of land use change. *Conference on Epistemological Perspectives on Simulation (EPOS06), 2nd, Brescia, Italy, 5-6 October 2006*. Online as at 05/10/2006: http://www.eco.unibs.it/ epos/Papers/GottsPolhillEPOS2006Abstract.pdf.

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

Polhill, J.G. and Gotts, N.M. (2006) A new approach to modelling frameworks. *Proceedings of the First World Congress on Social Simulation, Kyoto University, Kyoto, Japan, 21-25 August 2006*, Vol 1. pp215-222.

2.4 Politechnika Wroclawska

2.4.1 Biophysical Model for the Odra Case Study

During the last few months a simplified version of the biophysical model was under development. It was requested by the Kassel team as a preliminary step to integrating the final, more sophisticated biophysical model with their agent based model.

The simplified biophysical model, called Simple Hydro-Agricultural Model (SHAM) is a one-dimensional abstraction of a real world system. The model assumes a specified number of parcels lying on a terrain of small and homogeneous slope, along a homogeneous channel, that runs through the centre of every parcel. Each parcel has the same area, but the types of land use can vary. At this point there are only two types of land use, arable land and fish ponds, but more are to be added soon. Along the channel, there are sluice gates at the bottom of every parcel that can be operated by agents. The model works with a monthly time step and in each time step it calculates average soil water level values for every parcel. Furthermore, at the end of every simulation year, it calculates yields from every parcel. The agent that operates on a given parcel can perform certain operations like cleaning the segment of the channel that lies within the parcel, or closing/opening a sluice gate. At this point an agent cannot change the land use type during a simulation run, however this possibility will be added in the near future. Yields depend on land use type and the water regime in a given parcel. Because the model handles some spatial effects, like water spilling from a given parcel to neighbouring parcels, the yields are affected not only by phenomena taking place on a single parcel, but also by those which happen in the surrounding.

Future versions of SHAM will give agents the ability to change the land use type on their parcels and model the transition process. They will also include information about soil type in each parcel and provide a wider choice of land use types.

2.4.2 Discrete Choice (Opinion Dynamics) Models

Based on our analysis of mean-field solutions of binary choice models we extended the analysis to include different social networks structures. The attention is focused on finite-size systems of interacting individuals (agents). The sites interacting with a given agent are the only that directly influence his/her decision. The interaction structure is conceptualized as a graph (network) with edges connecting the agents that interact. Each edge constitutes so called neighbourhood relation which is assumed reciprocal - the graph is undirected. In general, we do not impose any constraints on the graph structure, however some more detailed results regard the graph that is connected, i.e. there is a path of edges connected any two sites.

First we focus on the static case, i.e. the properties of (once attained) stationary state are considered. The equilibrium distribution assigns each state (state is the configuration of decisions all over the network) a probability of occurring in stationary state. We study the equilibrium distribution relative to the considered graph. The Gibbs distribution is of special interest because of its relations with statistical physics. However, the crucial aspect emerges from the Hammersley-Clifford theorem that establishes the equivalence between the global characteristics - joint Gibbs distribution and local characteristics - the conditional probabilities for each individual's decision given the attitudes of his/her neighbours (referred to as Markov random field). In fact the problem may be inverted so that Gibbs distribution determines the neighbourhood structure. Any joint distribution over finite network may be regarded as Gibbsian relative to complete pairwise interaction structure. The question is: when the equilibrium distribution is consistent with the assumed neighbourhood structure, i.e. when it is Gibbsian relative to this structure. In particular, we focus on the special case of inter-neighbourhood conditional dependence, i.e. given total numbers of neighbours sharing the alternative attitudes (N) - we specify the constraints on the functional form of equilibrium distribution and corresponding conditional probabilities for an arbitrary connected graph.

Further, we introduce dynamics and study the evolution of the system toward the equilibrium state characterized by the equilibrium probability distribution. We focus on the relations between the agents' decisions rules and the equilibrium distribution. We model the time evolution of attitudes all over the network by a Homogeneous Markov Chain (HMC) - the sequence of random vector of binary decisions that satisfy the Markov property, i.e. the future state is directly influenced only by the present one and the history is irrelevant in determining the next step in the chain. The homogeneity reads: the rules of transition from present to future state do not depend on time. We formulate and prove the sufficient conditions for the transition probabilities of HMC to yield Gibbs distribution relative to given network as well as its special form under assumption (N). Whether these conditions are also necessary remains an open question. We investigate under what conditions the decision rules assigned to each individual coincide with the local characteristics of equilibrium distribution and we formulate the sufficient and necessary conditions for the individuals' decision rules. This coincidence means also that the neighbourhood system emerging from the stationary distribution coincides with the assumed interactions structure that drives the system toward equilibrium.

