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

CAVES Progress Report No. 5

April 2008 • Project 012816: CAVES –
Complexity, Agents, Volatility, Evidence and Scale

Duration: 2005-2008

Funded under the EU 6FP NEST programme.

Table of Contents

1 Introduction

This is the last progress report from the CAVES Project. The end of this reporting period marks the end of the project.

The purpose of the CAVES project was to couple policy concerns for complex human-environmental systems with linked physical, biological and social models based soundly in complexity science. We have produced a constructive demonstration of modelling procedures for the formation of social policy in conditions of uncertainty due to complexity. The demonstration entailed the development of clusters of models for each of three case studies.

All objectives have been met:

1. *Evidence*: The substantial qualitative and statistical evidence brought to the project by partners responsible for the case studies was developed further to inform the design and stakeholder validation of models of land use or HIV/AIDS, respectively.
2. *Models*: Agent based social models linked to biogeophysical models were developed to investigate why some external shocks to complex social networks are followed by volatile episodes and some are not. The models were based on and constrained by qualitative and statistical evidence.
3. *Generalisation*: Clusters of consistent models were developed at several levels and axes of aggregation that, at the most descriptive level, captured the relevant detail of complex systems and, at the most abstract level, allowed for ready comparison with models in the literature on complex systems.

For a complete overview of the work undertaken, the methodology applied, the achievements made, and the impact of the CAVES project we refer to the final report (Deliverable 16), which is available online from the project web-site at http://caves.cfpm.org.

As part of the final project meeting in Manchester, an international workshop has been held aimed at both academic experts and key officials involved in climate policy in general and land use issues in particular. The objective of the final workshop was both to present our results and to engage the experts and officials in exercises with the procedures developed by the CAVES project. A report of the workshop including a list of participants and speakers can also be found on the project web-site.

2 Modelling issues

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

2.1 Centre for Policy Modelling

Work has concentrated on further developing the declarative model for the South African case study to incorporate the mining scenario and allow for more statistical output to be collected in order to support evaluation of different policy measures. We have also continued our work on social networks and statistical signatures and have undertaken a model-to-model comparison of the procedural and the declarative model with regard to the sexual mixing scheme these models apply. In addition, several academic papers have been written and submitted to international conferences.

2.1.1 Declarative Model

The declarative model has been extended to incorporate mining. It has also been adapted to include different transmission rates for the different stages of HIV, as reported in a recent study from Uganda (Wawer et al. 2007).

Mining

Due to the lack of empirical data, the representation of mining in the declarative model is largely based on assumptions. A mine will need a certain number of skilled and unskilled workers, some of which may be recruited from the local population. In the current version only unskilled jobs are advertised since skilled jobs require an engineering or management degree and we know from anecdotal evidence that only a fraction of the village population achieves tertiary education of any kind.

When a mine opens in the case study area, it will attract workers from outside the region (in-migrants). We assume that the mine already brings a number of them with it at the start, while more will arrive over the course of the simulation to apply for jobs. As we don't have any quantitative data concerning numbers and dates of arrival this is modelled via statistical distributions: a negative-exponential distribution for the arrival of in-migrants sand a normal distribution for the number of in-migrants arriving at the same time. Further distributions determine the characteristics of the in-migrants like age (Normal distribution with mean 29 and standard deviation 7 years), previous mining experience (Random uniform between 2 and 6 years) and HIV infection (probability of 30%). We assume that all in-migrants are male.

In-migrant mine workers are only relevant to the modelled village population in their direct interactions with them. These include: establishing friendships with co-workers from the village, having sexual relationships with women from the village and renting accommodation. Therefore these are the only aspects considered in the model. This means for example that in-migrants aren't assigned to households and their home economy isn't computed at all.

Most mine workers will stay in hostels around the mine or a nearby town, thus being outside the model's scope. Some in-migrants will try to stay in the village, though, and rent a shack from households who (a) have one available on their premises (assumed value: 40%) and (b) are in need of money. They will then form neighbourhood links with their landlord household and its neighbours.

As in-migrant mine workers are represented as agents of the class `Person` like all other individuals in the model, they use the same endorsements mechanism to find friends and sexual partners. We have introduced a new endorsement token "has-job" to allow female agents to take this fact into consideration when evaluating their pursuers. There is now also a chance of 10% for any male agent to encounter a possible female partner randomly, as opposed to being restricted to agents known via friends, neighbours or joint activities.

Any in-migrant who can't get a job after 4-6 months or who develops AIDS (stage 4 in the progression of HIV) is assumed to leave the area (return home). The latter is in accordance with a "current trend of repatriating incapacitated mineworkers to their homes", where incapacitated refers to "injured ex-mineworkers that have been rehabilitated and mineworkers with terminal illnesses"[1].

Mines are represented by the class `Mine`. They will advertise jobs when there are vacancies (number of workers < demand) and evaluate applicants according to their

[1] http://www.teba.co.za/products_mineworkerbenefits.asp

experience. When not enough experienced applicants are available to fill the vacant posts, a mine will employ up to 10% inexperienced applicants, thus giving local people a chance. Usual contract lengths seem to be 12 or 18 months[2]. In the model, experienced workers (experience > 4 years) will be given a contract of 18 months, all others receive a contract of 12 months. If a mine worker develops AIDS his contract is ended (see above).

HIV Transmission Rates

In the first version of agent-based transmission of HIV infection, the declarative model made use of a constant transmission rate of 0.003 per coital act[3]. Recent studies have found that infectivity is actually dependent on the stage of disease; a newly infected person having a much higher chance of transmitting the virus than a person in a later stage. This can influence the rate of spread of the infection in the community.

To incorporate these findings into the model, both stages of HIV progression and different transmission rates according to stage have been introduced. The following table gives the transmission probabilities for different stages used in the model, which have been adapted from a study in Uganda (Wawer et al. 2007):

Stage	Average Duration	Transmission Probability
1	3 months	0.0082
2	12 months	0.0015
3	96 months	0.0007
4	14 months	0.0028

As before, selected agents are initialised with HIV/AIDS based on their gender and age at the start of the simulation. In contrast to the previous model version, these agents are now also assigned one of the different stages of HIV. The probabilities for the stages are directly related to their respective average duration.

2.1.2 Model-to-model comparison: Sexual mixing schemes

Model-to-Model (M2M) analysis has grown in recent years into a major research area in the field of multi-agent based simulations as more and more researchers from different backgrounds apply agent-based modelling to the same problems. As the M2M 2007 workshop website reads, "the specific aim of [M2M is] to sustain the development of techniques for comparing models used in social simulations or computational social science". One of the primary aims of such exercises is to identify suitable abstraction levels where model comparisons can be made across different scales, statistical signatures as macro-level output, and micro simulation results with regard to particular case-studies addressing similar social problems. The question we address is whether different styles of implementations – in our case declarative vs. procedural programming – lead to qualitatively different model results.

We have undertaken a model-to-model comparison of the procedural and the declarative model with regard to the sexual mixing scheme these models apply. Both models assume a traditional society where males take the first step in a sexual partnership. Thus male agents look for potential partners whereas female agents may accept or reject the courtship offers. In both schemes, incest and inbreeding are prohibited.

The formation of sexual networks takes into account that people may have several concurrent sexual partners. The only exception is married female agents, who do not take part

[2] http://www.teba.co.za/products_mb_other.asp#cdp
[3] http://library.med.utah.edu/WebPath/TUTORIAL/AIDS/HIV.html

in extra-marital sexual relationships. Since same-gender partnerships are more or less taboo in the case study area the models only consider heterosexual relationships. The resulting sexual networks are therefore two-mode networks with males and females forming the two distinct sets of nodes.

Procedural Model: Scheme based on Simple Aspiration and Quality Measure

In this scheme, we assign aspiration and quality ("attractiveness") log-normally to agents when they are created (cf. Heuveline et al. 2003, Simão/Todd 2003). If the number of current partners is lower than his upper limit a male agent of age between 16 and 55 looks for a female between 14 and 40 whose attractiveness exceeds his aspiration level. He then sends her a courtship offer. Female agents evaluate all offers received during the current time step. All offers from agents whose attractiveness is below their own aspiration level are rejected immediately. From the remaining suitors a female agent will choose the best and accept his offer. The criteria to determine the best suitor change with age. For instance, young female agents prefer males of similar age, while more mature female agents may prefer unmarried employed suitors.

Agents search for sexual partners mostly among their friends and acquaintances. The friendship network is dynamic and based on the rules outlined by Jin et al. (2001) for evolving clustered networks. There is also 5-10% chance for picking a female as potential partner at random. As the agents' search is dependent upon their friends and friends-of-friends circles, there is high likelihood that the potential partner is of similar age.

Agents without partners have their aspiration level successively decreased after a particular waiting time. For those satisfied with their current sexual partner(s), the aspiration level is updated incrementally. This is a simplified version of the original aspiration and choice criteria as described in (Simão/Todd 2003).

From the anecdotal accounts of the villagers we know that it takes about 1-2 years before a couple gets married. Since a male has to offer lobola (bride price) to the female's household before marriage, the marriage of the couple may further be delayed. We initialize each agent with a maximal courtship duration sampled from a log-normal distribution with the minimum and maximum cut-off values as 12 and 36 months. The courtship duration time (i.e. the time for dating until an agent decides to marry) is updated with age.

Declarative Model: Scheme Based on Endorsements

The declarative model makes use of endorsements and individual tags in the agents' decision making. Endorsements can be thought of as labels used by an agent to describe certain aspects of other agents[4]. Some endorsements are static in that, once applied, they don't change over the course of the simulation (e.g. is-kin), while others are dynamic and may be revoked or replaced according to an agent's experiences (e.g. is-friend, has-job) (Werth et al. 2007). All agents use the same list of endorsements but differ in how they assess them.

In case of the sexual network, endorsement values are used to model the attractiveness of potential partners. A male individual finds a female attractive if (i) she is of the same age or younger and (ii) her overall endorsement value is higher than a certain threshold particular to the male individual. We thus assume that partner choice is biased by a preference for particular attributes.

[4] 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.

Adults between 15 and 64 are considered to be sexually active. Males of this age group look for potential partners among the siblings of friends, work colleagues and the social groups they belong to. When they encounter a female adult they find attractive (see above) they make a pass at her, modelled as sending her a partner request.

Females evaluate all requests received at the same time and pick the best of the "applicants" if (i) their overall endorsement value is higher than the female's threshold and (ii) in case they already have a certain number of current partners, if this applicant's endorsement value is higher than the lowest endorsement value of the current partners. In this case, the current partner with the lowest endorsement value is dropped, thus ending their sexual relationship.

Relationships otherwise may end due to marriage of one of the partners or – in absence of detailed knowledge about real reasons – are broken up randomly. The probability applied is influenced by the number of current partners and the person's age; it is highest for young people with a maximum number of concurrent partners.

Simulation Results

In this section we compare the results from the two sexual-mixing schemes introduced above: the first based on aspiration and quality (scheme-1) and the second based on endorsements (scheme-2). A typical run for each scheme is presented comprising of 35 (simulation) years. We apply typical network signatures to compare the structure of the resulting sexual networks and look at HIV prevalence and population size as relevant macro-level measures.

Both models use the probabilities of HIV transmission reported by Wawer et al. (2005). The chance for mother-to-child HIV transfer was assumed as 30% (Newell et al. 2004). Finally, both models were initialized with detailed household composition data derived from RADAR surveys in the Sekhukhune District, South Africa.

Characteristics of the Simulated Heterosexual Networks

Heterosexual networks are by construction two-mode and exhibit distinctive signatures different from other social and sexual networks. Typical signatures for heterosexual networks include spanning-tree like structure, large geodesic distance and diameter for any two connected pairs of nodes, low density and low clustering (Bearman et al. 2004, Robins et al. 2005). Another important characteristic of heterosexual networks is the existence of very few cycles (i.e. 4-cycles and cycles of higher order).

As the study of an empirical heterosexual network [3] shows, spanning-tree like structures and few cycles reflect the influence of a bias in partner selection, which is not observable in random mixing schemes (Koopman et al. 1997, Kretzschmar/Morris 1996). Large geodesic distance and diameter of the network also reflect that a majority of the individuals (nodes) may not be aware of others' sexual links in the population. In the case of HIV spread, being linked in a long chain of people can drive the prevalence much faster.

Figure 1 shows snapshots of the sexual networks resulting from the two schemes taken in the middle and towards the end of the respective simulation runs. As can be seen, the networks from both schemes exhibit the characteristic spanning-tree like structures and few cycles. In addition, a number of 3-stars and dyads can be found. 3-stars are indicative of promiscuous behaviour (as are the spanning-tree like characteristics of the larger components) while dyads may include married couples where the male was unable to find another female sexual partner.

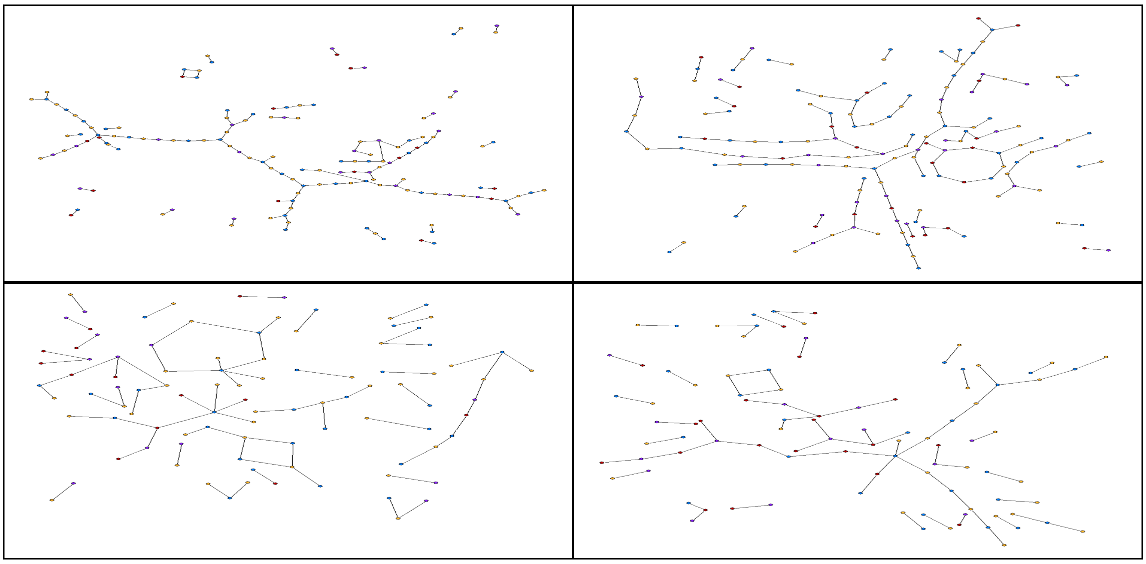
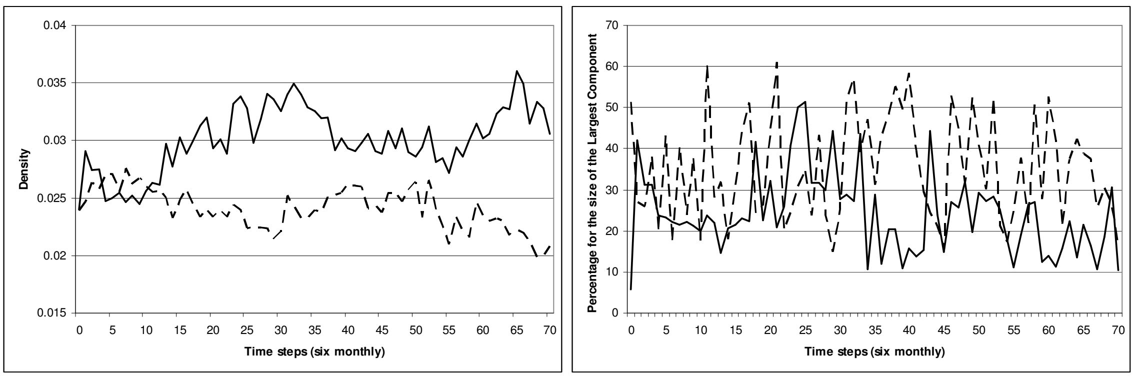


Figure 1: Snapshots of the heterosexual networks resulting from scheme-1 (top), taken at the 21st and 33rd year, and scheme-2 (bottom), taken at 21st and 35th year.

