



J. Gary Polhill, Lee-Ann Sutherland and Nicholas M. Gotts (2010)

Using Qualitative Evidence to Enhance an Agent-Based Modelling System for Studying Land Use Change

Journal of Artificial Societies and Social Simulation 13 (2) 10

<<http://jasss.soc.surrey.ac.uk/13/2/10.html>>

Received: 18-Feb-2009 Accepted: 15-Nov-2009 Published: 31-Mar-2010

Abstract

This paper describes and evaluates a process of using qualitative field research data to extend the pre-existing FEARLUS agent-based modelling system through enriching its ontological capabilities, but without a deep level of involvement of the stakeholders in designing the model itself. Use of qualitative research in agent-based models typically involves protracted and expensive interaction with stakeholders; consequently gathering the valuable insights that qualitative methods could provide is not always feasible. At the same time, many researchers advocate building completely new models for each scenario to be studied, violating one of the supposed advantages of the object-oriented programming languages in which many such systems are built: that of code reuse. The process described here uses coded interviews to identify themes suggesting changes to an existing model, the assumptions behind which are then checked with respondents. We find this increases the confidence with which the extended model can be applied to the case study, with a relatively small commitment required on the part of respondents.

Keywords: Agent-Based Modelling, Land Use/Cover Change, Qualitative Research, Interdisciplinary Research

Introduction

- 1.1 In this paper, we suggest a methodology for using qualitative research to inform the incremental development of an agent-based model of land use change, showing explicitly how research findings are related to modelling decisions. Agent-based modelling is particularly well suited to studying coupled human-natural systems (Hare and Deadman 2004). Boulanger and Bréchet (2005), highlighting the promise of agent-based modelling in the study of sustainable development, note that it allows an intuitive representation of the environment and the embedding of agents within it. Parker and Meretsky (2004) contend that agent-based models can capture dynamic feedbacks among land use patterns, spatial location and land use composition. Bousquet and Le Page (2004) conclude that researchers in ecology and the social sciences can use agent-based modelling to study the interactions among spatial, network and hierarchical levels of organisation, a view supported by Huigen (2004). There is also a considerable body of literature on land use change across the social sciences, and increasing pressure from funding bodies to integrate not only research findings from other disciplines, but to undertake interdisciplinary work that will result in data that are useful to multiple fields. The methodology used here was developed in conjunction with a rural sociologist, and used to ground a largely abstract model in empirical evidence.
- 1.2 The work was conducted as part of the CAVES (Complexity, Agents, Volatility, Evidence and Scale) EU project, which ran from March 2005 to April 2008. The purpose of the project was to study complexity, taking responses to shocks in coupled human-natural systems as the research domain. The project gathered evidence from three case studies: one in the Vhembe district in the Limpopo province of South Africa, looking at AIDS; another in the Odra river valley in Poland looking at irrigation; and one in the Grampian region of Scotland. The focus of the Grampian region case study was on land use change patterns in the Upper Deeside region of North East Scotland (figure 1) between the mid 1980s and 2030. Specific foci were on the drivers and processes of land use change, and the particular role of social networks in these processes. The case studies were used to inform the development of three agent-based computer models, which were then to be used to examine statistical signatures of complexity. This paper concentrates on the work done for the Grampian case study and the development of its associated model, rather than issues of complexity.



Figure 1. Map (courtesy of Freemap) with the approximate location of the Grampian case study area highlighted using an orange ellipse.

- 1.3 The phenomenon of particular interest in the Grampian case was the apparent resilience in land use and land ownership change to various shocks over the past 20 years: changes to the EU Common Agricultural Policy (CAP) in 1992 and 1999, the exit of sterling from the Exchange Rate Mechanism (ERM) in 1992, the Bovine Spongiform Encephalopathy (BSE) epidemic, and the foot-and-mouth outbreak of 2001. These could be expected to result in considerable changes to land use. Instead, farmers in the region largely continued to produce what they always have: beef and lamb, with some arable crops primarily used as livestock feed. Possible reasons for this resilience include unsuitability of land for other commodities, uncertainty over the permanence of the identified 'shocks' and the difficulty shifting from one commodity to another particularly in terms of livestock. However, the academic literature also identified a number of sociocultural reasons for this lack of change. Burton's (2004) work just north of the study site found underlying cultural sources of resistance to change, hinging on an embedded farming culture of independence and maintenance of farming traditions. Shucksmith and Herrmann's (2002) paper also based on responses from farmers near and in the study site, emphasised the differential responses of farmers: given similar situations, specific cohorts of farmers respond differently (e.g. supplementing income through off-farm employment versus expanding production to create economies of scale). Only a small subset favoured experimentation with new business options. It was therefore decided to address primarily sociocultural issues empirically as a means of grounding the model.
- 1.4 Robinson et al. (2007) review various methods for building empirical agent-based models in the domain of land use change. Of these, participant observation and 'companion modelling' (Barreteau et al. 2003) are qualitative, involving narrative rather than statistical analysis of evidence. Both are varieties of participatory agent-based modelling, in which members of the study population become actively involved in model design and validation (e.g. Bharwani et al. 2005). Participatory ABM is one of a suite of tools applied to environmental integrated assessment (Hisschemöller, Tol and Vellinga 2001). Whilst Robinson et al. (2007) note problems with generalisation among the weaknesses of these techniques, a further issue is the considerable amount of interaction with stakeholders involved. This entails an investment of money and time by researchers and stakeholders that will not always be available, particularly if some parties are unconvinced of the benefits. Farmers especially, busy for most of the year, can be wary of engaging in such processes. Further, where the project is committed to a particular framework or theoretical basis for modelling, as in CAVES where the emphasis was on complexity, allowing stakeholders too much influence over model structure could result in a model inappropriate for the study originally funded (for example, if the model exhibited none of the statistical signatures associated with complex systems). Indeed, stakeholders may in any case not have been convinced that academic study of complexity was sufficiently relevant to justify their personal participation. Instead, it was decided to use a qualitative strategy that collected data through interviews on the topic of decision-making and change processes, rather than asking participants to give input directly to the model.
- 1.5 Qualitative research techniques extend beyond participatory approaches, and include qualitative interviewing, discourse and conversational analysis, and analysis of texts and documents (Bryman 2008), although interviewing is the most favoured tool (Denzin and Lincoln 1998). There is no widely agreed definition of qualitative research, but most proponents include an emphasis on research which is inductive (theory generated from data), interpretivist (reality as socially constructed) and emphasising meanings and descriptions. As such, the resultant data are typically composed of words and images, rather than numbers. The choice of qualitative, as opposed to quantitative (number based) methodologies, depends on how the data will be used: where statistical treatment is required, such as for model parameterisation, quantitative methods are the better choice. Qualitative approaches may also be used when quantitative data are not available or are too expensive to obtain, and the lack of quantification is not too significant an obstacle.^[1] Qualitative interviewing is particularly useful for studying complicated relationships and slowly evolving events or processes (Rubin and Rubin 1995) and can therefore be seen as useful for informing model ontology. Findings from qualitative research tend to be specific to the population studied: an advantage when considering issues affecting a particular population, but a disadvantage for generalisation. However, Mason (2002) argues that qualitative research can be generalised: that concepts and theories generated from the research should have relevance beyond the study site, even if the precise data cannot be replicated.
- 1.6 Here, we are interested in making improvements to the FEARLUS (Framework for the Evaluation and Assessment of Regional Land Use Scenarios) agent-based model, which pre-existed the CAVES project, to better fit it to the requirements of the Grampian case study. FEARLUS was first conceived in 1998, and on-going work with FEARLUS has explored the relative performance of various heuristic strategies for choosing land uses in different environmental contexts (Polhill, Gotts and Law 2001; Gotts, Polhill and Law 2003; Gotts, Polhill and Adam 2003; Gotts et al. 2003; Gotts and Polhill 2009b). Other work has explored the use of collective rewards in managing pollution from land uses (Gotts and Polhill 2007), and more recently FEARLUS has been coupled with an ecological metacommunity model to explore the impact of land use policy on biodiversity (Polhill et al. 2008).
- 1.7 There were good reasons to believe that the case study would be a reasonable fit to the functionality provided by FEARLUS: both the case study and the modelling system pertain to land use decisions made by farmers. However, there were also reasons to expect that we would have to make changes. Boero and Squazzoni (2005) suggest three broad categories of agent-based model: 'theoretical abstractions',