We also discuss the obtained results in the framework of utility function models. We assume very general form of utility function - the additive one with two components: the systematic (encompassing the individual, external and interaction-based preferences) and random one. There is an essential symmetry between the random (e) and systematic part (V). We study the conditions under which the pairs (V,e) leads to Gibbs distribution in equilibrium.

As a parallel effort in binary choice models theme, we investigated a continuous time version of Nowak-Latane model. A probabilistic version of this deterministic model is proposed in order to include the possibility of non-standard behaviour of individuals. In the present system interactions determine the probability of decision change per time unit instead an analogical decision rule proposed by Lewenstein, Nowak, and Latane. In our construction, the rule of individual-decision change may not lead the system tends to a stable equilibrium state. This is because every individual whose opinion is represented by the Isinig spin is a complex thermodynamical system and the informational entropy can flow from the Ising subsystem. Thus, the subsystem is thermodynamically open and the law of entropy maximization does not apply to it. Furthermore, the number of individuals can be much smaller than the number of particles expected in thermodynamics.

2.4.3 Evolution of Social Networks -- Theoretical model of evolution of an affiliation social network

The classical approach to collecting data on social networks involves asking all members of the examined population about their social links. However, obtaining full coverage of the population is usually impossible due to significant costs and high refusal rate that result in incomplete data. Incomplete data loses much of the power and value and effectively does not allow for many kinds of structural analysis. To simplify the process of data collection, data on group affiliation are collected and the structure of the network is reconstructed based on group memberships. Such a network is called a two-mode network or an affiliation network. It is then assumed that each two individuals belonging to the same group must know each other (form a link).

However, the assumption that all members of any given group know each other is unrealistic. In larger groups (or groups that do not operate long) it seems more appropriate to talk about increased probability of knowing each other. Presented model aims to generate a social network based on affiliation data, with an underlying assumption that shared group membership (affiliation) increases the probability of link formation but it does not assure it.

To diagnose whether obtained network structure is typical for a social network, the clustering coefficient (indicating existence of well interconnected subgroups) and degree correlation (correlation between degrees of neighbouring nodes) are computed. Both indexes were shown to have significantly higher values for social networks than for other types of networks.

The model shows that increasing probability of interaction with group members causes increase of both the clustering coefficient and the degree correlation, as predicted. The results of simulations offer a new hint towards understanding of human tendency to socialize with similarly social people. It turns out to be at least partly driven by the tendency to build the network of social connections through group memberships – as opposed to the classical explanation of (semi)intentional search of people with similar individual characteristics, such as age, sex, education or sociability (so called homophily). In fact, many formal or even informal groups gather people who share common characteristics. The model shows that the actual content of this similarity is not the only driving factor for creation of a social network and the mere group affiliation is a sufficient factor to obtain a network that will show characteristics typical for social and only social networks.

This method of generating network structure seems a promising tool to simplify the process of data collection for social networks. However, it should still be validated to what extent the obtained structure matches the actual structure of the examined network. The validation should ideally take place in an experimental or quasi-experimental setting in which it would be possible to map the network by asking all respondents for their social links and group affiliations (to obtain the structure most similar to the "actual" network). Then a comparative analysis between the "actual" network, the network based on the classical affiliation approach and the network generated by our evolutionary model should be performed. Further works should also include introduction of overlapping group memberships in our model as well as further sensitivity analysis.

2.4.4 Plans how to proceed

We will continue our work on different versions of the Odra biophysical model and, together with the Kassel team, we will test their applicability and usefulness.

We will also continue to analyse binary choice models on complex networks. We will also apply such models to collective action problems.

The structure of social networks has proven to be very difficult to elicit directly from farmers. Building on our preliminary results of networks built based on affiliation, we will improve these models and try to validate the results.

3 Case studies

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

3.1 South African case study (SEI Oxford)

3.1.1 Introduction

Over the past 6 months the SEI team has focused on three areas linked to the South African case study. Firstly, there have been activities to inform the model development by researchers feeding information back to the modellers. Secondly, a validation working paper has been produced and thirdly, a fieldtrip has been planned for April that assesses assumptions adopted in the case study model. These are described in more detail below.