To compare the differences of the two simulated networks we calculated some basic characteristics for every sixth month using Pajek[5]. Membership in the simulated sexual networks requires the individuals to have at least one sexual partner (married or unmarried) at a particular time step.

Figure 2 (left) shows the overall density for both networks. The decline in network density for scheme-1 coincides with the decline in the overall population size (cf. Figure 4; right). On the other hand, the stable population in the first 20 years for scheme-2 enables sexually-active agents to form more concurrent partnerships than in scheme-1. The weekly time step used in the declarative model (scheme-2) as opposed to the monthly time step in the other model (scheme-1) might influence this outcome as well. We will need to investigate this further.

Comparing the relative size of the largest network component depicts volatile effects in both cases (Figure 2; right). However, scheme-1 shows a relatively higher proportion than scheme-2. This can be explained by a chance, albeit small, for male agents to randomly pick a female partner. This may result in merging two or more smaller components into one large connected component. A more significant measure of comparison would probably be the complete distribution of components' sizes, which we will aim for in future work.



[5] http://vlado.fmf.uni-lj.si/pub/networks/pajek/

Figure 2: Density (left) and percentage size of the largest component in the network (right); dashed line represents scheme-1, solid line scheme-2.

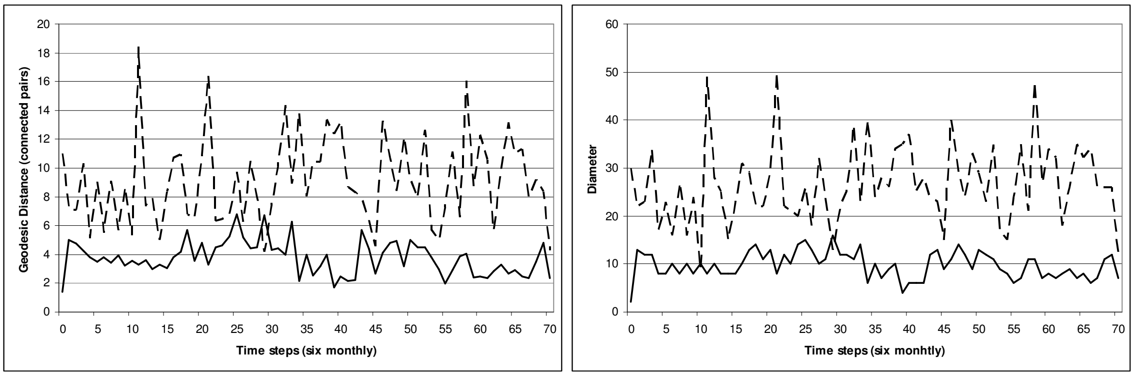


Figure 3: Geodesic distance (left) and diameter of the networks (right); dashed line represents scheme-1 and solid line scheme-2.

Figure 3 (left) depicts the geodesic distance for any connected pair of individuals and figure (right) shows the diameter of the two networks. Unsurprisingly, scheme-2 exhibits a much lower geodesic distance due to the higher density of the network. Similarly, the diameter of the network for scheme-1 is generally higher since the size of its largest component is almost always greater than for the network of scheme-2. It shows that scheme-1 has more long 'chainlike' structures as observed by (Bearman et al. 2004) although we do not know if similar characteristics are found in our case study region as well. As both the schemes do not assume a fixed population size, the size of the network is also a major influence on the outcome of these network statistics. As next steps we will investigate the subgraph characteristics for comparing networks of different sizes.

HIV/AIDS Prevalence

As expected, the different network structures arising from the two different sexual-mixing schemes have different outcomes on the macro level. It is interesting to observe clustered volatility in the time series of the network characteristics (cf. figures 2 and 3). This effect, however, is not significant in the prevalence of HIV/AIDS and the population size.

Figure 4 (left) shows the prevalence of HIV/AIDS over 35 years for the two sexual mixing schemes at a monthly scale. In both cases, a number of agents have been initialized as HIV-positive at the start of the simulation, based on the current HIV/AIDS prevalence in the case study region. While in the case of scheme-1 prevalence progresses slowly it quickly reaches about 23% in the case of scheme-2. This can be explained by the higher density of the sexual network and the time scale of the declarative model, where new cases of HIV (via sexual transmission per coital act and external incidence among migrants) are introduced at a weekly basis. After a short decline the prevalence reaches a plateau of about 20% around the 15th year (month 180). In contrast, with scheme-1 the prevalence shows a significant drop between months 225 and 275 (year 18-23) and then climbs again.

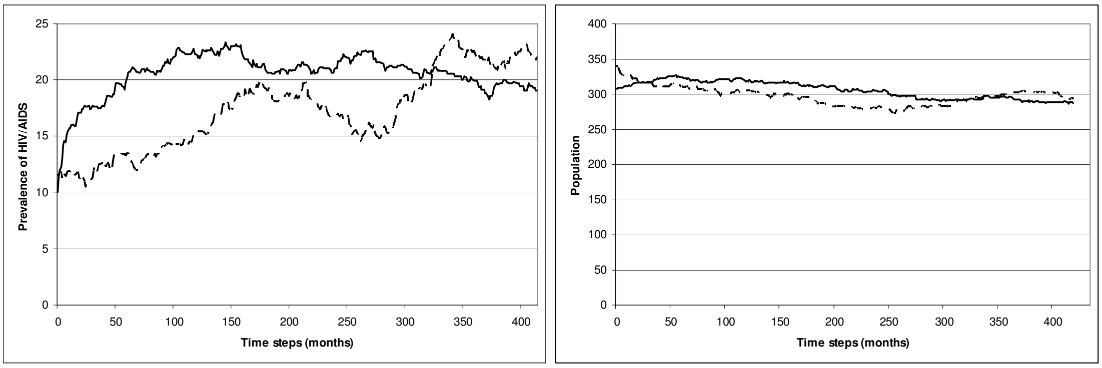


Figure 4: HIV/AIDS prevalence (left) and population size (right) over 35 years for a typical simulation run; dashed line represents scheme-1, solid line scheme-2.

Typically, the spread of an epidemic is characterized by the rise in the prevalence, followed by a decline and then attaining equilibrium for some time. Scheme-2 shows this characteristic of the curve while scheme-1 shows more volatility. As both simulations were run for only 35 years, it would be interesting to investigate whether the prevalence stabilizes in the next 50 years for the two cases.

Conclusion

Evidence-driven modelling of the spread of HIV in the Sub-Saharan region requires extensive fieldwork, clinical trials and involvement of the local stakeholders. Sexual interaction among promiscuous heterosexual partners has been identified as the primary cause of HIV/AIDS prevalence in the region. This involves individuals' sexual preferences, migration and the role of commercial sex workers. In the absence of detailed empirical data, any individual-based modelling of the spread of HIV has to depend on plausible assumptions regarding the sexual behaviour of the population. It is therefore important to investigate the effects of a chosen model mechanism on the resulting sexual network and overall simulation outcomes.

2.1.3 Publications

In the period of October 2007 to April 2008, the following papers have been written:

- Shah Jamal Alam, Ruth Meyer and Emma Norling: *A Model for HIV Spread in a South African Village*. Accepted for MABS 2008, 9th International Workshop on Multi-Agent-Based Simulation, May 12-13, Estoril, Portugal.
- Shah Jamal Alam and Ruth Meyer: *Comparing Two Sexual Mixing Schemes for Modelling the Spread of HIV/AIDS*. Submitted to WCSS-08, World Congress on Social Simulation 2008, George Mason University, Fairfax, July 14-17, 2008. Online available as CPM-Report No. 08-189 <http://cfpm.org/cpmrep189.html>
- Scott Moss: *Simplicity, generality and truth in social modelling. Part 1: Epistemological issues.* Submitted to WCSS-08, World Congress on Social Simulation 2008, George Mason University, Fairfax, July 14-17, 2008.
- Scott Moss: *Simplicity, generality and truth in social modelling. Part 2: Demonstration.* Submitted to WCSS-08, World Congress on Social Simulation 2008, George Mason University, Fairfax, July 14-17, 2008. Both parts are online available as CPM Report No. 08-187 <http://cfpm.org/cpmrep187.html>

2.1.4 References

Bearman, P.S., Moody, J. and K. Stovel. Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks. American Journal of Sociology. 110(1). 44-91 (2004)

Cohen, P. Heuristic Reasoning – An Artificial Intelligence Approach. Pitman (1985)

Heuveline, P. et al.. The Structure of an Epidemic: Modelling AIDS Transmission in Southern Africa. Proc. Sym. Agent-based Computational Modelling, Vienna (2003)

Jin, E.M., M. Girvan and M.E.J. Newman. Structure of growing social networks. Phys. Rev. E. 64.4 (2001)

Koopman, J.S. et al. The role of early HIV infection in the spread of HIV though populations. J. Acq Imm Def Syn Hum Retrovirol. 14. 249-258 (1997)

Kretzschmar M. and M. Morris. Measures of concurrency in networks and the spread of infectious disease. Math Biosci. 133(2). 165–195 (1996)

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

Newell, M. et al. Mortality of infected and uninfected infants born to HIV-infected mothers in Africa: a pooled analysis. The Lancet. 364. 1236-1243 (2004)

Robins, G., Pattison, P. and J. Woolcock. Small and Other Worlds: Global Network Structures from Local Processes. American Journal of Sociology. 110(4). 894-936 (2005)

Simão, J. and P.M. Todd. Emergent Patterns of Mate Choice in Human Populations. Artificial Life. 9 (4). 403-417 (2003)

Wawer, M.J. et al. Rates of HIV-1 Transmission per Coital Act, by Stage of HIV-1 Infection, in Rakai, Uganda. Journal of Infectious Diseases. 191. 1403–1409 (2005)

Werth, B., Geller, A. and R. Meyer. He endorses me - he endorses me not - he endorses me ... contextualized reasoning in complex systems. AAAI Fall Symposium (2007) <http://cfpm.org/cpmrep184.html>

2.2 Universität Kassel

2.2.1 Introduction

The work of the Kassel modelling team over the period of the last six months has concentrated on the further development of the two versions of the SoNARe[6] agent-based model. The rationale for this separation of the modelling effort into two strands has been to be able to investigate more or less independently the abstract features and phenomena of the case study, i.e. those which are presumed to be found also in other problem settings, and the more case-study specific features and phenomena. The centre of interest of both strands has been to explore those consequences for the collective action necessary on the part of farmers to reactivate a neglected land reclamation system (LRS) that arise from the relation between the social dimension and the economic dimension of farmer decision making against the background of the structure of their social network and the inherent hydrological inter-dependencies caused by the spatial location of the land parcels the farmers manage.

SoNARe-A, the abstract version, takes a quantitative approach to modelling farmers'

[6] **So**cial **N**etworks of **A**gents' **Re**clamation of Land

perceptions and their consideration of social and economic factors that bear on their decisions. Amongst other things, this allows the investigation of sets of scenarios that show shifts from more economically driven farmer populations to more socially driven farmer populations and vice versa.

SoNARe-D on the other hand includes more data from the actual case study (e.g. an elicited network structure and an additional farmer type) and is thus more evidence-driven than the more abstract version. In addition, it attempts to more closely capture the symbolic/qualitative nature of agents' decision-making and differentiates a set of social influencing factors.

While the abstract model has undergone some detailed sensitivity analyses, the evidence-driven model — largely owing to its novelty and added complexity — has still to be thoroughly investigated.

2.2.2 The SoNARe-A Model

The goal of SoNARe-A has been to investigate how a collective effort of the farmers involved can be triggered to transform the neglected LRS found in the Odra region at present to a collectively managed working LRS.

Assumptions

Farmers have individual perceptions of their economic success over a past sequence of years. The economic success is determined by the farmers' profit from their farming activities. Farmers evaluate their profit with respect to a certain (fixed) threshold below which the profit of a year is considered "too low". Profit is composed of production costs, costs for LRS maintenance (if maintaining), attained yield from a farmer's field (which is influenced by climate conditions and the global effectiveness of the LRS), and compensation payments in case of crop losses.

Farmers are embedded in a social network; their opinion about LRS maintenance (pro/con) is propagated over the network links and perceived by others. The sum of all influences received over the social network links constitutes a farmer's perception of social support of his opinion concerning LRS maintenance.

We consider two decision scenarios: In the Selfish Scenario farmers decide about whether to contribute to the LRS maintenance solely based on their economic success, i.e. they disregard all social aspects. In the Social Scenario farmers balance their decision equally between individual economic success and perceived social support.

Simulation

All simulations start out from the status-quo observed in the Odra region (on the given level of abstraction, of course a bit "sketchy", etc.): The LRS is not functioning and none of the farmers is willing to start maintaining it. There are no economic incentives for maintaining the LRS because farmers who experience crop losses in wet years are fully compensated.

The first 10 years of the presented simulation runs are "warm-up years" that simulate the status-quo and initialise profits, profit memories and social support perceptions of the farmers. Starting from simulation year 11 an LRS initiator gets active. This actor (someone from an NGO, a mayor, etc.) observes the individual profits of the farmers and starts to promote LRS maintenance (by exerting social influence pro LRS) if a certain proportion of farmers (here 10%) have very low profits. The activity of the LRS initiator is only relevant in the Social Scenario because in the Selfish Scenario farmers disregard all social influences.

In addition, also starting from year 11, one of five compensation policies takes effect:

Compensation Policy	
0	pay no compensation
1	pay compensation always to all farmers
2	pay compensation always to farmers who maintain LRS
3	pay compensation to all farmers, but only when LRS initiator promotes maintenance
4	pay compensation only to farmers who maintain LRS, but only when LRS initiator promotes maintenance

Unlike during the initial 10 (status-quo) simulation years when farmers are always fully compensated the compensation policies assume an upper limit of compensation that is paid to an individual farmer – in fact farmers are not always fully compensated. Under full compensation and with a neglected LRS the maximum amount of compensation paid to an individual farmer reaches from 7 units in a wet year to 2.5 units for the most downstream farmers in a normal year. For the policy scenarios we set the maximum compensation to a value of 2.5. All simulation runs use the same maximum value.

Simulation results

Figures 1 and 2 on the following pages show sets of diagrams of the simulated dynamics for different combinations of decision scenarios and compensation policies. Fig. 1 illustrates the effects of the 5 different compensation policies on the Selfish Scenario, Fig. 2 shows the Socially Active Scenario. In the diagrams, dashed vertical lines mark the end of the warm-up years, the displayed mean values cover only years 11 to 120, i.e. the years when the respective compensation policies become active. Strategy adjustments summarise opinion shifts in either direction. Relative changes are the differences of these percentages relative to the previous year (first differences). Compensation payments and total farmer profits are shown as rolling means over 3 years.