'typifications' and 'case-based models', proposing FEARLUS as an example of the second kind. These categories cover a spectrum from conceptual to empirical agent-based models. In being applied to a specific case-study, changes would be needed to enable FEARLUS to work closer to the empirical end of the spectrum.

- 1.8 With any new scenario, there is the question of whether to use existing modelling software, or to create a new model. One of the supposed benefits of object-oriented programming is the reuse of code (Lewis et al. 1991). The advantage of adapting an existing piece of software is that previously tested functionality can be preserved. The disadvantage is that legacy functionality is imported with it, and design issues associated with that functionality may inhibit the creation of desired features in the new model. However, given sufficient correspondence between the desired features of a model of the scenario and an existing piece of software, the effort required to change the existing software to meet the requirements will probably be less than the effort required to build a wholly new model. Given a fixed amount of time, therefore, an adapted piece of software will have more of the required features of a model for a scenario than a new piece of software would. Frenken (2006), advocating a method for building models he terms 'TAPAS' (Take A Previous model and Add Something), points out that incremental modelling strategies are more successful, faster to build, and easier for others (presumably familiar with the previous model) to understand. This is the approach adopted here.
- 1.9 The paper is structured as follows. In the next section, we describe the FEARLUS modelling system prior to making changes to better fit it to the Upper Deeside region for the Grampian case study of CAVES. A methodology section then outlines the qualitative interviewing approach, how the evidence was then used in making changes to FEARLUS, and how the assumptions behind these changes were subsequently checked with respondents. In contrast to the normal approach in describing how a case study and a model interrelate (in which the case study is described in one section, and then the model in another), we use the results section to show in a detailed and systematic way how findings did or did not lead to changes to FEARLUS. These changes, and the checking of them, are indeed the results of this piece of methodological research. A discussion and conclusion follow.



The original FEARLUS modelling system

- 2.1 This section describes the FEARLUS modelling system before changes were made from the qualitative research findings discussed in this paper. Entities in the model are indicated using a `fixed width font`.
- 2.2 A FEARLUS model consists of a raster `Environment of Land Cells`, each representing a square region of fixed area. A `Cell's` neighbourhood is determined by parameters to the model, which include Moore and von Neumann options (Figure 2). Each `Cell` has associated `Biophysical Characteristics`, representing local conditions affecting `Yield`. `Cells` are grouped into `Land Parcels`—each `Land Parcel` is a unit of decision and ownership, having one `Land Use`, and one `Land Manager`. One `Land Manager` may own several `Land Parcels` at a time. The neighbourhood of a `Land Parcel P` is the set of `Land Parcels` with `Cells` neighbouring `Cells` of `P`.

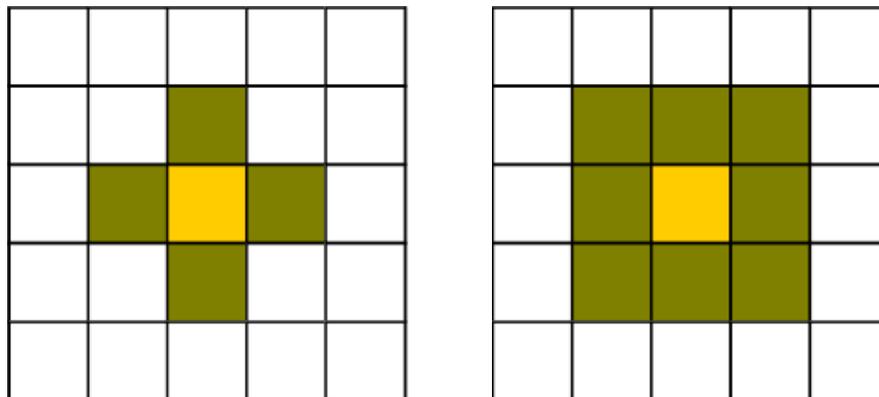


Figure 2. (a) Von Neumann and (b) Moore neighbourhoods of cells in a raster grid.