3.1.2 Informing model development

Through an iterative process between the modellers and researchers, modellers' assumptions have been refined. This process has also included the modellers speaking to researchers about the areas they are interested in developing further in the model. The researchers have provided as much information as they can and sought further information to add to the knowledge base. In particular, there has been a focus on social networks related to savings clubs and village social dynamics.

3.1.3 Validation working paper

The SEI has been pulling together information from the project partners for a validation working paper. The aim of the paper is to discussion validation and its role in the project. Validation is used to show the validity of agent based models and their findings without forfeiting their ability to capture complex social realities. The paper includes the following:

- Section 2 reflects on the broader issue of validation, and particularly on the challenge of validating agent-based models. The section contains perspectives from people working in field-based studies and from people who approach validation as a model issue.
- Section 4 involves a synopsis of protocols for model validation that may be useful as a basis for thinking and practice in this area. This section includes actions planned in the CAVES

project such as methodology, approaches and protocols. It also includes model design issues and the manner in which they impact upon the validation process.

- Section 5 involves practical steps - validation protocols - that can be taken to aid validation in the CAVES project.
- Section 6 begins with an overview of the three case studies, focussing on how agents were designed and the role of stakeholders in that process. Plans are noted in the case studies for further stakeholder engagement and validation exercises.

3.1.4 The first step of a validation exercise: Selection of assumptions

The fieldwork trip in April will try and test some of the assumptions used in the South African model. There are too many assumptions to be tested within the 4 days of the April field work. So, it is inevitable to select a sub-set of assumptions from the entire assumption lists. One of the most important issues about validation is the robustness of the process, i.e. the validation process has to be replicable and transparent for future references. Therefore, the selection process has to be as objective as possible. However, the objective selection process such as random selection may not be optimal as most social science assumptions may not be possible to validate. On the other hand, it is also important to consider the usefulness of a social science model for the user of the particular model. In this case, it may be justifiable to focus certain assumptions to make model more useful. For example, a local authority may be interested in the social network aspect from an ABM as it is not captured by other conventional models.

We propose a two-step process to select assumptions: 1) random selection, and then 2) subjective systematic selection of assumptions. In the first step, we will sample assumptions completely randomly. If too many of these sampled assumptions are not executable due to lack of time, resource, etc., the random selection approach will not be practical. This 1st step, itself, is a finding for this validation exercise, i.e. assumptions are hard to validate. In this case, we will switch to the subjective systematic selection, which is based on the testability of assumptions, the importance and interests of model for the users of the model, and the availability of previous validation exercises. The system of selection process has to be recorded as much as possible to try to be replicable and transparent. This selection process can be considered as KISS, TAPAS, and GEF, i.e. the selection process is stylised, based on previous work, and for the purpose of model.

3.1.5 April Sekhukhune fieldtrip overview

The aim of the fieldwork is to go back to people and places where previous research was undertaken, to

- Feedback results from research
- Test assumptions based on previous research as outlined above
- Get input on how to take research forward to impact on policy and practice

The fieldwork will be carried out in Tubatse municipality, Sekhukhune District, Limpopo Province, focusing on the village of Ga-Selala. The fieldwork will start by interacting with the village stakeholders. Various methods might be used to test assumptions, including role-play, participatory video and structured questions. Other researchers in the area will also be contacted to get their feedback on assumption. In particular, the Radar staff who have worked in this area will be contacted.

There will be a workshop at the municipal scale (in Burgersfort) after the initial village fieldwork. Again, there will be feedback and some exercises to test assumptions. Village stakeholders and other key stakeholders (provincial or district stakeholders, researchers) will also be invited to the meeting.

3.2 Grampian case study (Macaulay Institute)

3.2.1 Overview

During the September – March time period, Dr. Small completed the primary field research, initiated a new knowledge transfer activity, and worked with the Grampian team to draft a validation protocol. Data analysis of the primary field research remains ongoing. As part of her post-doctoral position on the CAVES project, Dr. Small has also spent considerable time preparing and revising papers for publication. In order to facilitate her ongoing involvement in the CAVES project, as of January 1, 2007 she moved from 1 to 0.5 Full Time Equivalent on the CAVES project, thus extending her work on the CAVES project (which was due to be completed August 2007) until the end of the CAVES project (March 2008). Timelines for Dr. Small's activities in the project have been revised to reflect this altered time allocation.