The following table compares the calculated mean values as indicators for the different scenario combinations:

	Selfish Scenario			Social Scenario		
Compensation Policy	Total Compensation per year (avg. / final)	Total Profit (avg. / final)	Avg. Mobilisation in % (avg. / final)	Total Compensation per year (avg. / final)	Total Profit per year (avg. / final)	Avg. Mobilisation in % (avg. / final)
0	0 / 0	94 / 99	68 / 78	0 / 0	101 / 109	81 / 100
1	46 / 48	107 / 103	35 / 38	41 / 42	121 / 119	2 / 2
2	20 / 28	105 / 112	64 / 74	22 / 28	103 / 109	66 / 81
3	17 / 0	97 / 106	64 / 93	21 / 0	101 /109	47 / 100
4	6 / 10	102 / 111	70 / 82	2 / 0	102 /109	79 / 100

The next table shows the relative reductions of compensation payments and profits when comparing to the warm-up years (full compensation)

	Selfish Scenario	Social Scenario

Comp. Policy	Reduction of Compensation	Reduction of Profit	Reduction of Compensation	Reduction of Profit
0	100%	36%	100%	31%
1	16%	27%	24%	17%
2	63%	29%	60%	30%
3	69%	34%	61%	31%
4	89%	31%	96%	31%

Interpretation

Compensation Policy 0:

- Cutting back compensation payments to 0 introduces a phase of economic stress (decrease in profits). This breaks up the initially passive behaviour of the farmers and triggers a phase of volatility in the opinion dynamics.
- In the Selfish Scenario a majority of around 80% gets mobilised to maintain the LRS. For the remaining portion of farmers the (then reduced) economic stress does not offer sufficient incentives to maintain the LRS.
- In the Social Scenario (with timed social activity of the LRS initiator) the economic stress induces a mobilisation of 100% of the farmers, whereas during the warm-up years (full compensation of all crop losses) the social influence of the LRS initiator does not suffice to mobilise farmers.

Compensation Policy 1

- In the Selfish Scenario a minority of the farmers (around 40%, located upstream) starts to maintain the LRS because the reduced maximum amount of compensation does not equalise their crop losses. The remaining farmers may well live with the compensation paid and the benefits from the partly installed LRS.
- In the Social Scenario the permanent flow of compensation prevents the LRS initiator from breaking up the social coherence between the farmers. A small proportion of (upstream) farmers have high crop losses in wet years which makes them start to maintain but in subsequent normal years social pressure makes them shift back to passive behaviour (see relative changes). In total almost none of the farmers maintain; profit is mainly generated by compensation payments.
- In the Selfish Scenario a higher total of compensation is paid because those farmers who invest in the LRS have a higher chance of getting compensated. Here, compensation is partly a subsidisation of LRS maintenance.

Compensation Policy 2:

- In both decision scenarios the selective input of compensation payments to farmers maintaining the LRS reduces the total payments required and provides sufficient economic incentives to mobilise a majority of around 80% of the farmers.

Compensation Policies 3 and 4:

- For both decision scenarios the “pulsed” input of compensations further reduces the total payments.
- When pulsed compensation is paid to all farmers (policy 3) in both decision scenarios nearly 100% of the farmers are mobilised. In addition, compensation payments decrease over time and fall to zero at the end of the simulation.
- When pulsed compensation is paid selectively to LRS maintainers (policy 4) in both

decision scenarios compensation inputs are further reduced while in the Selfish Scenario profits increase and in the Social Scenario remain constant.

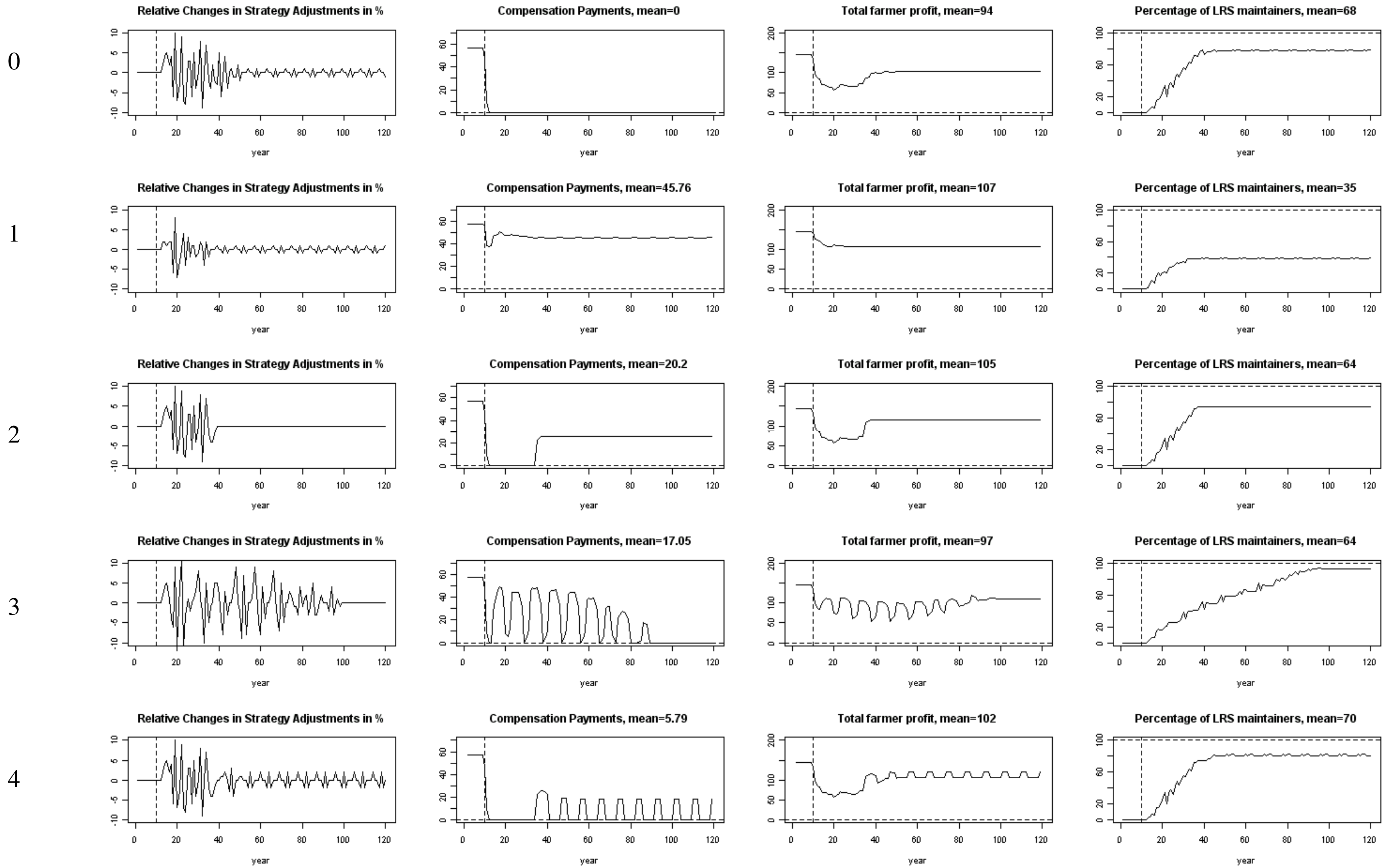


Figure 1: Simulated dynamics for Selfish Scenario, compensation policy 0 to 4 per row. Strategy adjustments summarise opinion shifts in either direction. Relative changes are the differences of these percentages relative to the previous year (first differences).

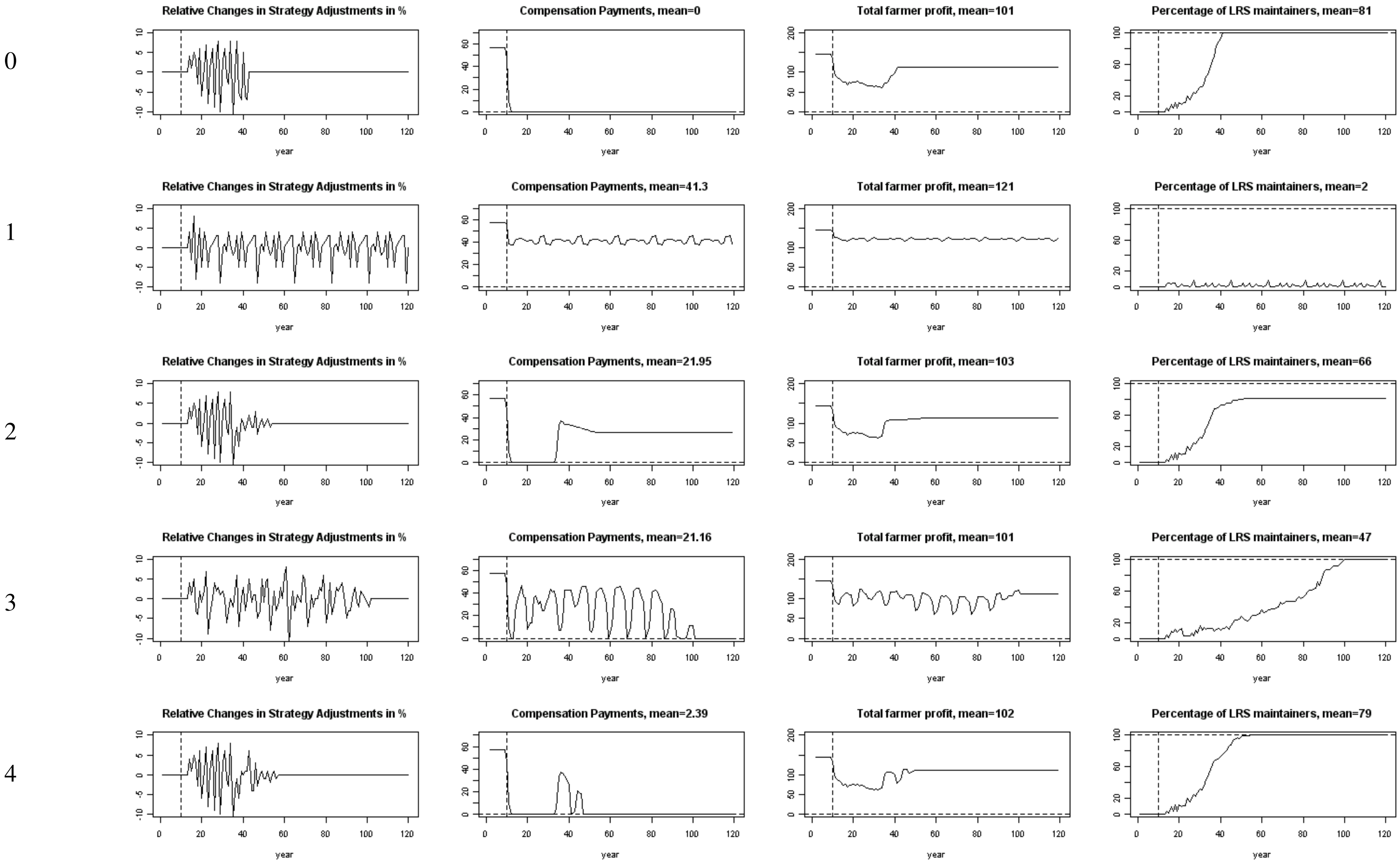


Figure 2: Simulated dynamics for Social Scenario, compensation policy 0 to 4 per row. Strategy adjustments summarise opinion shifts in either direction. Relative changes are the differences of these percentages relative to the previous year (first differences).

2.2.3 The SoNARe-D model

The overall goal for the development of the SoNARe-D model of the Odra case study has been to incrementally and systematically integrate as many as possible of the features and drivers — as extracted from the collected data of the field research — that bear on the social dilemma assumed to be inherent in the establishment and maintenance of a functioning land reclamation system (LRS) of irrigation and drainage ditches and canals (and as described in detail in earlier reports). To this end, the model comprises three interconnected submodels capturing biophysical, economic and social aspects of the investigated problem area. The biophysical and economic aspects are covered by the so-called Simple Hydro-Agricultural Model (SHAM) and the Simple Agro-Economics Model (SAEM) were developed by Grzegorz Holdys at Wroclaw University of Technology.

In the following, recent improvements to the SoNARe-D model, i.e. the agent-based social model, developed by the Kassel group, will be characterised. Finally, an example scenario run will be described.

In the main, the social model was enriched in two ways: First, the structure of the social network (the acquaintance network) was modelled on the basis of a set of case study data regarding the neighbourhood and family relations of the farmers present in the case study area. Second, in order to better reflect the qualitative nature of the data concerning the decision criteria and decision making of the relevant actors, the agents were equipped with a quasi-qualitative decision mechanism based on endorsements.

The social environment of the agents

A social network was extracted from the interview data provided by the Odra case study team. Whilst the data comprises a total of 116 farmers, it contains neighbourhood and family network information for only 74 farmers. In this context neighbourhood refers to the neighbourhood of where the farmers live rather than to the neighbourhood of the land parcels they manage. The data thus obtained has been modified to account for and ensure the transitivity of family relations, i.e. it is assumed that if a_1 is related to a_2 and a_2 is related to a_3, then a_1 and a_3 must also be related. Figure 3 (on the left) shows the resulting combined neighbourhood and family network, i.e. a kind of acquaintances network.

In order to be able to scale up this network to more than the 74 nodes without jeopardising its general signature, a first primitive algorithm was devised.

The algorithm first chooses at random one node from the set of existing nodes to act as the template for the next new node to be generated. This template node is then copied including all of its edges, i.e. the newly generated node shares the template node's network neighbourhood. However, it is possible to set a random rewiring parameter which determines the percentage of these edges which will be randomly rewired. (Both random processes use the same uniform distribution.) Figure 3 (on the right) shows a network of 100 nodes generated by the algorithm described without rewiring. The network measures calculated for the two networks depicted in figure 3 are given to indicate that the general signature is largely conserved.

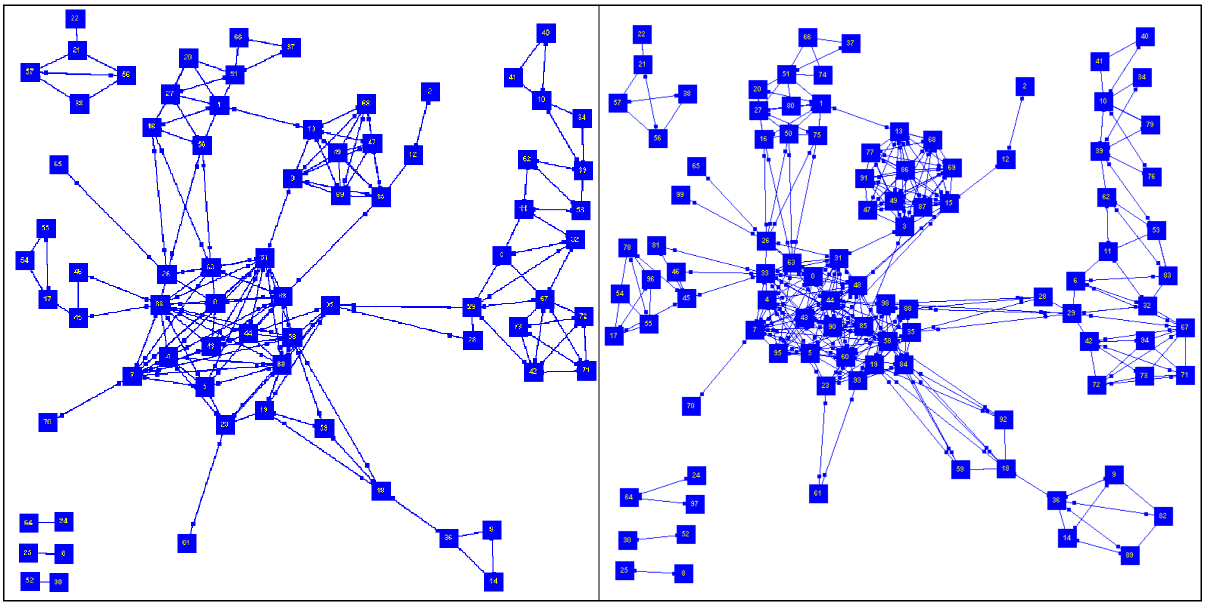


Figure 3: On the left: An acquaintance network obtained from Odra case study data (N=74). On the right: The acquaintances network scaled up to 100 nodes (without random rewiring).

```
N = 74
Clustering Coefficient                = 0.5715294840294839
Average Path Length                   = Infinity
Density                               = 0.05715850986121256
Diameter                              = Infinity
Average In Degree                     = 4.22972972972973
Average Out Degree                    = 4.22972972972973
Network Degree Centrality             = 0.11853881278538812
Variance Network Degree Centrality    = 6.822580801911454E-31

N = 100 (random rewiring proportion = 0.0)
Clustering Coefficient                = 0.538132356850236
Average Path Length                   = Infinity
Density                               = 0.0631
Diameter                              = Infinity
Average In Degree                     = 6.31
Average Out Degree                    = 6.31
Network Degree Centrality             = 0.1569156915691569
Variance Network Degree Centrality    = 8.863641131063289E-29
```

It should be remarked, however, that the algorithm does not guarantee the conservation of network signatures in all cases. This particularly holds true when the random rewiring proportion is sufficiently great. Thus, it is, at present, imperative to carefully select a desired network from the algorithms output.