- 2.3 Each `Year`, `Land Managers` make `Land Use` decisions about the (one or more) `Parcels` they own, and experience the consequences of those decisions (Figure 3). There are four consequences. First, a `Government` agent may issue rewards or fines to `Land Managers` according to its `Policy`. Second, the `Land Managers` accumulate money to their `Account` by harvesting the `Yield` from their `Land Use` decisions. The `Yield` is influenced by the `Biophysical Characteristics` of the `Land Cells` of the `Land Parcel` (which are constant over time), and a `Climate` that changes over time, but not over space. The `Economic Return` is influenced by the `Yield` and the state of the `Economy`, which like the `Climate` changes over time but not over space. (Figure 4 shows the layers of spatial representation in FEARLUS.) The `Climate`, `Economy` and `Biophysical Characteristics` are all exogenous driving variables of the model. The amount accumulated to the `Account` of a `Land Manager` is the total `Economic Return` from all their `Parcels`, less a `Break Even Threshold` per unit area, to account for costs of running the farm. Third, some classes of `Land Manager` may `Approve` or `Disapprove` of each other. Finally, `Land Managers` with less than zero `Account` are declared bankrupt, and sell all their `Parcels`, which are then bought by neighbours or by in-migrant `Managers` in an auction. A detailed discussion of the system for exchanging `Land Parcels` in the `Land Market` is given by Polhill, Parker and Gotts (2005; 2008).

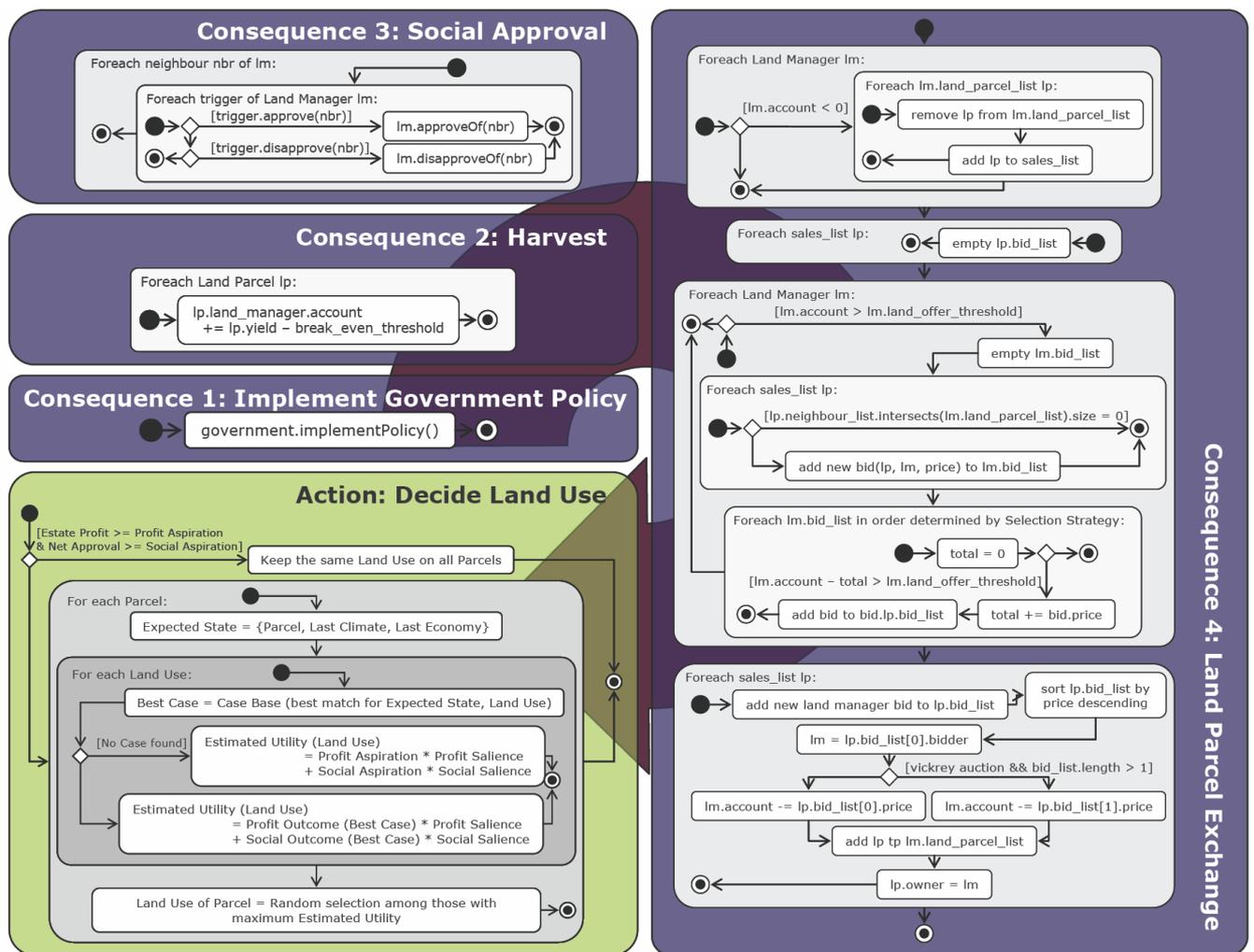


Figure 3. Loop of a single Year in a FEARLUS schedule, characterised as a Land Use decision and subsequent action carried out by each Land Manager, followed by a series of consequences. UML activity diagrams (Booch, Rumbaugh and Jacobson 1999) in each action/consequence box give some details on the processes undertaken at each step.