3.2.2 Primary Field Research

Primary field research was completed in March. In addition to the interviews completed in the March – September 2006 period, Dr. Small completed interviews with two identified farming successors (to farmers in the study), one key informant interview and two estate manager interviews. This brings the total interviews to date to 44 (49): 24 farmers (29 including spouses or partners also participating), 6 identified successors (to farmers in the study), 5 estate managers, and 9 key informants. Dr. Small has continued her participant observation of key farmer interaction settings, attending a biofuel conference and a National Farmers' Union meeting, as well as interactions included in the knowledge transfer activity. Data analysis of the field research is ongoing.

3.2.3 Knowledge Transfer (KT) Activity

Dr. Small, in conjunction with Dr. Polhill and Dr. Gotts have accessed funding through the Scottish Executive and Rural Affairs Department to disseminate CAVES-Grampian study findings to stakeholders and key informants. This funding is part of a Scottish Executive initiative to increase communication about research findings to research participants and other stakeholders.

Primary activities include:

4. Networking with the agricultural community at agricultural events to discuss project goals and identify key stakeholder issues. (e.g. demographics and Single Farm Payment).
5. Development of a checklist of issues for use in consulting with people in different sections of the agricultural community (i.e. farms, advisors, equipment suppliers, estate managers).
6. Development of materials developed in CAVES for use in networking with members of the community at events (e.g. an information note on project findings to date).
7. Circulate an Information Note on land manager perspectives on drivers of land use change to members of the agricultural community, SEERAD, relevant public organisations (through face-to-face interactions at agricultural events) and via the world wide web.
8. Feedback/evaluation to assess whether Institute activity in this topic was known beforehand, means of engagement (i.e. face-to-face), and relevance of Information Note.

The knowledge transfer project was initiated in October 2006 and will continue until June 2007. Activities to date include attendance at the Thainstone Christmas Classic and the Royal Northern

Spring Show, and the creation of a two page information note. Results of the feedback activity indicate that the study findings identified in the information note are consistent with the experiences and perspectives of farmers who reviewed the note.

3.2.4 Validation Protocol

The validation protocol has been drafted in conjunction with Dr Polhill and Dr Gotts. A final version will be completed after consultation with other CAVES project members at the March meeting in Vienna.

The draft validation protocol focuses on validation of the modelling framework components: comprehensiveness and relative importance of factors in land use change, principles of land use change, and the decision-making process. As agent-based models are not intended to be predictive, the outputs of the models will not be validated by stakeholders. The focus is instead on model inputs and processes. Heuristically, the FEARLUS models generated through the CAVES project follow the 'TAPAS' approach (Take A Previous Model and Add Something). Emphasis in validation is therefore on demonstrating improvements to the model resulting from CAVES field research. Similarly, the particular focus on social networks within CAVES is also reflected in the validation protocol.

Individual interviews will be completed with 8 – 10 land managers and stakeholders (a combination of old and new respondents). During the interview, a variety of techniques will be utilised, including open ended questions, ranking of identified factors, future scenario discussion and decision-tree processes. Results will be combined with validation activities in the KT project, which focuses more generally on the principles of land use change processes.

3.2.5 Academic publications

As part of her post-doctoral position, Dr. Small has been writing and submitting papers to peer reviewed journals. In this reporting period, she submitted two papers: "East Meets West: Utilising Western Literature to Conceptualise Post-Soviet Agrarian Change", which has been accepted to the Journal of Peasant Studies; and "Farm Family Coping With Stress: The Impact of the 1998 Ice Storm", which was submitted and subsequently revised for the Journal of Comparative Family Studies. She also completed revisions on a paper written earlier in her post-doc, entitled "Sustainable (Rural) Livelihoods: A Critical Review", which has been accepted to the Canadian Journal of Development Studies. Both accepted papers are scheduled for publication in 2007.

3.2.6 Directions for Future Work

Primary activities for the next six months are as follows:

- Data analysis (and paper writing)
- Implementation of validation protocol; subsequent feedback to modelling team.
- Completion of CAVES primary field research report.
- Initiation of fourth wave of field research. This is expected to begin in October, after harvest. This research will reflect the results of the validation process.

3.3 Odra Valley case study (Uniwersytet Wroclawski)

3.3.1 Description of the Work Undertaken

Field work and data analysis

After rethinking the results of the first wave of field research, the interview questionnaire was significantly modified in order to get more deeply in the institutional aspects of collective action and land reclamation problems, as well as economic issues, which seem to be of great importance. Thematic modules regarding land use remained almost the same.