Agent types and their decision-making

SoNARe-D models three basic types of agents in accordance with three main actor types identified in the case study region. The water partnership initiator agent (WPI) plays the role of an opinion leader (NGO, local mayor etc.) who attempts to convince farmers to maintain their sections of the LRS. It is connected to all farmers via the social network. The WPI gets active once a certain number of farmers have big losses. The second type of agent is the big farmer (BF). It owns several land parcels (in the present version each BF owns ten parcels) and always maintains the LRS of these parcels. Finally and most importantly, small farmers (SF) — also called part-time farmers (PTF) —each own a single land parcel and decide once a year whether to maintain that parcel's section of the LRS in the following year or not. They do so by way of an endorsement mechanism as defined by Moss (1995). Small farmers are further subcategorised as willing PTF (WPTF) and unwilling PTF (UPTF) depending on whether they are susceptible to the influence of the WPI or not.

As in previous model versions, a SF currently has only one decision to take, namely whether or not to maintain its section of the LRS in the current year. In this version, however, it does so by annually endorsing its two options, i.e. it collects "arguments" for the one and for the other option, and deciding on this basis if it maintains or not. Similar to previous model versions and as indicated by the data collected from interviews with actors in the case study, these influencing factors are either economic or social in nature. The latter are, again, modelled as and derived from endorsements, namely how the agent has endorsed its neighbours in the social network and the maintenance strategy they currently employ.

Each agent possesses two separate memories for storing endorsements, one for its maintenance decision related endorsements (*mem*) and another for its agent related endorsements (*aem*).
First of all, each year each farmer agent (henceforth called "the endorser") endorses each of its neighbours in the social network (henceforth called "the endorsee") on two dimensions:

4. If the endorsee uses the same strategy as the endorser, a "*sameStrategy*" endorsement is stored in the endorser's *aem*, otherwise a "*differentStrategy*" endorsement is stored;
5. Profit-related endorsements are stored as follows:

   if $lastProfit_{endorsee}$ > (1 + sgn($lastProfit_{endorser}$) * moreProfitThreshold) * $lastProfit_{endorser}$
   then store "moreProfit" in aem
   else if $lastProfit_{endorsee}$ < (1 + sgn($lastProfit_{endorser}$) * lessProfitThreshold) * $lastProfit_{endorser}$
   then store "lessProfit"
   else store "similarProfit" in aem

   with *moreProfitThreshold* ϵ [0;1] and *lessProfitThreshold* ϵ [-1;0].

These endorsements serve as the basis for the endorser to rank its neighbours with respect to how similar to itself it considers them to be and which neighbours it deems worthy of being imitated because of their economic success. To this end also endorsement schemes are defined and assigned to each SF to express each agent's individual preferences, i.e. how it "defines" similarity and success (or rather "worthy of imitation due to success"), respectively. Currently, all SF are assigned the same similarity and success endorsement schemes:

similar	
Token	***Value*** *(ev)*
similarProfit	2
sameStrategy	1
	0
differentStrategy	-1
moreProfit, lessProfit	-2

better off	
Token	***Value*** *(ev)*
moreProfit	2
	1
	0
similarProfit	-1
lessProfit	-2

We use the following overall endorsement function to evaluate the overall endorsement value *(oev)* of each endorsee with respect to similarity and success given the endorsements stored in *aem*:

$$oev = \sum_{ev} (\text{sgn}(ev) \cdot base^{|ev|})$$

Having endorsed its network neighbours, the farmer agent then endorses its options with respect to LRS maintenance. Currently, maintenance is endorsed on four dimensions by each agent. Whether all or just some of them are actually taken into account when evaluating the two options and making a decision depends on the maintenance endorsement scheme assigned to the individual agent (see below).

The four dimensions and their respective endorsements are:

1. Economic success is endorsed depending on the agent's "interpretation" of when a profit is to be considered a "big loss". In the present version of the model farmer agents consider their attained profit, i.e. the profit from that year's harvest, to be a big loss, if it is negative or zero or, in case it is positive, if the following statement holds true:

   *lastProfit < lossToleranceFraction * meanProfitCurrentStrategy*

   with *lossToleranceFraction* ϵ [0;1] and *meanProfitCurrentStrategy* being the mean of the memorised profits that were attained in those past years in which the current strategy was also used or zero, if the agent does not remember any such years. If any of the above statements apply, a "bigLossesWhenMaintaining" endorsement or a "bigLossesWhen-NotMaintaining" endorsement is stored in *mem*.

2. If the endorsee which is considered the most similar to the endorser currently maintains its LRS, i.e. the farmer agent which is endorsed best according to the endorser's similarity scheme and corresponding overall endorsement function , a "*mostSimilarNeighbour Maintains*" endorsement is stored in *mem*, if it does not maintain a "*mostSimilarNeighbour DoesNotMaintain*" endorsement is stored.

3. Analogous to 2., a "*bestOffNeighbourMaintains*"/"*bestOffNeighbourDoesNotMaintain*" endorsement is stored in *mem*, if the endorsee which is considered to be best off of all neighbours maintains/does not maintain. However, this endorsement is only issued, if the endorsee is better off than the endorser itself.

4. If at least half of the agent's network neighbours maintain, a "*mostNeighboursMaintain*" endorsement is stored in *mem*, otherwise a "*mostNeighboursDoNotMaintain*" endorsement is stored.

Again, to reflect agents' preferences in weighting the different endorsements — this time when it comes to deciding for or against maintenance — one endorsement scheme of the set of schemes defined below is assigned to each SF and the overall endorsement function defined above is applied (to the endorsement memory *mem*). If the result is positive, then the agent will maintain its LRS section(s) in the next year; it does not maintain otherwise.

(1) **economic**: *bigLossesWhenMaintaining/NotMaintaining* = 1/-1; all others set to 0.
(2) **most +eco**: same as economic, but *mostNeighboursMaintain/DoNotMaintain* = 1/-1
(3) **most similar +eco**: same as economic, but *mostSimilarNeighbourMaintains/DoesNotMaintain* = 1/-1
(4) **best-off +eco***:* same as economic, but *bestOffNeighbourMaintains/DoesNotMaintain* = 1/-1
(5) **most and best-off +eco***:* combines most and best-off, but *bigLossesWhenMaintaining/NotMaintaining* = 2/-2

(6) **most and most similar +eco***:* combines most and most similar, but *bigLossesWhenMaintaining/NotMaintaining* = 2/-2
(7) **most, most similar and best-off +eco***:* combines most, most similar and best-off, but *bigLossesWhenMaintaining/NotMaintaining* = 3/-3

The distinction between WPTF and UPTF, i.e. the willingness to maintain and the lack thereof, is realised by adding the endorsement token "*WPIExertsInfluence"* with an endorsement value of 1 to the maintenance endorsement scheme of WPTFs, but not to that of UPTFs. A "*WPIExertsInfluence*" endorsement is added to an agent's *mem* each year that the WPI is active. The WPI is active as long as the mean number of agents with big losses over the past 6 years is greater than *WPIActivationThreshold* (currently set to 33).

Simulation results

The model run described in the following was performed for each of the seven maintenance endorsement schemes defined in the previous section, i.e. seven scenarios are compared and within a scenario all SF use the same scheme. 200 parcels in all (see figure 4), i.e. 20 independent channels with 10 parcels each, were assigned to a total of 110 farmers: 100 SF (75 UPTF and 25 WPTF) each managing one land parcel, and 10 BF each managing 10 land parcels. The 10 land parcels of a BF are situated in a row along one or two channels[7]. The model parameter *distRngSeed* (the seed for the random distribution) determines the way in which parcels are assigned to farmers. In order to show a characteristic run for each scenario, the *distRngSeed* was set to 1 for all scenarios except for scenario 4 for which it was set to 7. Initial test runs indicated that this *distRngSeed* is not sensitive for these scenarios except for scenario 4. However, the general trend in this scenario remains the same, namely that 100% maintenance are eventually achieved and maintained.

Small farmers endorse one another in their social network (BF are not included). Its topology is that shown on the right in figure 3. In addition, they are also connected to those SF that own a parcel along the same channel as theirs. The WPI is connected to all SF. It observes all SF and gets active when on average over the last six years 33 or more of them had big losses; it is inactive otherwise.

The endorsement/decision relevant parameters for the SF were set as follows:

- *moreProfitThreshold* = 0.2
- *lessProfitThreshold* = -0.2
- *lossToleranceFraction* = 0.9
- retention time of endorsements in *aem* = 10 years
- retention time of endorsements in *mem* = 10 years
- capacity of the profit memory = 10 memory items (= 10 years)

The model was run for 114 years, twice repeating a historical water level sequence of 57 years. The data represent an average monthly water balance measured in Wroclaw in the period between 1947 and 2003. It was provided as part of SHAM by the project member Grzegorz Holdys, Wroclaw University of Technology.

The economic model SAEM was configured with the following default settings:

- *unitYieldBasePrice* = 61
- *unitProductionCost* = 1704
- *channelMaintenanceCost* = 10

[7] Note again that, if the land parcels are spread over two channels, the two rows of parcels – and thus the channel sections – are not connected.

- *notChannelMaintainedFine* = 20
- *parcelArea* = 2

SAEM calculates the profits for individual parcels by the function:

*profit = (yield * unitYieldBasePrice) – (parcelArea*unitProductionCost) – channelMaintenanceCost – notChannelMaintainedFine*

with *channelMaintenanceCost* only applying in the case of maintainers and *notChannelMaintainedFine* applying only in case of non-maintainers. The maximum *yield* that can be obtained is 70.

At the start of the simulation no SF maintain their LRS.

Figure 5 shows some results for all SF. Strategy adjustments summarise opinion shifts in either direction. Relative changes, an indicator for volatility, are the differences of these percentages relative to the previous year (first differences). Profits are shown as rolling means over 5 years for maintainers, in blue, and non-maintainers, in red.

The results show that volatility is far higher in the purely economic/egocentric scenario than when social considerations/aspirations play a role. This can be seen as an indication for a social lock-in or coherence effect in the more social scenarios (2-7), as had already been identified in early versions of the abstract model. This finding is mirrored in the percentage of LRS maintainers which stabilises much more in those scenarios than in the economic case. However, if agents only compare themselves with the neighbour they consider to be most similar to themselves — with regard to profits and maintenance strategy application — it stabilises on a much lower level than in all other cases (around 40%).

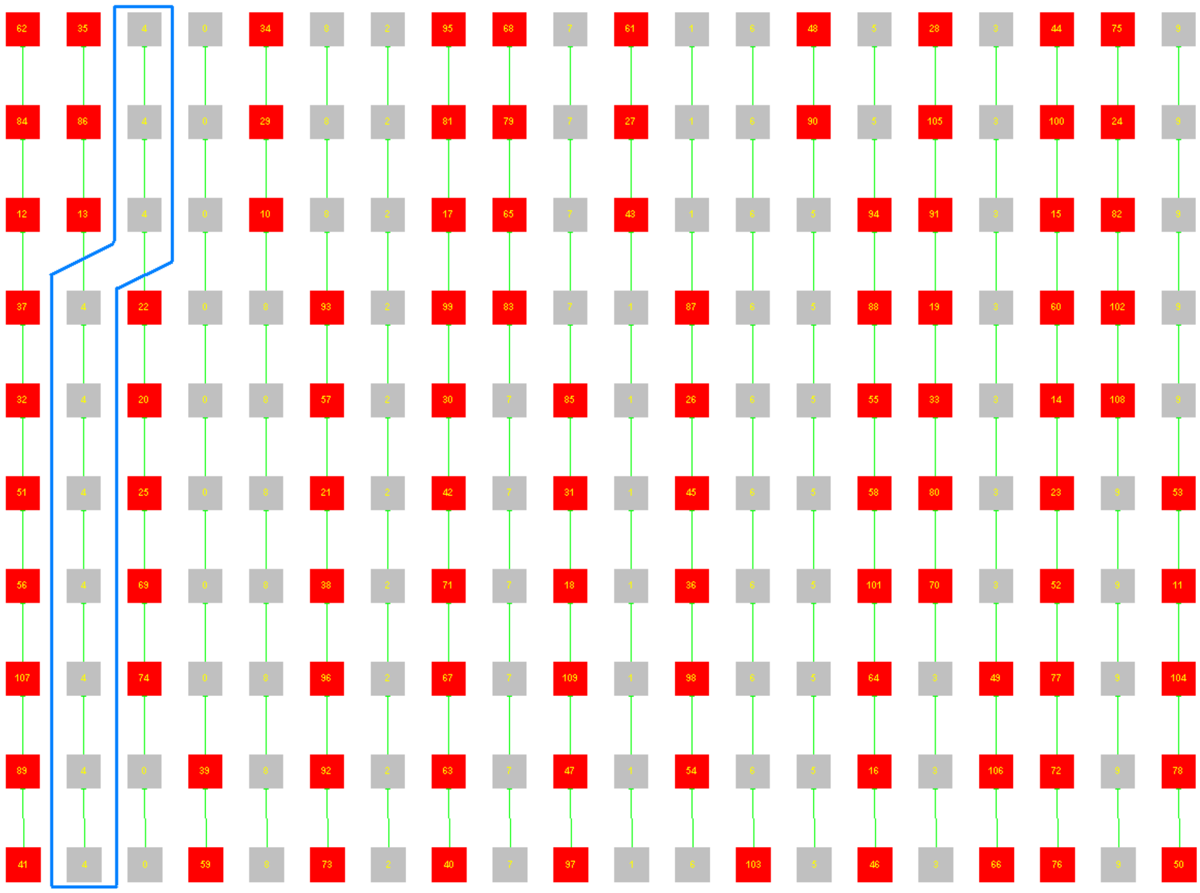


Figure 4: 20 simulated channel (left to right) with 10 land parcels each (top to bottom). Red squares represent land parcels managed by small farmers. Grey parcels represent land parcels managed by big farmers. As an example, the blue box indicates the ten parcels of a single big farmer situated along two (independent) channels.