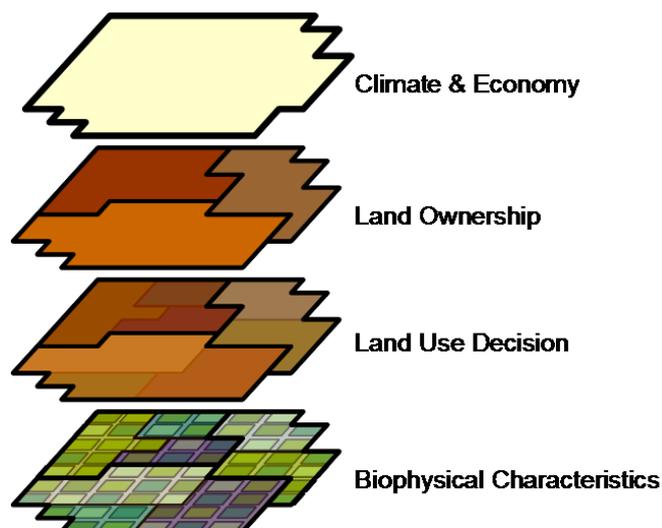


Figure 4. Layers of spatial representation in FEARLUS.

2.4 **Yield** and **Economic Return** are both calculated using lookup tables. The **Yield** lookup table takes the **Climate**, **Biophysical Characteristics** and **Land Use** as input, and returns the **Yield** per unit area as output. The **Economic Return** lookup table takes the **Land Use** and **Economy** as input, and returns the **Income** per unit **Yield** as output. FEARLUS offers considerable flexibility in configuring these lookup tables. Each of the **Land Use**, **Biophysical Characteristics**, **Climate** and **Economy** can be described using an arbitrary number of properties, each property having an arbitrary number of symbols. For example, a **Land Use** might have a 'commodity' property and a 'management strategy' property, with the commodity property having symbolic values 'Wheat', 'Carrots', 'Turnips', 'Beef' etc. and the management strategy property having symbolic values 'Organic', 'Conventional' and 'TransitionToOrganic'. In another scenario, the **Land Use** could be described using different properties and symbols, however. Similarly, the **Climate** could be described using properties such as 'rainfall', 'temperature' and 'variability' in one scenario, or 'length of growing season' and 'length of harvest season' in another. This flexibility in describing the determinants of **Yield** and **Income** potentially enables FEARLUS to be fitted to the requirements of a range of different scenarios. The disadvantage is that outputs must be supplied in table form for all combinations of inputs of **Biophysical Characteristics**, **Land Use** and **Climate** for **Yield**, and **Economy** and **Land Use** for **Income**. An example lookup table is shown in Table 1.

Table 1: Example lookup table for **Yield**, in a case study for which **Climate** is described in a single property (also called 'Climate'), with values 'Bad', 'Medium' and 'Good'; **Biophysical Characteristics** are described in a single property 'GrazingQuality', with values 'Poor', 'Rough' and 'Improved'; and **Land Uses** are described using two properties: 'Intensity', with values 'VeryLow', 'Low', 'High' and 'VeryHigh', and 'BoughtInFeedSupplements', with values 'None' and 'Lots'. The value in the rightmost column is the consequent **Yield** per unit area, and there is one row for each of the $3 \times 3 \times 4 \times 2 = 72$ possible combinations of property value symbols. (Not all rows are shown here—ellipses are used to show one or more skipped rows.)

Climate	GrazingQuality	Intensity	BoughtInFeedSupplements	Yield
Bad	Poor	VeryLow	None	32
Medium	Poor	VeryLow	None	40
Good	Poor	VeryLow	None	40
Bad	Rough	VeryLow	None	32
Medium	Rough	VeryLow	None	40
Good	Rough	VeryLow	None	40
Bad	Improved	VeryLow	None	32
Medium	Improved	VeryLow	None	40
Good	Improved	VeryLow	None	40
Bad	Poor	Low	None	32
Medium	Poor	Low	None	40
Good	Poor	Low	None	40
Bad	Rough	Low	None	56
Medium	Rough	Low	None	80
Good	Rough	Low	None	80
...
Bad	Poor	High	None	32
Medium	Poor	High	None	40
Good	Poor	High	None	40
...
Good	Improved	VeryHigh	None	160
Bad	Poor	VeryLow	Lots	40
Medium	Poor	VeryLow	Lots	50

Good ...	Poor ...	VeryLow ...	Lots ...	50 ...
Bad	Improved	VeryHigh	Lots	100
Medium	Improved	VeryHigh	Lots	150
Good	Improved	VeryHigh	Lots	200

- 2.5 Land Managers each belong to one of possibly several Subpopulations, the latter determining parameters for the Land Managers' decision-making algorithms. (Members of Subpopulations are all Land Managers.) Many of these parameters are distributions, allowing Land Managers to be configured as individuals different from others in the same Subpopulation. A FEARLUS model can consist of one of two classes of Subpopulation. One class is for configuring Land Managers using Case-Based Reasoning (CBR; Aamodt and Plaza 1994) to choose Land Uses for the Land Parcels, another is for configuring Land Managers using combinations of heuristic strategies for this task. Land Managers belonging to this class of Subpopulation have been discussed at length in existing published work with FEARLUS, and will not be described further here.
- 2.6 Land Managers belonging to the CBR class of Subpopulation implement a simple social approval model. As well as having an Aspiration for Profit, they have an Aspiration for Approval—recognition by the farming community that they are a 'good farmer' (Burton 2004). They also have rules determining whether or not they approve or disapprove of their neighbours; these include the ability to approve or disapprove of certain Land Uses, and to disapprove of those using more polluting Land Uses than they are. Land Managers count the number of times they have been approved or disapproved of, and their Net Approval: the number of times they have been approved minus the number of times they have been disapproved. Two classes of CBR Land Manager implement different conditions for reviewing Land Uses on the farm. That shown in the bottom left of Figure 3 is the simpler, and is designed for runs where Land Managers only ever own one Land Parcel. A more complicated algorithm is used in runs where Land Managers may own more Parcels: If the Net Approval is less than their Approval Aspiration, then Land Managers will review the Land Uses of all Land Parcels they own using CBR. If the Approval Aspiration is met, then the Managers consider each Parcel in turn. If the Profit on that Land Parcel meets the Profit Aspiration, then the Land Use is not reviewed unless the Land Manager has just bought the Parcel; if it is less than the Profit Aspiration, then the Land Use of that Land Parcel is reviewed using CBR.
- 2.7 Land Managers using CBR have an Episodic Memory recording their Experience of Land Uses they have tried. Each Experience records the Land Parcel, Climate and Economy, the Land Use tried, and the amount of Profit made and Net Approval received. When reviewing the Land Use of a Parcel, a Manager generates an expected Climate and Economy (by assuming no change from those the previous Year), and searches the Episodic Memory for each Land Use to obtain the expected outcome (Profit and Approval) were they to use that Land Use on the Parcel in question. If they have no experience of a Land Use, they expect that the outcome will just meet their Profit and Approval Aspirations. The expected Utility of each Land Use is then computed by weighting the expected Profit and expected Approval by a contextually varying Salience. This determines the relative importance to the Manager of Profit and Approval, and may be changed by a set of rules. Examples of such rules are: 'increase the Salience of Profit if the Yield is less than a threshold'; 'increase the Salience of Approval if Net Approval is less than a threshold'; and 'decrease the Salience of Profit if the Account is more than a threshold'. The Land Use (which may be the same as that applied previously) is then selected at random from those with equal maximum expected Utility. If all known Land Uses are expected not to meet Approval and Profit Aspirations, the algorithm will select at random among Land Uses not so far tried.
- 2.10 Each Land Use may have associated with it an amount of Pollution it generates on the Land Cells of the Parcels to which it is applied. Two classes of Government agent are available, each creating a social dilemma if it is assumed that Land Uses generating more wealth tend to be more polluting. Both classes of Government compute the total Pollution in all Cells of the Environment, and issue a Reward to all Managers if this Pollution is less than a threshold. In one class, the Reward is issued per Farm (so Managers each get the same amount regardless of the number of Parcels they own), in the other it is issued per Parcel.
- 2.11 Figure 5 shows a screenshot from a model run based on a configuration derived from data pertaining to the Grampian case study.