The second wave of field research was conducted in 10-11. 2006. Next 17 interviews were performed in villages Rogow Legnicki, Kawice and Kwiatkowice. This means that altogether we have 30 interviews regarding land use, land reclamation and social networks in the study area. Processing and analysis of the results should be finished soon.

The first version of the agents' types and rules set for the social model was written together with Kassel team, based on narrative storylines developed in summer. This work included short visit in Kassel 23-24.11.06.

The rules set concerning relations between soil classes, economic factors and land use was developed and validated with the key domain expert.

Representativeness indicators for the study area were developed based on statistical data on agriculture and land use.

Spatial interdependency of agents along channels of land reclamation system was elicited from expert knowledge during the set of meetings with the agricultural engineers and farmers.

Validation

A document "Odra Case Study Contribution to Deliverable No. 8 Working paper on case study structure, stakeholder/agents and validation data" was written and submitted to SEI team working on Deliverable 8. This document contains: short description of both technical as well as social aspects of LRS maintenance, brief summary about activities in stakeholder participation, chapter on agents design including construction of networks and properties of agents types, detailed part regarding study area representativeness, and finally some suggestions for validation.

Karolina Krolikowska participated in a workshop 'Vulnerability and Validation Roundtable: From the local and tacit to general and formal, and back again' organized by SEI in Oxford, 15-16.02.2007 and devoted to validation issues.

Due to availability of data for Strzegomskie Hills region, it was taken into consideration as a candidate for the biophysical model validation. In order to asses the applicability of these data for validation, the analysis of topographic and other environmental properties of the region was conducted and compared with Odra river case study region parameters. The significant differences found, let to the conclusion that it is not suitable for biophysical model validation.

GIS

A drainage enforcement tool was developed, as a GRASS GIS extension program v.breach, in order to improve the hydrological soundness of the biophysical model. It has been released under GPL license, on the GRASS GIS WIKI page, in vector addons section: `http://grass.gdf-hannover.de/wiki/GRASS_AddOns#Vector_add-ons`

Another GRASS GIS extension program, v.line.center, was developed as a tool supporting the transfer of attributes from the land cover vector polygon data into watercourses vector line data. The tool was necessary to perform the analysis of the hydrological net structure and density per land use/land cover type and per land owned by particular agents. v.line.center is available on http://grass.gdf-hannover.de/wiki/GRASS_AddOns#Vector_add-ons, under GPL license.

3.3.2 Next steps

- Final social research report will be written and submitted to CAVES Knowledge Base.
- Three publications are planned, probably to the following journals: Landscape Ecology, Agricultural Systems, and Agriculture, Ecosystems and Environment.
- A protocol for validation should be established based on ABM simulation outputs and implemented.
- Further social research in other villages within the area of Odra case study region will be conducted, using the same method. It will enable us to assess the representativeness of the social situation in Rogow Legnicki and provide reference data for validation.

4 Publications

The CAVES project is dedicated to publish results in relevant scientific journals and conferences. Nearly all of the submitted papers from the last reporting period have been accepted at their respective conferences and presented successfully.

During the fourth six months of the project, the following papers and posters have been written:

- S. Alam, R. Meyer, E. Norling: Agent-based Model of Impact of Socioeconomic Stressors: A Dynamic Network Perspective. Accepted for AAMAS 2007 (International Conference on Autonomous Agents and Multi-Agent Systems), Hawaii, May 14-18, 2007
- S. Alam, R. Meyer, S. Moss, G. Ziervogel: Impact of HIV/AIDS in the Context of Socioeconomic Stressors: An Evidence-driven Approach. Submitted to JASSS (Journal of Artificial Societies and Social Simulation), currently under peer review.
- N. Gotts: Resilience, panarchy and world-systems analysis. Ecology and Society, in press.

5 Deliverables

5.1 Current Deliverables

There is one Deliverable due at the end of project month 24. This is Deliverable No. 9, the prototypes of the coarse grain models. All three prototype models – one for each case study -- are now available from the CAVES CVS server.

Deliverable No. 8, the working paper on case study structure, stakeholder/agents and validation data, was due in the last reporting period but has been delayed in order to enable us to base the working paper on experience instead of an ideal outline. As expected in the last progress report, this delay has enhanced the value of the Deliverable significantly. It will be available in project month 25 from the CAVES internet portal at http://cfpm.org/caves/CAVESWiki/Publications.

5.2 Future Deliverables

There are no Deliverables due at the end of the next six months of the project (project month 30).