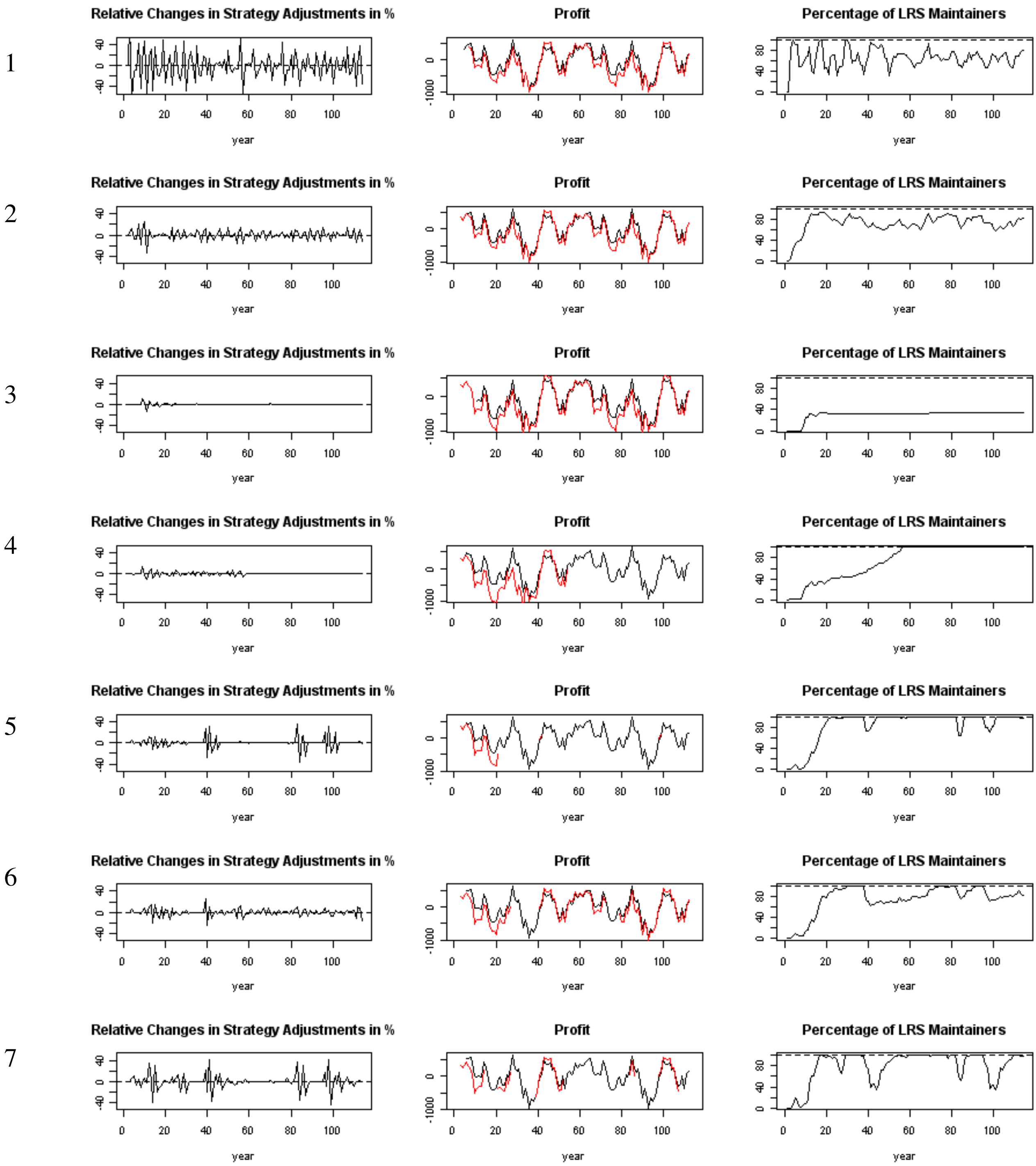


Figure 5: Simulation results for seven maintenance endorsement scenarios over 114 simulated years. Endorsement schemes used: 1=economic, 2= best off +eco, 3=most similar +eco, 4=most +eco, 5= most and best off, 6= most similar and best off, 7=most, most similar and best off +eco.

2.2.4 Publications

Krebs, F., Elbers, M. and Ernst, A. (2008). Modelling the social and economic dimensions of farmer decision making under conditions of water stress. Proceedings of the 1st ICC workshop on complexity in social systems, Lissabon.

2.2.5 References

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

2.3 Macaulay Institute

The modelling report consists of three main parts: a description of the modifications to FEARLUS since September 2007, an account of the progress made toward construction, quantiative calibration and validation of a fitted FEARLUS model of the Upper Deeside area of the Grampian Region from 1980-2005, and an outline of future work, which will be funded by the Scottish Government.

2.3.1 Modifications to FEARLUS

The following is a summary of the modifications made to FEARLUS since September 2007, creating the final version 1.1.4:

- Enabled compilation on Linux (Fedora) and Apple (Leopard) platforms.
- Created numbers of Government policy classes to implement various ways in which a government might try to influence land use choice.
- Added functionality to enable Managers to sell up and quit farming with a small probability (rather than just quitting through being bankrupt).
- Added truncated-normal distribution for initialising Land Managers from their Subpopulation parameters.
- Added time series of Subpopulation in-migrant selection probabilities.
- Added facility for Land Parcels to be compared using their Biophysical Characteristics when determining similarity of cases, under some restricted conditions.
- Added facility for Land Managers to wait for a number of Years of failing to achieve Aspirations before reviewing Land Use on the Farm.
- Created a random number generator class to seek data from a Quantis true random number generator device.
- Added reports to generate output conforming to the CAVES generalisation framework document sent by the Polish team.
- Prepared final version of user guide.

The following provides more detail on the above.

Compilation on Linux and Apple platforms

The FEARLUS 1-1-4 software can now be compiled on most Linux platforms with the Swarm software installed. On Fedora Core 6, Paul Johnson's binary RPMs were found to work successfully once all prerequisite software was installed. On SUSE Linux and Redhat Enterprise Linux, both x86_64, compilation was successful using CVS download versions of Swarm. On an Apple running Leopard (OS X 10.5), with some assistance from the Swarm Support mailing list, a CVS version of

Swarm was compilable that enabled FEARLUS to run. A binary distribution of Swarm for Leopard is apparently under development.

Government classes

Various Government classes have been added to FEARLUS to enable the model to explore the consequences of different grant policies on land use change. These classes are described as follows:

- RewardActivityGovernment. Reward for specified activities by the Land Manager (i.e. Land Uses).
- ClusterActivityGovernment. Reward for specified activities by the Land Manager, with extra given to the Manager if their neighbours also have the same activity.
- TargetActivityGovernment. Reward for specified activities when their coverage (i.e. proportion of use in the Environment) is less than a target coverage.
- TargetClusterActivityGovernment. Reward for specified activities with extra for neighbours using the same Land Use when their coverage is less than a target coverage. If the target is exceeded, only reward those managers rewarded in the previous Year.

The above classes all issue rewards, and are available in four forms: issue a fixed specified reward for particular Land Uses; issue rewards as a share of a fixed budget; issue fixed rewards but with a cap on what any one Land Manager can receive; issue rewards as a share of a fixed budget but with a cap on what any one Land Manager can receive.

There are four further new classes:

- SubsetActivityGovernment. Reward a fixed amount for specified Land Uses, but put a limit on the number of awards made, and allocate them randomly.
- SortSubsetActivityGovernment. Reward a fixed amount for specified Land Uses, but put a limit on the number of awards made, and allocate them to Managers in descending order of the number of awards they are eligible for.
- NbrSubsetActivityGovernment. Reward a fixed amount for specified Land Uses where they occur with at least n (a parameter) neighbours using the same Land Use. Put a limit on the number of awards made, and allocate them randomly.
- SortNbrSubsetActivityGovernment. Reward a fixed amount for specified Land Uses where they occur with at least n neighbours using the same Land Use, with a limited number of awards made to Managers in descending order of the number of awards they are eligible for.

Sell up probability

Before this modification, the only way for Land Managers to leave the simulation was bankruptcy. This modification introduced the facility for a small probability to be specified that the Managers would ‘sell up’ in any given Year: leave the simulation even though they are not bankrupt. The probability for an individual agent is specified from a distribution at Subpopulation level.

Truncated normal distributions in Subpopulations

The facility was added to FEARLUS for certain Land Manager parameters to be set from their Subpopulation using a ‘truncated normal’ distribution, in addition to the pre-existing uniform and normal distribution options. For truncated normal distributions, a sample is taken from a normal distribution with the specified mean and variance, but if the sample is less than a specified minimum, it is set to the minimum, and if greater than a specified maximum, it is set to the maximum. Since this effectively results in a probability density function with infinite spikes at the minimum and maximum, such a distribution could be problematic analytically. A better approach might have been to use a different distribution, such as one with a specific lower bound, where this would achieve the desired effect, and this will be considered for future versions.

Subpopulation in-migrant selection probability time series

Modifications have been made to allow the probabilities that in-migrant Land Managers will belong to particular Subpopulations to change over time. This is provided simply through specifying the relevant probabilities over time in a file (that is, the in-migrant demographic change is exogenous to the model). This allows us to reflect changes in real-world land manager populations when modelling scenarios. For example, many Scottish farms are being bought by Irish farmers, and other farms are bought by the wealthy and used for recreation or keeping horses rather than growing food (at least on a commercial scale).

Case similarity for Land Parcels based on Biophysical Characteristics

Prior to version 1-1-4, similarity of Land Parcels when evaluating cases was based on Euclidean distance. This was done as a simple solution to the issue that Biophysical Characteristics are specified at the sub-Parcel (Cell) level, meaning that Parcels could potentially have areas of different Biophysical Characteristics. It was felt that the complexity of comparing Parcels' Biophysical Characteristics in detail was probably unrealistic, and in the absence of an alternative approach, Euclidean distance was used under the assumption that closer Parcels are more similar than distant Parcels (not unreasonable given that this is a basic geographical assumption). However, in some simulations, it is possible to stipulate that FEARLUS allocate only one Cell per Parcel, and in others, it is possible that there are Parcels with homogeneous Biophysical Characteristics across their Cells. In such cases, an option has been provided for them to be compared using the Biophysical Characteristics property value lists rather than Euclidean distance.

Delays to changing land use

Evidence from the case study suggested that managers did not immediately switch commodity when they were dissatisfied with the returns on their current activities, as was the case in FEARLUS. Instead, managers would 'ride the curve', waiting for things to improve in the marketplace, often under the impression that the year they changed commodity would be the year that things got better. To model this, we introduced parameters allowing individual Land Managers to have a set number of consecutive Years of not achieving Aspirations before reviewing Land Use on the Farm, Land Use otherwise not changing.

Quantis true RNG compatibility

Issues with pseudo random number generators in spatial modelling are highlighted by van Niel and Laffan (2003), and an inexpensive true random number generator device has been acquired by the Grampian modelling team to provide a check that results are not biased by any of the problems that pseudo random number generator (pRNG) algorithms are known to have. Van Niel and Laffan (p. 52) cite Hellekalek (1998) on this issue, stating that period length, correlation and structure are the main issues with pRNGs. Pseudo-RNGs generate numbers that after a length of time, repeat in cycles. The length of one cycle thus limits the number of random numbers that can be meaningfully obtained from a pRNG. The period length of the MT19937 generator (the default in the Swarm simulation libraries) is 219937 – 1; an extremely long cycle that wouldn't be an issue in simulation models, but some pRNGs, and in particular those available by default in many programming languages, have much shorter periods; and rules of thumb for sampling from pRNGs do not recommend using all numbers in a cycle (the square root or even cube root of the cycle length has been recommended, according to van Niel and Laffan). Even within a cycle, random number sequences may be correlated to some degree or another. The ‘structure’ issue manifests itself in more problem-specific terms: van Niel and Laffan cite literature on using pRNGs to generate lattice structures, stating that for some classes of pRNG the results do not evenly occupy the n-dimensional space.

The Quantis ‘true’ (‘hardware’ might be a more accurate term) random number generator works by firing photons at a semitransparent mirror, and using two detectors to see if the photon went through the mirror, or was reflected by it, is able to generate a series of bits. Having obtained the device, we wrote a class for inclusion in the FEARLUS model enabling, on compatible machines, random numbers to be sourced from this device rather than one of the Swarm library’s pRNGs.

Generalisation framework reports

The document circulated by Paulina Hetman and Piotr Magnuszewski was used as a basis for implementing a number of new reports in FEARLUS, enabling output to be gathered that hopefully can be compared with other CAVES models. There are three new report classes, generating data for each of the three main categories of information in the generalisation framework document: networks, space and time.

- NetworkStatisticsReport. Various statistics can be computed for each of the Neighbours, Approvers, Disapprovers, Approved, Disapproved, Advisors and Advisees networks. Where the network is not fully connected (i.e. where there exist Land Managers between which there is no path), the statistics are computed for fully connected subsets of Land Managers. The statistics include the diameter, mean path length, assortativity, density, and clustering coefficient. For undirected networks, the closeness centrality and centralisation index are also computed.

  Example output (undirected):

  ```
  Neighbours network 1
  Number of edges        150     Number of nodes        93
  Diameter:        inf        Largest subgraph diameter:     15
  Mean path length:      inf     Unconnected value:     inf
  Assortativity:    0.349631        Density:        0.0350631
  Clustering coefficient:  0.175182       Local clustering coefficient:      0.288588
  Closeness centrality:    0.0038702      Closeness variance:     1.21725e-06
  Centralisation index:    16.9457        Degree variance:        870.796
  ```

  Example output (directed):

  ```
  Advisors network 6
  Number of edges        7       Number of nodes        6
  Diameter:        inf        Largest subgraph diameter:     2
  Mean path length:      inf     Unconnected value:     inf
  Assortativity:    0.259259        Density:        0.309524
  Clustering coefficient:  0.375          Local clustering coefficient:      0.277778
  ```

- SpatialAutocorrelationReport. The autocorrelation index and p-value are computed (the latter using non-free sampling of a specified number of samples) for the requested spatial data, which may be one of Price (the most recent Price of Land Parcels), Land Use, number of times Parcels have changed Land Use, number of times Parcels have changed Land Manager, Yield, or Income.

Example output:

Price space	Autocorrelation index	0.413367	p-value (non-free sampling)	0	
LandUse space	Autocorrelation index	0.221383	p-value (non-free sampling)	0	
OwnerChange space	Autocorrelation index	0.807106	p-value (non-free sampling)	0	
LandUseChange space	Autocorrelation index	0.56181	p-value (non-free sampling)	0	
Yield space	Autocorrelation index	0.331911	p-value (non-free sampling)	0	
Income space	Autocorrelation index	0.339612	p-value (non-free sampling)	0	

- TimeSeriesReport. The kurtosis, tail exponent, and autocorrelation function are computable for 24 time series, covering all aspects of the model from Approval to Yield. An option is available to configure the statistics to be computed on the relative change of the time series, rather than the raw time series itself. More details are available in the user guide.

  Example output:

Time series LandUseChange Kurtosis -1.04821
Tail Exponent nan ACF k= 1 to k= 100 :

0.966154	0.932257	0.898688	0.865195	0.831944	0.798432
0.765202	0.732372	0.700004	0.668229	0.636991	0.606676
0.57718	0.548663	0.521119	0.49423	0.468023	0.442565
0.418122	0.394137	0.370651	0.34768	0.32524	0.303318
0.282233	0.261838	0.241141	0.220138	0.198772	0.176983
0.154889	0.132455	0.109697	0.0868488	0.0639019	0.0407982
0.0177737	-0.00521206	-0.0271158	-0.0479703	-0.0686663	-0.0891472
-0.109601	-0.129981	-0.15038	-0.170573	-0.189571	-0.207449
-0.22425	-0.240075	-0.254953	-0.268869	-0.28174	-0.293424
-0.303833	-0.31282	-0.321306	-0.329496	-0.337304	-0.344762
-0.351639	-0.357956	-0.362805	-0.366109	-0.368874	-0.371091
-0.373015	-0.374597	-0.375821	-0.376917	-0.377916	-0.378784
-0.379677	-0.380574	-0.381455	-0.382344	-0.382238	-0.380969
-0.378812	-0.375731	-0.371731	-0.366838	-0.361071	-0.353809
-0.345263	-0.335444	-0.324338	-0.311549	-0.297021	-0.280852
-0.262858	-0.243354	-0.222195	-0.199453	-0.175144	-0.149364
-0.122598	-0.0942864	-0.0645357	-0.0330143		

Final version of user guide

A user guide was prepared to accompany the release of FEARLUS 1-1-4, and passed to the Kassel team.