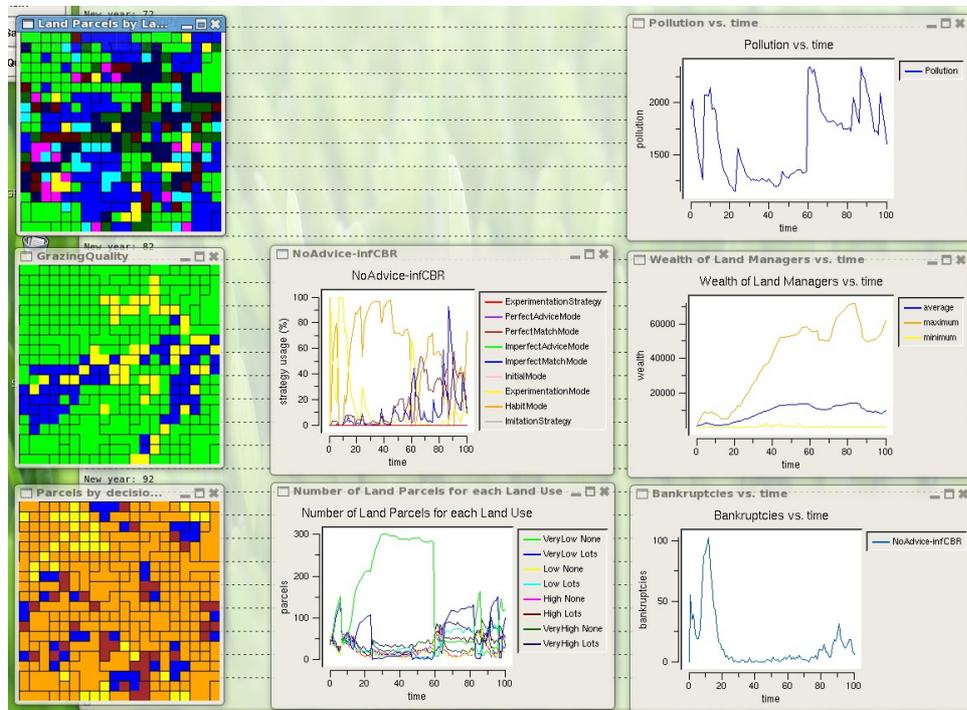


Figure 5. Screenshot from a model run. On the left hand side, starting at the top, the rasters show the spatial distribution of the most recent Land Use, Biophysical Characteristics (grazing quality in this case), and mode of decision-making. Clockwise from the top right, time-series graphs show Pollution, distribution of the Accounts of Land Managers, bankruptcies, Land Use history, and decision-making mode history.

Methodology

- 3.1 The purpose of the Grampian case study is to draw on field research to inform and develop the existing FEARLUS modelling system, with a view to generating scenarios of possible future land use patterns that would be used to study complexity with an evidence base. Qualitative research was conducted in two stages: first to gather the evidence used to inform a series of refinements to the pre-existing FEARLUS modelling system to create a framework capable of modelling the CAVES Grampian scenarios with an acceptable level of detail; second, to check the assumptions behind the ontological changes made to FEARLUS with the respondents. The qualitative findings themselves were validated separately through rigorous application and documentation of methods (see Silverman 2001). Overall, the method takes an iterative approach, in which questions to be addressed in qualitative interviews are derived from issues arising from model development, and changes to the model are suggested by findings from qualitative interviews. Figure 6 summarises the process.

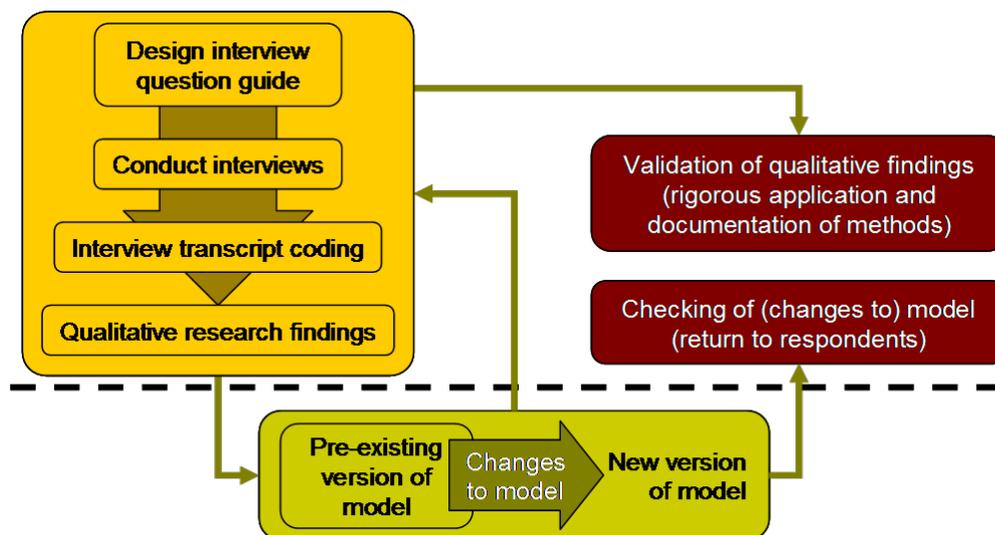


Figure 6. Summary of methodology. The dashed line represents the boundary between the model and the real world.