Initial runs with FEARLUS-1.1.4

Test runs have been made with FEARLUS-1.1.4. No errors in the code have become apparent so far. In particular, the generalisation framework reports appear to be produced satisfactorily, although it has been noted that in the case of the Neighbours network, the fact that Land Managers may disappear from the simulation means that the network may consist of multiple components at the point in the annual cycle where the network is recorded.

2.3.2 Progress toward a fitted FEARLUS model of the Upper Deeside area of the Grampian Region from 1980-2005

Available quantitative information on the changes in land use in Grampian region (and information on farm size change so far as this can be obtained) will be used for input calibration and macro-validation of the model as applied to the past. 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 quantitative data available on both inputs (climatic, land suitability and price data, and data on the effects of climate and land suitability on yields) and outputs (land use and farm size data) is limited, and in some case affected by confidentiality restrictions. Therefore we will be assessing whether the model is able to reproduce the direction and magnitudes of the trends found in the data concerning land use and farm size, given the best available data relevant to model inputs.

The general approach being taken is as follows (these steps are overlapping rather than sequential; all but the last two are underway):

- Select those aspects of the world that can be represented in some way by inputs or outputs of the model. Some of these aspects (e.g. farmer decision-making procedures, climatic and economic conditions, available land uses) are inputs to the model; others (land use distribution and farm size) are outputs.
- For each of these aspects, determine what data are available for the period from the mid-1980s to the present. Farmer decision-making is discussed below under the heading of stakeholder/data validation exercises; other aspects are discussed here.
- Where there are data relevant to input parameters, determine how it can best be encoded in those parameters.
- Where there are no data good enough to be worth using for a particular input parameter, select a range of plausible combinations of parameter values with which to run the model.
- Select an area similar to, but not too close to, the Upper Deeside case study area, within the Grampian region. Use this to calibrate those aspects of the model (primarily concerning land manager decision-making) which cannot be estimated from the data sources described below. (Ideally, data used for calibration and validation should be independent; selecting a functionally similar but non-adjacent area for calibration is the closest we can get to this ideal.) This calibration exercise should narrow the range of plausible combinations of parameter values, since any combinations that produce clear mismatches with the real-world output data can be dropped.
- For each remaining plausible combination, run the model on the case study area data, as many times as is feasible. For each important quantitative variable adequately documented in the data sources available (e.g., change over a specific span of time in the number of farms in the study area, or in the proportion of land devoted to different land uses, we will be testing whether the variable's real-world value is within the range of values produced by ensembles of model runs: any such real-world value more extreme than 95% of an ensemble of runs would suggest a possible flaw in the model, and merit detailed investigation.

The remainder of this section lists the main kinds of model input and output to be considered, and the main sources of quantitative data. While far from ideal, these do appear sufficient to undertake quantitative validation work. The process of interpreting and combining the potential sources of quantitative information listed below has been more complicated and arduous than expected.

Model inputs and outputs

- Input parameters specify:
  - Initial land uses.
  - Initial farm boundaries, size distribution.
  - Land capability.
  - Climatic conditions.
  - Expected yields under the relevant combinations of land quality and climatic condition.
  - Economic conditions, primarily prices for farm products and the inputs necessary to produce them.
  - Land use selection strategies.
  - Land purchase strategies.
  - Land managers' social interactions (approval or disapproval of neighbours, and the importance a land manager attaches to their neighbours' good opinion, relative to that assigned to financial considerations).
  - Off-farm income.
  - Economies of scale.
  - Government policies.
- Outputs that could be compared with empirical evidence (in all cases, both time series and spatial distributions need to be considered).
  - Land uses
  - Farm boundaries, farm size distribution
  - Land prices

Main Data Sources

Data on land uses and farm sizes

The most readily available and processed data on land use and farm size come from the EDINA transformation of the annual Agricultural Census data (http://edina.ac.uk/agcensus/). Land uses are given at various levels of specificity (e.g. numbers are given both for "total cattle", and categories as specific as "bulls for service"). For use with the current version of FEARLUS, broad categories are considered to be most suitable. Land use data is georeferenced using 10km, 5km and 2km squares, although created from parish-level data: the 2km square data will be used in quantitative validation. Data about farm size is actually expressed as the number of farms – and of farms of different types – within each grid square. More detailed data on farm size will be difficult to obtain: land sales are recorded in public document repositories, but not in a readily accessible way; and tenancy agreements do not appear to be so recorded.

Doubts have been raised about the quality of the EDINA data, in particular about the effect of the parish-to-grid transformation, but since much of our other data (see below) is also in gridded form, it is a considerable advantage to have land use and farm size data in this format. Additional information concerning farm size (specifically, changes in farm sizes over time) may be drawn from Scottish Agricultural College's annual Farm Management Handbook (Beaton, Catto and Kerr 2007), which makes use of Farm Accounts Scheme sample data.

Land capability data

The EDINA version of Agricultural Census information also provides (at a 1km2 scale) a division of land into a small number of categories which it describes as "land use", but which are actually a combination of information about land capability and availability for agricultural and other uses: the categories used for Scotland being "urban", "water", "upland", "woodland", "agricultural land (Scotland)" and "restricted agricultural land" (golf courses and similar areas). However, it is explained in Hotson (1988) that only "agricultural land" and "moorland" (which we assume to be the same as "upland") are regarded as suitable for any agricultural use, so in fact we need just three categories: "agricultural land", "upland" and "other".

Climate data

The UK Met Office provides free access to sufficient weather station data for FEARLUS purposes, back to 1961. Since FEARLUS works on an annual timescale, the annualised data are the most obviously relevant. This includes, at a 5km2 scale:

- Annual extreme temperature range (Highest daily maximum - lowest daily minimum) (°C)
- Heating degree days - S (15.5 - daily mean temperature), ignoring negatives, over the winter months.
- Growing degree days - S (daily mean temperature - 4), ignoring negatives, over the summer months.
- Growing season length - Bounded by daily mean temperature being >5 °C for >five consecutive days and <5 °C for >five consecutive days.
- Summer heat wave duration - Total days with maximum temperature >3 °C above the 1961-90 normal for >five consecutive days (May to October)
- Winter heat wave duration - Total days with maximum temperature >3 °C above the 1961-90 normal for >five consecutive days (November to April)
- Summer cold wave duration - Total days with minimum temperature >3 °C below the 1961-90 normal for >five consecutive days (May to October)
- Winter cold wave duration - Total days with minimum temperature >3 °C below the 1961-90 normal for >five consecutive days (November to April)
- Maximum number of consecutive dry days (days with less than or equal to 2 mm of rain) in a year
- Greatest five-day precipitation total in a year (mm)
- Rainfall intensity (annual rainfall total on rain days (days with >1 mm of rain) / number of rain days)

Since the annual data are drawn from a limited set of locations, interpolation will be necessary if the area of interest is not sufficiently near to a station. Which of these types of data are most relevant, and whether they provide sufficient information between them, depends on the particular land uses being modelled. So far as the upper Deeside study area is concerned, the most relevant are "growing season length", "summer heat wave duration", and "maximum number of consecutive dry days". All these affect the growth of grass: a long, warm, wet summer will produce the best growth, but a hot, dry summer will be particularly poor.

Data on yields

Data on yields, given land use, climatic conditions and land quality at the levels of detail appropriate to FEARLUS, have not been easy to find. The best general source appears to be Scottish Agricultural College's annual Farm Management Handbook (most recent edition: Beaton, Catto and Kerr 2007). This series should be particularly useful with regard to changes in expected yield over time. Yield figures are always accompanied by a specification of the physical conditions and management practices assumed, and often in the form of high, medium and low estimates; this should make it feasible to adjust figures to take into account local conditions.

Data on economic conditions

The most readily useable source of information on both input and output prices are the time series produced by the UK's Department for Food and Rural Affairs (DEFRA). These series also include information on land prices (which could be used for comparison with those produced by the model, and potentially as a guide to incorporating exogenous influences on land prices); and of labour and fixed capital: factors not currently represented in FEARLUS, but potentially useful in interpreting simulation results.

The Farm Management Handbook also contains price data on inputs, outputs and land, and these can be used as a check on the DEFRA prices. However, the FMH, being designed to provide guidelines in preparing forward budgets, concentrates on figures for gross margins, based on forecasts for the coming season/year. These gross margins are given per hectare (and per acre) for cash crops; per head for livestock. Gross margin figures are calculated by multiplying physical yield by price, and subtracting "variable costs" (seed, feed, fertiliser, medicine etc.). No account is taken of "fixed" or "overhead" costs such as labour, power and machinery, property upkeep, rent or finance charges.

Planning

The approach being taken is to specify in advance the procedures to be used in generating the input parameters needed for a specific model, from the sources described above. This specification process is not yet complete. Progress so far is described here, for each type of input parameter:

- Land capability (Biophysical Characteristics).
  - Source: EDINA "Land Uses" 1 km squares. Use the categories "upland", "agricultural land", and "unuseable" (this includes "urban", "water", "woodland" and "restricted agricultural land" - see notes on EDINA classification above).
  - Procedure:
    1) Decide the question above. Probably not initially.
    2) Make a map at 1km square scale.
    3) Assign all Land Parcels to the appropriate biophysical Characteristics class.
- Initial land uses.
  - Categories: cattle, sheep, arable. For cattle and sheep, at least two levels of intensity.
  - Source: EDINA transformation of Agricultural Census data.
  - Decisions:
    - Size of Land Parcel. This should be considered in relation to farm size, so as to give a small farm perhaps 10-20 Parcels (note that in the Upper Deeside region, most of the land is not this finely subdivided physically; divisions on the proposed scale allow for modelling mixed use of larger physical parcels by sheep and cattle). A 2km square is 400 hectares; a division into 16, so that each Land Parcel is a ½ km square, will be used initially.

– Number of levels of intensity for livestock. Initially, 5 land uses will be employed: arable, high-density and low-density cattle and sheep.
– How to "translate" from the figures for 2km squares, to a set of Land Uses for the Parcels.

- Procedure:
  1) Select the exact sets of EDINA squares to use (one for calibration, and the Upper Deeside area for validation).
  2) For each EDINA square, "translate" from the figures for 2km squares, to a set of Land Uses for the Parcels. This requires two steps:
     a) Determine how many Land Parcels are to be assigned to each Land Use.
     b) Distribute these Land Parcels within the EDINA 2km square. Since land capability is available at a 1km square resolution, any difference could be used to assign Land Uses preferentially to the most appropriate subgroup of Parcels; otherwise, assign at random.

- Initial farm boundaries, size distribution.
  - Sources: EDINA transformation of Agricultural Census data.
  - Procedure for producing an initial distribution of Estates.
    1. Begin with the map divided into agricultural land/upland/other (these will be the Biophyisical Characteristics used by FEARLUS) at the 1km scale (each 1km square being divided into 4 Land Parcels – see above), and a list of the types of holdings each can be part of. Superimpose agricultural parish boundaries.
    2. Draw maps superimposed on this, with the given figures in the 2km squares:
       a) For all holdings.
       b) For holdings of different types.
    3. Calculate the total number of holdings over the entire area to be modelled:
       a) For all holdings.
       b) For holdings of different types. Check whether these sum to the total for all holdings, and if not, adjust the totals by 1 at a time, starting with the total for all holdings, then the largest total for a type (in the opposite direction), and so forth.
    4. For each agricultural parish and each map in (2), sum the numbers in the 2km squares assigned to the parish, and attach both the exact total, and the total to the nearest whole number, to the parish (for totals of n.5, round up and down alternately)
    5. For each map in (2), check the total calculated in (3) against the sum of the parish numbers. Adjust the latter as required, starting with totals of n.5 and proceeding to rounding up n.4, n.3… etc. if the overall total is too low, and rounding down n.6, n.7… etc if it is too high. If necessary, continue with adjustments of n-1 to n or n+1 to n, and so forth.
    6. Check the parish totals for all holdings against the sum of parish totals for types of holdings. If there are inconsistencies, adjust the "total holdings" totals. Due to steps 3b and 5, all parish totals and overall-area totals will now be mutually consistent.
    7. Determine ideal proportions of Land Parcel Biophysical Characteristic types for each EDINA-defined farm type.
    8. Taking farms in decreasing order of size, allocate Land Parcels in rotation. Choose the first Parcel of each farm at random, among those that could be part of it given its type. Subsequent Parcels should be selected among neighbours of existing Parcels in the farm if possible, otherwise at random.
- Climatic conditions.

- Sources: UKCIP annualised data files at 5km2 scale, for "growing season length", "summer heat wave duration", and "maximum number of consecutive dry days".
- Questions to be decided:
  - Are these the correct choices of UKCIP data files?
  - How should interpolation be handled?
- Procedure: to be determined.

- Expected yields under the relevant combinations of land quality and climatic condition.
  - Sources: SAC Farm Management Handbooks.
  - Questions:
    - Do SAC Farm Management Handbooks contain enough information?
    - If so, which years do we need.
    - If not, where can we get more?
  - Procedure: to be determined.
- Economic conditions, primarily prices for farm products and the inputs necessary to produce them.
  - Sources: DEFRA price series.
  - Questions to be decided:
    - Which price series to use.
  - Procedure: to be determined.
- Land use selection algorithms.
  - Sources: CAVES qualitative interviews, psychology and rural sociology literature.
  - Questions to be decided:
    - Which of the algorithms currently possible with FEARLUS-1.1.4 to use for initial runs.
  - Procedure: to be determined.
- Land purchase strategies.
  - Sources: CAVES qualitative interviews, psychology, farm/land economics and rural sociology literature.
  - Questions to be decided:
    - Which of the algorithms currently possible (with FEARLUS-1.1.4) to use for initial runs.
  - Procedure: to be determined.
- Land managers' social interactions (approval or disapproval of neighbours, and the importance a land manager attaches to their neighbours' good opinion, relative to that assigned to financial considerations).
  - Sources: Lee-Ann's interviews, psychology and rural sociology literature.
  - Questions: Should any social (dis)approval be included in initial CBR runs? My current inclination is not to do so, as it is not particularly clear what would be (dis)approved of in the period up to 2005, when the SFP came in. However, it is worth reviewing the interviews and reconsidering: there are possibilities – in attitudes to organic farming, hobby farming, and tourism-related activities.
  - Procedure: to be determined.
- Off-farm income.

  - Sources: CAVES qualitative interviews.
  - Questions:
    - Whether to include off-farm income in initial runs.
- I would be inclined to try runs without off-farm income, and with off-farm incomes drawn from normal distributions (and fixed over time), to see how much difference the amount and spread of such income matters in the model.
  - Procedure: to be determined.
- Economies of scale.
  - Sources: SAC Farm Management Handbooks, CAVES qualitative interviews.
  - Procedure: to be determined.
- Government policies.
  - Sources: European Commission, MAFF/DEFRA and Scottish Office/SERAD/SEERAD documents.
  - Questions:
    - Whether to explicitly include policy in the initial runs, or deal with it by using changes in returns and costs.
    - What policies and policy changes to include.
    - Procedure:
    - Produce a list of policy changes that will be included in the model.
    - Work out how to implement them as changes in returns and in costs.
- Check subsequent changes in the initial distribution of Estates against subsequent figures for the distribution of holdings.
  - Effectively, the same procedure as for input can be used, and the results compared with what the model has produced.
  - Alternatively, use current actual boundaries from IACS data (how far back does this go?)
  - Repeat runs of the model, and test whether the actual figures are sufficiently within the "envelope" of the model runs. For this purpose, we will need to:
    - Decide on a set of measures (e.g. numbers of farms, largest farm, smallest farm, ratio between largest farm and the total area).
    - Perform a large number of runs.
    - Check where the "real world" measure lies in relation to the range of measures from the runs. I am not sure it would make sense to set precise accept/reject standards here (e.g. that every value must lie within the 95% confidence limits), since taking a greater number of measures would increase the chances one of the real world measures lay outside these limits, but the different measures will not be independent: describing the "profile" of results might be the most useful approach.