Collecting qualitative data

- 3.2 Methodologies in qualitative research are far from being ad hoc, so those gathering evidence and analysing it in such an exercise should be trained in qualitative research methods to ensure they adhere to recognised standards of research, and understand how to interpret their findings rigorously. The quality of results depends on a well informed research methodology, as is characteristic of any research process. Rubin and Rubin (1995, p.2) raise the following questions to illustrate the meaning of an informed methodology:

How do you think of questions for a topic that you have chosen? How do you get people to stay focused on what you want to hear about? Whom do you interview and why? How can you trust what people are telling you? On a more technical level, how do you persuade a person to become an interviewee? How specific should a question be? Is the wording too biased? How do you get people to elaborate on what they say? How do you put together different tellings of the same event? When do you take on-scene notes and why should you rely on memory? Do you use a tape recorder or a camcorder?

- 3.3 Numerous books have been written on the topic of qualitative research (see in particular Mason 2002; Rubin and Rubin 1995; Silverman 2001); only a few basic principles will be included here. The first step in any research process is to identify the research questions, and the most appropriate means for addressing them. In this study, the research questions were very general: nature and causes of land use change in the study site, and role of social networks in these. Qualitative interviews were chosen in order to facilitate identification of a wide range of influences on land use, and exploration of social dynamics in relation to land use change.
- 3.4 This study focused on interviews from (i) key informants^[2], (ii) active farmers, and (iii) their identified successors. In quantitative work, the sampling frame would be defined statistically, in order to determine representativeness, and therefore validity of results. For qualitative work, validity and reliability need to be achieved through different means, in this case the use of these three 'types' of interviewee, in order to allow comparison of responses by interviewees with different perspectives ('triangulation'). In general qualitative research methodologies seek as wide a diversity of respondents as possible; triangulation confirms a finding through different types of respondent making much the same point.
- 3.5 It was particularly important to include key informants, as literature had already demonstrated that farmers may deny social influences due to cultural norms of independence, even when these social norms are clearly evident to industry experts (Burton 2004). Due to the small number of respondents typically drawn on in qualitative work, respondents are often selected initially on the basis of recommendation: an individual in the study site who is familiar with most possible respondents and can therefore identify which are most likely to have experience with the study topics and represent different viewpoints, and perhaps more importantly, be willing to discuss these^[3]. From the initial recommendations, further candidates are identified by the interview respondents themselves, a process referred to as 'snow-balling' (see also Burton et al 2008)^[4]. To ensure that interviewees did not represent a single farming type or social circle, initial interviewees were recommended by two key informants with very different connections in the study site: a Scottish National Farmers' Union representative and a Farming and Wildlife Advisory Group representative. Interviewing continues until a point of 'theoretical saturation' is reached: no new insights are being revealed through the interview process (Bryman 2008). In this research, interviewees included 24 farmers and 6 of these farmers' successors. Eight key informants from a range of farming service providers (e.g. accountant, bank manager, agricultural advisor) were also interviewed. The interviews were recorded^[5] and transcribed either in full or in note form, depending on the length of the interview, which ranged from 45 minutes to 2 hours. Interviews were conducted in the Upper Deeside region (Finzean to Braemar—see figure 1) of north eastern Scotland. Pilot interviews were conducted from February–April 2006, with primary field research in June–July 2006 and from November 2006 to March 2007. This reflects several phases of research, which were scheduled into periods of high farmer availability, and allowed for a process of discussion and further questioning with the modelling team.
- 3.6 Qualitative research is typically iterative, meaning that research questions evolve over the course of the research project, as new findings emerge (Rubin and Rubin 1995). There must be flexibility in the interview topics and format, to ensure detailed exploration of the range of topics and processes raised by respondents. In this study a question guide was used to structure the interviews. The final question guide can be found in Appendix 1. This guide identified basic questions, which were then followed up with further open-ended probing of topics raised by the interviewees.
- 3.7 The guide was formulated in conjunction with the agent-based modelling team in late 2005, and evolved over the course of the interviews. Mason (2002 p. 69-72) sets out the basic approach adopted in this study. Major topics for exploration (in this case, causes of land use change and patterns of reciprocity) were broken down into questions addressing specific issues of interest, largely based on academic literature. A structure for the interview was developed, designed to put the respondent at ease (beginning with 'easy questions' about demographics and change on the farm, and progressing to less concrete discussion of informal exchanges and interactions).
- 3.8 The interview itself is flexible: the interviewee may raise issues spontaneously, and it is up to the interviewer to decide whether to explore the issue at the time it is raised, or to return to it later in the interview. Specific methods to expand (through open ended questioning, asking for specific examples of phenomena mentioned by the interviewer, or suggesting possible alternatives to the scenario presented) or limit (through closed ended questions, redirection or body language) the interview are used, as appropriate to the talkativeness and focus of individual respondents. As a consequence, it is inevitable that no two interviews will contain exactly the same questions or cover precisely the same ground.
- 3.9 It is basic research ethics to explain the recording process and gain the interviewee's consent prior to the interview beginning. Most processes include a confidentiality agreement, in which the interviewer and interviewee agree on who may access the transcribed information, and what purposes it can be used for. In this case, participants were informed that the transcriptions were accessible solely to the research team, and limited to academic use.
- 3.10 The analysis process involved comparing responses from different interviewees, and interviewee types, looking for the range of issues identified, underlying concepts, evident responses to the research questions, any new questions or topics of interest which arise, and links between different ideas. Use of data analysis software to facilitate this process is now standard practice, and in this case NVIVO was used. NVIVO software enables typed transcripts of interviews (or indeed any text or image) to be 'coded'—sections of text are assigned to topical 'nodes', such that all of the sections of text assigned to the node can be viewed as a group. For example, a node entitled 'decision-making' would contain all of the interviewees' comments relevant to decision-making, making it easier to review these together. Information from nodes can then be organised into a matrix, with columns corresponding to several nodes of interest, and rows to respondents. The analyst can then easily compare individual responses across the selected topics. This may be done several times to look for different relationships, or different questions. It is important to note that the purpose of the exercise is not to determine how many respondents have the same responses, but to look for links between similar responses, and the range of issues identified by respondents who appear similar in some respects. This process is typically used to provide foundations for a sociological paper (for an example based on the data used in this process, see Sutherland (2010)). Integrating qualitative findings into the model is further discussed in the next section.

Changing the modelling system

- 3.11 Findings from the qualitative evidence-gathering exercise were used to create a discussion document outlining a set of proposed changes to the ontology of the FEARLUS modelling system and how they might be implemented algorithmically. This was used in deciding which changes would actually be made, based on the importance of the change, the amount of effort required to make it, and the time available.
- 3.12 The field researcher and model developers also met regularly to discuss findings and their potential implications for the model structure. Sometimes, these discussions led to different questions being asked in follow-up interviews, completing the iterative loop of integrated evidence gathering and model development. For example, the slow farmer response to changes in profitability of specific commodities was identified during initial interviews. Discussion with modellers raised the question of parameterisation: how much time lag could be expected in specific circumstances. Subsequent interviews identified the length of time it would realistically take for a farmer to change commodities, and how this varied between commodities. This yielded the finding that establishment of a beef or dairy herd would take five years, while changing a field crop could be achieved in one.
- 3.13 For various reasons (including the use of FEARLUS in other work (e.g. Gotts and Polhill 2007; Polhill et al. 2008), it was desirable to retain backward compatibility of the FEARLUS modelling software. This also meant that if a change was not supported when checked with respondents, the associated functionality could be switched off.
- 3.14 It is important to note that there is no established methodology for implementing qualitative interview findings in agent-based (or indeed other) models. The approach taken here was to consider options for the implementation of each finding, with a tendency to prefer simpler options. Such a prejudice has been criticised (Edmonds 2007), however, our concern was primarily with whether the phenomenon occurred in the model—we were unable to justify more elaborate implementations from the qualitative research findings we had.

Checking the changes

- 3.15 There is much dispute over the degree to which measures of reliability and validity—the hallmarks of quantitative research—can be applied to qualitative work. Two of the most common approaches for qualitative research are triangulation and respondent validation. Triangulation involves the use of multiple sources of data to corroborate data or findings (Bryman 2008). Respondent validation involves returning the research to respondents to validate either the original transcripts, the researcher's analysis of transcripts, or sections of resultant academic texts (Bryman 2008). Silverman (2001) is critical of these approaches, arguing that both involve removing the original data from its original or intended context, and thus losing some of its value. Silverman (2001) argues instead for rigorous application and documentation of methods—a process that would not be regarded as 'validation' by those more familiar with a quantitative understanding of the term (Rykiel Jr. 1996). This approach was used to validate the qualitative study findings. Triangulation was also undertaken through the use of key informants, farmers and successors, as well as comparison to similar academic research. However, we recognised that the translation from qualitative research finding to modelling principle could result in a loss of accuracy. Thus, we decided that checking the changes by returning the assumptions behind them to the respondents would most appropriately demonstrate the credibility of new model components.
- 3.16 To implement the return of findings to the study respondents, a questionnaire was derived from model components and assumptions. The opportunity was also taken to check some pre-existing processes within FEARLUS, such as the range of factors included in calculating farm profitability. However, consistent with the TAPAS approach, the emphasis was on demonstrating improvements to the model's ontology arising from changes derived from the field research. The questionnaire covered the comprehensiveness and relative importance of the proposed factors influencing land use change, and the decision-making process. Since the purpose of the exercise was to check (changes to) the structure rather than the behaviour of the model, outputs from the model were not given to respondents for validation; the focus was instead on processes within the model, and, to a lesser extent, input to it, such as the classification system used for agricultural land. This is broadly consistent with the first half of Moss and Edmonds's (2005) 'cross-validation' procedure: checking the rules of behaviour at the micro level qualitatively, whilst using statistical methods to validate the emergent behaviour at the macro level.
- 3.17 In this checking phase, questionnaires were completed with eight respondents from the primary field research (five farmers and three key informants), between April and July 2007. The farmers were selected to represent a range of tenures (tenanted and owner operated), ages (41-70), and approaches to farming (traditional and farm expansion-oriented) which were most common among study respondents. These also correspond to the 'conservatives' and 'accumulators' identified in Shucksmith's (1993) study of farmers in Grampian region, in which the study site was located. The key informants were also chosen to reflect different perspectives. These individuals were a farm business advisor, an environmental programme advisor, and a representative from the Scottish National Farmers' Union. No factors (estate managers) or farm successors were drawn on for checking, as the variety of their responses in the primary field research suggested that a single representative of either group would be unlikely to typify the perspective of the whole. The questionnaire involved a variety of approaches, including open-ended questions, ranking of identified factors, and weighting of responses.



Results

- 4.1 The results of the exercise are presented in three tables, with a view to documenting more specifically how findings, model implementation and checked statements are related than is current practice in modelling literature. Table 2 summarises the findings from the primary field research as they relate to changes to the model, with a column reserved for indicating how the changes were checked. In some cases it was possible to accommodate the finding by changing model parameters, rather than model structure. "Checked statements" represent statements to which respondents were asked to 'strongly agree' (SA), 'generally agree' (GA), 'disagree' (D), or 'strongly disagree' (SD), which is shown in the "Response" column. Table 3 outlines changes made to the model related to findings that were not checked with respondents through the questionnaire for one of three reasons: the finding is relatively obvious, the change to the model is difficult to check, or support for it can be obtained in the literature. Finally, Table 4 details findings that did not result in a change to the model, and were not already implemented.
- 4.2 The following summarises the changes made to FEARLUS:
- Farm-Scale Fixed Costs added to the economic model.