2.3.3 Future Work

Future work on modelling Upper Deeside within FEARLUS will be funded by the Scottish Government. The work will consist of the following steps:

1. Completing, calibrating and validating a model of Upper Deeside agricultural land use and changes in farming, 1980-2005.
   a) Completing the specification of the process to be used in generating input parameters.

b) Selecting an area of Grampian region, not adjacent to Upper Deeside but similar to it in topography, climate and other relevant characteristics, and following the process to be specified in step 1a.

c) Define a calibration procedure to be employed in this region. This will involve searching the space defined by those parameters with values not determined by data sources, for the values best able to produce outputs resembling the real-world outcome.

d) Following this calibration procedure. The result will be a set of calibrated parameters.

e) Following the process of step 1a for the Upper Deeside area. Combine the parameters thus generated with the calibrated parameters to generate a model for validation.

f) Perform multiple runs of the model with these parameters, using the Quantis random number generator described above.

g) Check whether any quantifiable real-world outcomes differ significantly from the ensemble of simulation run results.

h) If so, use the differences, along with the CAVES qualitative interview transcripts to design changes to the model.

i) Repeat stages d through h as necessary.

2. Modelling policy-relevant scenarios. The model emerging from step 1 will be applied, with a range of assumptions about future economic and climatic scenarios over the period to 2050, to examine possible Scottish Government policies to encourage land use practices that will reduce net Scottish greenhouse gas emissions, while maintaining economic and social sustainability. Initially, the model will be applied in Upper Deeside, then in other parts of the Grampian region, and finally in other rural upland regions of Scotland. Potential policies to be investigated will be determined in discussion with policymakers.

3. Journal articles will be produced. The first of these, on the use of qualitative interview information to inform agent-based models of land use, is currently in preparation. Further articles will report the results of the calibration and validation procedures outlined in step 1 above, and the modelling of policy-relevant scenarios outlined in step 2.

2.3.4 References

Beaton, C., Catto, J. and Kerr, G. (eds) (2007) The Farm Management Handbook 2007/08. Scottish Agricultural College.

Hellekalek, P. (1998) Good random number generators are (not so) easy to find. Mathematics and Computers in Simulation 46, 485-505.

Hotson, J. McG. (1988) Land Use and Agricultural Activity: An Areal Approach for Harnessing the Agricultural Census of Scotland.

van Niel, K. and Laffan, S. W. (2003) Gambling with randomness: the use of pseudo-random number generators in GIS. International Journal of Geographical Information Science 17 (1), 49-68.

2.4 Politechnika Wroclawska

2.4.1 Role-Playing Game for the Odra Case Study

During the last eight months the WUT team in cooperation with the team from Wroclaw University worked on the role playing game, designed as a validation tool for the model built by the Kassel team. The development of the game, called AgroGame, was completed in November, just before the workshops with stakeholders that took place on the 20th, 21st and 22nd of November. The data gathered during those three days is still being analysed.

AgroGame is a simple multiplayer computer game. Each player of the game plays a role of a farmer, who owns a parcel of arable land. There are six such parcels in the game, each owned by a different farmer. The parcels lay on a piece of land of small and homogeneous slope, along a homogeneous channel, that runs through the centre of every parcel. This channel is one of the key elements of the game's model. Its purpose is to remove extra water from adjacent fields. The parcels' area and land use is homogeneous. The game works with a yearly time step. A single game's session consists of 15 time steps. In each time step the game's model calculates yields and a very simple economic balance for every parcel. At the start of the game, each farmer has the same amount of money. During the game farmers spend their money on agricultural production and optionally on maintaining the channel. They earn money by selling agricultural products.

There are two variables that are used for measuring players' performance. One is wealth accumulated during the game and another is reputation. Reputation is measured in special “reputation points”. The more “reputation points” a player has, the higher their reputation. At the end of the game, players are presented with game results that consist of values of wealth and reputation for all players. The values of wealth and reputation are treated as equally important and it is up to the player to decide which variable is more significant to them.

In every turn a player can undertake certain actions. These actions can be divided into two groups: channel maintenance related actions and social interactions. In the first group there are only two possibilities. A player can decide to either maintain or not maintain their segment of the channel. If a channel's segment is not maintained it cannot remove extra water from adjacent fields. This can lead to local flooding during wet years, which in turn leads to losses of crops. On the other hand, maintaining a channel costs money and during years with average rainfall the flood protection function of the channel is negligible. Even if the channel is not maintained, there will be no flooding. Hence, the player is faced with a dilemma: maintain the channel and have relatively small income, due to high maintenance costs, regardless of weather, or not maintain and have high income during normal years and losses during wet years. The players only have knowledge about the current weather. There are no forecasts available. The effectiveness of the channel in preventing floods on a given field depends not only on its condition along this field, but also along a number of fields below. In other words even if a player maintains their part of the channel, but his neighbours below do not maintain their parts, then the player can still have losses due to flooding. This is the worst possible scenario, as the player pays for channel maintenance and despite of that has poor crops.

The second group of actions consists of social interactions. These are limited to just two or three possibilities depending on the game's mode. The basic two actions are praising and criticizing. Every player can praise or criticize any other player. When a player is praised, their reputation rises by one point. If they are criticised, they loose one point of reputation. During a single time step a player can criticise or praise another player only once. The game can be run with an additional mechanism of social interaction. It's a complaint. Any player can file a complaint to the municipality office against any other player who is not maintaining their part of the ditch. The complaint takes effect in the next time step and causes the player against whom such a complaint was filed to be fined a penalty fine. This fine is high and can have a significant effect on the player's wealth. However, the player against whom such a complaint was filed will be informed about this fact during the same time step and can decide to start maintaining the channel and by so doing, evade the penalty. A player who decides to file a complaint against another player must be aware of the fact, that the other player will be informed about the complaint, will know who filed it and so can undertake some retaliatory actions like criticising.

AgroGame is designed to be played by six players. In order to play the game during workshops with stakeholders, we created small local networks made up of seven notebook computers connected by a router. Six computers served as client machines and one was the game server. The game's engine and the model were hosted on the server, while clients were only responsible for presentation of the game's state and sending players' input to the server. Both the server and client applications were written in Java. Communication was implemented using RMI, which greatly eased the development process. The figure below depicts the game's main screen.

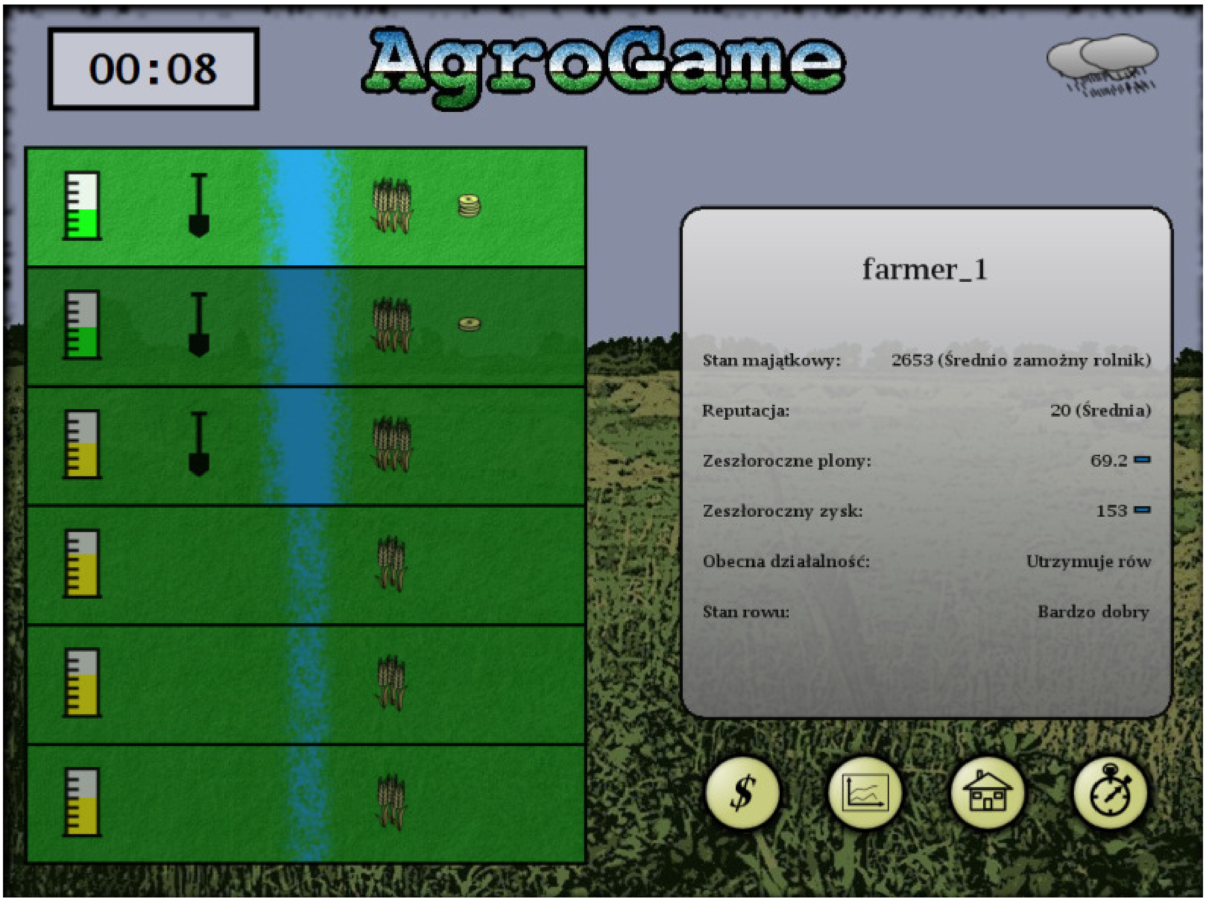


Figure 1: Main screen of the AgroGame

2.4.2 Binary Choice (Opinion Dynamics) Models

The results generated in earlier project phases were synthesized and written down in a series of publications – some of them are already submitted others are in the final stage of preparation.

2.4.3 Evolution of Social Networks – A Theoretical model of an evolution of an affiliation social network.

The work on the algorithm that allows for construction of social networks based on group affiliations has been completed, and the process of publication of results has begun. Based on the assumption of Newman and Park (2003) regarding positive assortativity as typical for social networks, we created an algorithm that constructs networks of desired high assortativity. The algorithm is based on mechanisms that are as plausible from the point of view of social science, as possible at this stage of research. The mechanisms include: group membership that allows forming social ties within the group with relatively high probability; inborn human tendency to maintain fairly stable (yet different for different people) number of friends; and human ability to adapt to social environment and modify these inborn preferences depending on the situation. Different attractiveness of social groups has been taken into account as well, operationalised as probability of becoming a member of each group. What is more, parameters of the algorithm, such as affiliation distribution or degree distribution were based on results of big social surveys, such as The International Social Survey, The General Social Survey and DDB Life Style Survey.

Operation of the algorithm has been tested for different shapes of distributions (e.g. normal, exponential, Burr distribution) of key variables (degree distribution, affiliation distribution).

Behavior of the algorithm in regard to assortativity changes has been tested in relation to the value of parameter that represents the process of adapting the pre-defined desired number of social ties depending on the social environment (operationalised as membership in one or many groups).

The first paper presenting results of our research has been accepted for publication in Acta Physica Polonica A (in English). Another paper is in final stages of preparation.

2.4.4 Articles and Conference Presentations

- P. Hetman, P. Magnuszewski; Equilibrium properties of evolving binary choice models on networks, XXIII Max Born Symposium "Critical Phenomena in Complex Systems" 2-6 September 2007, Polanica Zdroj, Poland.
- A. Janutka, P. Magnuszewski; Dynamics of probabilistic social-impact model, XXIII Max Born Symposium "Critical Phenomena in Complex Systems" 2-6 September 2007, Polanica Zdroj, Poland.
- Hetman P., Magnuszewski P., Stefańska J., Bujkiewicz L.; How groups shape social network - on assortativity of social networks; European Conference on Complex Systems – Dresden, Germany, 1-5 October 2007.
- P. Hetman, P. Magnuszewski, J. Stefańska, L. Bujkiewicz, K. Ostasiewicz; How nodes and groups properties influence assortativity in social networks?; accepted for publication in Acta Physica Polonica A.
- Ostasiewicz, K., Tyc, M. H., Goliczewski, P., Magnuszewski, P., Radosz, A., Sendzimir, J.; Multistability of impact, utility and threshold concepts of binary choice models; submitted to Physica A.
- Janutka A., Magnuszewski P.; Structural transitions in non-equilibrium model of social impact; submitted to European Physical Journal – Special Topics.

3 Case studies

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

3.1 South African case study (SEI Oxford)

Over the past eight months the SEI team has mainly focused on validation. A fieldtrip to Sekhukhune in late October 2007 was prepared in a meeting of the case study and modelling teams in Manchester. Also, SEI has taken the lead in producing three Deliverables for the CAVES project.

3.1.1 Meeting in Manchester, UK

On 10th October 2007 travel to Manchester was arranged to meet with the team from the Centre for Policy Modelling for a CAVES validation planning meeting ahead of upcoming fieldwork in Sekhukhune, South Africa (see below). The meeting established the validation protocol to be used and the selection of specific activities for the fieldwork. It was then decided that we would produce model traces (CPM team), corresponding narrative story lines (SEI team) and characterisations of model macro-outputs (CPM team). The list of validation activities devised at the project meeting in Kassel (September 2007) was reviewed.

3.1.2 Validation field trip, Sekhukhune District, South Africa

From 29th October to 2nd November 2007 fieldwork was carried out in Tubatse municipality, Sekhukhune District, Limpopo Province, with the aim of validating some of the model outputs that the Manchester modelling team had provided from their agent-based models. A participatory validation exercise was done with the villagers where the start of a story (based on storylines produced from model traces) was presented on cards and the villagers provided feedback on the plausibility and likely future situation. An exercise was also done with municipal managers on scenarios of macro-level output that included reference to in- and out-migration, income levels, food security and school attendance. The video that was produced after the April fieldtrip was shown and appreciated by many. Leonie Joubert, a South African author and journalist joined the fieldtrip. Her research will form a chapter in her new book, Boiling Point, on climate change and communities and she will write a popular article for SEI.

3.1.3 Validation Report

In the following few days the validation field trip was written up as the “validation report” (3 documents) and provided to the modelling team in Manchester. To summarise:

Villagers commented on the story and adjusted it according to their perceptions of the most likely case given their experience of living in the village. Additional cards were added to capture the changes presented by the villagers and lines were drawn to represent some of the connections between people. It was necessary to adjust some of the original storylines in order for the initial situation to be presented in a logical way and then commented on. Once the initial case was agreed, then the consequential events and actions were presented and commented on. Some stories were more appropriate to present for feedback than others, so one storyline was left out.

The new/revised logic would be available to be integrated into the model.

For the municipal meetings, several sets of macro-level outputs were selected and prepared as time-line plots. Also, some major movements in these plots were described by some pieces of text. Presenting the real plots might have been misread to suggest that the models produced some usable forecasting outcomes; therefore, schematic diagrams were used instead.

The officers were asked the future situation of out-migration in the case of 15 new mining companies coming to the GTM (and they mentioned there are currently another 8 permits awaiting finalization to add to the existing 17 mines). They said that if more people were educated properly to work on the mines, some people might not feel it necessary to migrate out of the GTM. If the job opportunity from new mining and education such as universities would be increased, the rate of out-migration would be also lowered. Nonetheless, they said that the rate would not be 0% as some people always left the municipality.

Considering the possibility of closing down of some of 17 existing mining companies, they responded that out-migration would increase: the men's migration rate will be between 2% and 5% but that of women will be no difference, i.e. between 1 and 3%.

The officers were asked about in-migration. They thought that the current in-migration rates for the whole of GTM were 30% for men and 15% for women. The officers said that the in-migration would increase to 65% and 30% at growth points and whole municipal, respectively with new 15 mining companies. There are no model outputs on this topic at present.

3.1.4 Deliverables

Deliverables coordinated by SEI for the CAVES project consist of the following working papers:

- Deliverable 8- Case Study Structure, Stakeholder/Agents and Validation Data. Produced by SEI and accepted by the CAVES management team August 2007.
- Deliverable 10 - Case Study Validation Protocol working paper. Produced by SEI and accepted by the CAVES management team March 2008.
- Deliverable 12 - working papers on results from each case study. Prepared April 2008 and submitted (awaiting response from management team).

3.1.5 Recent Publications

Publications comprise a journal article and a poster presentation:

- Gina Ziervogel and Anna Taylor: "Feeling Stressed: Integrating Climate Adaptation with Other Priorities in South Africa". Environmental Journal, March /April 2008.
- April 3rd 2008: Poster presented at Global Environmental Change and Food Systems (GCAF) conference "Food Security and Environmental Change: Linking Science, Development and Policy for Adaptation", session on Food Security

3.1.6 Future work

Future work will involve working on journal papers, book chapters, conference papers and other articles. These opportunities will be discussed at the final project meeting in Manchester in April 2008. Some further research areas identified to pursue are: Climate Change and Social Protection: viewing the Sekhukhune case through the lens of social protection.

There are many adaptation strategies that can be developed within sectors, such as the use of drought-resistant crops or the improvement in early warning systems. However, these require an understanding of the specific nature of climate impacts. Social protection, however, is intended to strengthen the social safety nets that underpin the livelihoods of vulnerable communities. If these safety nets are truly strengthened by social protection programmes, it is more likely that people will be able to respond to the many of the impacts of climate change themselves. This is particularly important in order to support the 'agency' of people and give them the choice in how they might choose to respond to the stressors they face.

3.2 Grampian case study (Macaulay Institute)

3.2.1 Overview

During the October – April reporting period, final data analysis and reporting were completed. In addition, a journal paper based on Grampian field research was submitted to Sociologia Ruralis, and two further papers are in progress. A proposal for further qualitative research following on from the Grampian study has been funded through the Scottish government.

3.2.2 Field Research and Data Analysis

Primary field research was completed in March 2007, followed by field research validation completed in September 2007. Due to the ample material collected during primary field research, and the absence of major issues raised in the validation process, it was decided not to undertake a further stage of field research, and concentrate instead on data analysis and journal publications.

Selected aspects of the final version of the FEARLUS model, implemented as a result of Grampian study findings:

- the decision rule that farmers do not change their current crop or type of stock when their aspiration threshold has been reached, even if there are higher prices in a different commodity.
- in ordering the factors which farmers take into consideration when changing commodity, the anticipated profitability of the new commodity in the near future is of primary importance.
- land is differentially desirable, on the basis of previous (and therefore anticipated) profitability.
- farmers will always bid on neighbouring land, if they have sufficient resources (according to a threshold specific to the individual farmer).
- decision rules about land acquisition can be made specific to farming cohorts (e.g. entrepreneurial, traditional, pluriactive, lifestyle/hobby/environmental).
- fixed costs can be associated with commodities, and differences between commodities in economies of scale can be defined.
- the principle that farmers will put up with losing money without changing commodity immediately, although restrictions to this principle on the basis length of time over which loss occurs and extent of loss have not been quantified.
- including more specifics on proximity of land and expansion plans in decisions regarding land acquisition. (Expansion plans can be handled through farmer type; the definition of proximity is based on geographic neighbourhood, but can to some extent be varied by the model user).
- off-farm income or initial wealth as an important factor in agents' decisions and ability to withstand shocks and stresses.
- allowing farmers to sell up and quit farming in circumstances other than bankruptcy

A final copy of study findings has been mailed to all study participants.

3.2.3 Policy Recommendations

The purpose of the CAVES study is to generate scenarios in order to inform policy. Several policy implications emerge directly from the qualitative study. These include:

- Differential response to policy based on farmer type must be taken into consideration.
- Pace of response: farmers seldom make major changes on-farm based on a single year's returns. Farmers are much more likely to make changes in response to permanent or long-term market and policy shifts. It is therefore important to implement policies around which

farmers can build long term plans.

- Farmers appear to respond more quickly to 'opportunities' rather that 'needs must' situations. A 'carrot' rather than 'stick' of many current policies would appear more likely to be successful in achieving policy objectives.
- Greater success in farmer response to grant programs can be expected if these are in line with current farming priorities, and definitions of 'good farming'.
- Although farmers are concerned about the environment, engagement in current environmental programs largely reflects a desire to recoup lost income from commodity-based payments.
- Although most farmers in the study site have reduced their level of chemical inputs, they resist 'the last 20%' it would take to achieve organic status.

Further policy implications will be drawn from the scenarios generated through the computer models.

3.2.4 Areas for further research

The qualitative field research generated multiple research questions for further investigation. The issue of amenity value farming, and how this plays out in integration into farming networks and orientation to environmental programs, has already been funded through the Scottish government for further research. This qualitative study will build directly on the interviews with amenity value farmers identified in the CAVES Grampian project.

Other areas for further exploration, either through theoretical application to existing data or further research include:

- An analysis of 'multifunctional' attitudes and behaviours in the study site.
- The social construction of 'opportunity' among different farming types
- The dynamics of time constraints in farm decision-making: commodity price cycles, seasonal variations and the physical limitations on changing commodity type
- The complexity of farming networks and farming norm generation
- The integration of environmental program participation into definitions of 'good farming'.

3.2.5 Academic Publications

A journal article by Lee-Ann Sutherland entitled "Farming in the Network Society: Social Capital in Grampian Agriculture" was submitted to Sociologia Ruralis in November, 2007.

Abstract

This paper addresses the issue of land-based industries in the network society (Castells 2004). Drawing on a case study of farming resource networks in Scotland, the author describes the information, labour and equipment sharing practices of study site farmers, and analyses these with a combined conceptual framework of Castells' (2004) network society and Bourdieu's (1986) social capital. Study findings demonstrate a clear distinction between highly localised patterns of labour and equipment sharing, and highly varied and geographically extended patterns of information exchange. While the dependence on dispersed information networks and technology are consistent with the network society thesis, analysis suggests that in combination with the temporal and spatial

limitations of agricultural production, ‘network society’ engagement facilitates ongoing productivist, rather than post-productivist agriculture in the region.

Two further papers, one addressing multifunctional agriculture, and a further joint paper with the modelling team, are in progress.

3.3 Odra Valley case study (Uniwersytet Wroclawski)

3.3.1 Social research

Role-playing game

We have designed and performed the role-playing game focused on LRS maintenance as a tool for model validation and a kind of sociological experiment. There were three sessions performed in three villages, each of them consisting of the 15-minutes training game and two main games lasting for about an hour. Every player was accompanied by two students, who made observations of the process and carried out interviews with players during and after the games.

Later on we developed the scheme of analysis of the game results, both quantitative and qualitative. First stage of data processing is now finished (preparation of quantitative data for statistical calculations and the code book for coding of interview transcripts). Further analysis is now being conducted.

Field research

The last wave of social field research was carried out in winter 2007. We collected 30 interviews in the Peclaw area, chosen as a reference area due to well maintained LRS and active Water Partnership.

3.3.2 Conceptual model of LRS

Based on all collected information we have developed a conceptual model of Land Reclamation System maintenance using the method of Causal Loop Diagramming.

3.3.3 Cooperation with the modelling team

We had regular e-mail / Skype contacts with the modelling team and we provided them with additional data necessary for the development of the agent based models.

3.3.4 GIS

- Another LULC (land use/land cover) time series, from 1975, was digitised within the Rogow Legnicki and Peclaw research areas from a 1:25 000 topographic map. It was used together with two other series already digitised, from 1930's and 1990's, to analyse LULCC (LULC changes) at three time points. The results of the analysis were also used for developing a framework for spatial error assessment in LULCC research.
- A raster DEM (digital elevation model) of cell size 10m was created in GRASS GIS by means of natural neighbour interpolation, from TIN (triangulated irregular network) data provided by Lower Silesia WODGiK (Voivodship Department of Geodesy and Cartography). It replaced DEM previously used the in biophysical model. Current DEM has is more reliable for hydrological modelling. According to provider's metadata the original

TIN data accuracy varies from 0.27 m to 0.45 m, 0.35 m in average, whereas the former DEM had accuracy of about 1.5m in average, as it was interpolated from elevation contour lines digitised from a 1:10 000 topographic map.

- An orthophotomap from aerial photographs taken in 1995-1996 was prepared for Rogow and Peclaw areas. It was used as an aid material in LULC interpretation.
- Previously released within the course of CAVES project GRASS script extensions r.surf.nnbathy and v.breach, were updated (note that GRASS AddOns website address has changed to http://grass.osgeo.org/wiki/GRASS_AddOns).
  - r.surf.nnbathy 1.96 changelog:
    - fixed handling spaces in pathnames,
    - got rid of non – POSIX Shell syntax,
    - better run-time error trapping,
    - require nnbathy 1.76 because it has an important bug fixed (refer to file CHANGELOG in the nn package source code for details; available at http://www.marine.csiro.au/~sak007),
    - cosmetics, the manual cleaned up.
  - v.breach 5.9.7 changelog:
    - workaround for a feature of GRASS g.region module that a syntax "g.region -a res= vect=" might not work as expected, depending on the initial region settings.

3.3.5 Presentations and publications

- The presentation: Karolina Krolikowska, Andrzej Dunajski, Piotr Magnuszewski 'The impact of institutional changes on land amelioration and land use in Odra River Valley, Poland' was presented during the international conference 'Impact Assessment of Land Use Changes', 6-9.04.2008, Humboldt University, Berlin. This presentation will be the basis for the paper in the special issue of the peer-reviewed journal chosen by the relevant session leader.
- Presentation at "Methodology of landscape research" conference, Sosnowiec-Krynica, Poland: Andrzej Dunajski, Maciej Sieczka, 2008, Land use/cover change analysis based on historical maps – a framework for spatial error assessment
- Andrzej Dunajski, Maciej Sieczka, 2008, The influence of the rectification error on the land cover change analysis results using topographic maps as an information source, (submitted publication)
- We have prepared the Odra Case Study Final Report. In this report we describe all steps undertaken by the Odra Case Study team during three years of the CAVES project and the results of our work, including the final results of legal and economic analysis, social survey research, biophysical research, and stakeholders' workshops as well as preliminary results of the role-playing game. The report was meant to be the input for Deliverable 12, as well as the base for further publications and case study area monograph.

4 Publications

The CAVES project is dedicated to publishing 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 last seven months of the project, the following papers have been written:

- Shah Jamal Alam, Ruth Meyer and Emma Norling: *A Model for HIV Spread in a South African Village*. Accepted for MABS 2008, 9th International Workshop on Multi-Agent-Based Simulation, May 12-13, Estoril, Portugal.
- Shah Jamal Alam and Ruth Meyer: *Comparing Two Sexual Mixing Schemes for Modelling the Spread of HIV/AIDS*. Submitted to WCSS-08, World Congress on Social Simulation 2008, George Mason University, Fairfax, July 14-17, 2008. Online available as CPM-Report No. 08-189 <http://cfpm.org/cpmrep189.html>
- Andrzej Dunajski, Maciej Sieczka: *Land use/cover change analysis based on historical maps – a framework for spatial error assessment*. Presentation at "Methodology of landscape research" conference, Sosnowiec-Krynica, Poland, 2008
- Andrzej Dunajski, Maciej Sieczka, 2008, *The influence of the rectification error on the land cover change analysis results using topographic maps as an information source*, (submitted publication)
- P. Hetman, P. Magnuszewski; *Equilibrium properties of evolving binary choice models on networks*, XXIII Max Born Symposium "Critical Phenomena in Complex Systems" 2-6 September 2007, Polanica Zdroj, Poland.
- Hetman P., Magnuszewski P., Stefańska J., Bujkiewicz L.; *How groups shape social network - on assortativity of social networks*; European Conference on Complex Systems – Dresden, Germany, 1-5 October 2007.
- P. Hetman, P. Magnuszewski, J. Stefańska, L. Bujkiewicz, K. Ostasiewicz; *How nodes and groups properties influence assortativity in social networks*?; accepted for publication in Acta Physica Polonica A.
- A. Janutka, P. Magnuszewski; *Dynamics of probabilistic social-impact model*, XXIII Max Born Symposium "Critical Phenomena in Complex Systems" 2-6 September 2007, Polanica Zdroj, Poland.
- Janutka A., Magnuszewski P.; *Structural transitions in non-equilibrium model of social impact*; submitted to European Physical Journal – Special Topics.
- Krebs, F., Elbers, M. and Ernst, A. (2008): *Modelling the social and economic dimensions of farmer decision making under conditions of water stress*. Proceedings of the 1st ICC workshop on complexity in social systems, Lissabon.
- Karolina Krolikowska, Andrzej Dunajski, Piotr Magnuszewski: *The impact of institutional changes on land amelioration and land use in Odra River Valley, Poland*. Presentation at the international conference 'Impact Assessment of Land Use Changes', 6-9.04.2008, Humboldt University, Berlin.
- Ostasiewicz, K., Tyc, M. H., Goliczewski, P., Magnuszewski, P., Radosz, A., Sendzimir, J.; *Multistability of impact, utility and threshold concepts of binary choice models*; submitted to Physica A.
- Scott Moss: *Simplicity, generality and truth in social modelling. Part 1: Epistemological issues.* Submitted to WCSS-08, World Congress on Social Simulation 2008, George Mason University, Fairfax, July 14-17, 2008. Online available as CPM Report No. 08-187 <http://cfpm.org/cpmrep187.html>

- Scott Moss: *Simplicity, generality and truth in social modelling*. *Part 2: Demonstration.* Submitted to WCSS-08, World Congress on Social Simulation 2008, George Mason University, Fairfax, July 14-17, 2008. Online available as CPM Report No. 08-187 <http://cfpm.org/cpmrep187.html>
- Lee-Ann Sutherland: *Farming in the Network Society: Social Capital in Grampian Agriculture*. Submitted to Sociologia Ruralis in November, 2007
- Gina Ziervogel and Anna Taylor: *Feeling Stressed: Integrating Climate Adaptation with Other Priorities in South Africa*. Environmental Journal, March /April 2008.

5 Deliverables

With the end of the project, several major Deliverables are due. These are:

No.	*Title*	*Lead*	*Project Month*
10	Working paper on case study validation	SEI	35
11	Working paper on modelling and case studies in policy analysis process	MMU	38
12	Working papers on results from each case study	SEI	38
13	Final models as open source libraries	UNIK	38
14	Technical reports on the final models	UNIK	38
15	Handbook on conceptual and formal models to measure resilience of complex systems	WUT	38
16	Final Report	MMU	38