- Enabling Land Manager agents with case-based reasoning to also use heuristic imitative and innovative decision-making algorithms.
- Land Manager agents wait for a number of consecutive years of unsatisfactory performance with a Land Use before making changes.
- Off-Farm Income added to the economic model.
- Land Manager agents may seek 'advice' from neighbours if their case base has no example of a Land Use matching their expected situation.
- The condition for meeting Social Aspiration is based on the mean Net Approval from those Managers not disapproved of only.
- Land Managers have a small probability of selling up even if they are not bankrupt.
- Time and size limits are imposed on the cases bases of Land Manager agents.

4.3 A full account of the functionality available in the most recent version of FEARLUS, incorporating all the changes outlined here, can be obtained from the user guide (Polhill, Izquierdo and Gotts 2008).

Table 2: Results from the qualitative research, the consequent changes to the model, the statement validated and the degree of agreement among the respondents.

Finding Description	Implementation in model	Checked statements	Response
In ordering the factors which farmers take into consideration when changing the profitability of the new commodity is of primary importance	Already in model: parameterisable using the <i>Saliency of Profit</i> over <i>Social Approval</i>	"Please order the following one to ten, where one is the most important reason for changing to a new commodity, and ten is the least important."	8/8 (GA or SA): Profitability was consistently ranked first.
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	Already in model: <i>Aspiration Thresholds</i>	"Farmers do not change a current crop or type of stock if they are satisfied with the profit margin."	8/8 (GA or SA): Unanimous support.
Economies of scale: The larger the farm, the more they can produce with the same equipment.	New: This is implemented as a parameter (<i>Farm-Scale Fixed Costs</i>), which is subtracted from all <i>Land Managers</i> when calculating the <i>Economic Return</i> . As an option, the parameter can be given as a time-series, loaded from a file. Original behaviour is enabled by setting this parameter to 0.	"The larger the amount of a commodity you produce, the cheaper it is to produce each unit."	5/8 (GA or SA): Clarification that economies of scale can be limited by other aspects of the farm business such as land, labour and capital availability.
Fixed costs per land use: There has been a marked decrease in the number of commodities produced on individual farms over the past two decades. This appears to be due to the fixed costs of producing each individual commodity.	New: Fixed costs per <i>Land Use</i> are representable using economies of scale, implemented thus: Each <i>Land Use</i> is given three parameters, used to derive a co-efficient <i>c</i> in the range [0, 1] by which to multiply the break-even threshold per unit area for that <i>Land Use</i> . When <i>c</i> is 0, there are no costs for the <i>Land Use</i> , and when <i>c</i> is 1, there is no economy of scale. The three parameters specify the	"There are fixed costs associated with each crop and land use adopted in a farming operation. The more different commodities produced, the higher these are."	6/8 (GA or SA): While this statement is generally true, having more than one commodity can spread resources like labour over the calendar year. Thus while the fixed costs of five commodities is higher than two, having two commodities can be more efficient than having only one.

area below which c is 1, the minimum value of c , and the area above which c takes the minimum value; between the two areas, c descends linearly from 1 to its minimum value. Original behaviour is enabled by setting the minimum value of c to 1.

Some farmers are more innovative than others; new commodities or technologies are introduced by a subset of the total population.

New: Facility for CBR Land Managers to use imitative and innovative strategies. Each Land Manager is given a probability, taken from a distribution at Subpopulation level, that when deciding the Land Use of a Land Parcel, they will use an innovative or imitative strategy rather than consulting their Case Base. Original behaviour is enabled by setting this parameter to 0 in all Subpopulations.

"New crops or types of stock are introduced by a small number of highly innovative farmers."

7/8 (GA or SA) (to both): One farmer clarified that many farmers are innovators in some way.

"Other farmers watch to see how successful these innovations are, before adopting the innovation themselves."

Due to the high opportunity cost of changing commodities, farmers require several years of poor returns before changing commodities.

New: Land Managers wait for a number of consecutive Years of failing to meet Aspirations before reviewing Land Use on the Farm. Each Land Manager is given a number of consecutive Years, d , taken from a uniform distribution at Subpopulation level, and keeps track of the number of consecutive Years, y , for which Aspirations have not been met. If $y = d$, then the Manager reviews the Land Use on all Parcels they own. Profit Aspiration is met if the mean Profit per unit Area meets a threshold (obtained per Manager from a distribution at Subpopulation level). Social Approval Aspiration is discussed in Table 3. This new condition for reviewing

"Farmers will accept financial losses for several years before changing their crop or type of stock."

6/8 and 7/8 (GA or SA) respectively:
The same farmer respondent disagreed with both statements, arguing that farms could not afford to lose money indefinitely; he agreed in principle that there is a delay between making a loss and changing commodities. The issue is therefore about parameterisation of time and amounts of loss, rather than the principle itself.

"Farmers will continue producing their current crop or type of stock at a small loss, if they see no viable alternative except leaving the farm."

	Land Uses was implemented using subclasses, allowing original behaviour to be used if required.		
Off-farm income can offset losses from farming.	New: Each Land Manager has a normal distribution of Off-Farm Income, a sample from which is added to their net Profit each Year. The parameters for the normal distribution (mean and variance) are each taken from a uniform distribution at Subpopulation level. Original behaviour is obtained by setting the minimum and maximum mean and variance to 0 in the Subpopulation parameters.	"Off farm income allows farms to stay viable without generating a profit."	8/8 (GA or SA): Unanimous support.
Until the present time, farmers would bid on neighbouring land if they have sufficient resources. This is less likely to be true in future, due to changes in the Single Farm Payment (a government subsidy). Farmer types are entrepreneurial, traditional, pluriactive and lifestyle/hobby/environmental	Already in model: Land Managers currently only bid on (socially or physically) neighbouring Parcels.	"Farmers will always try to buy neighbouring land when it comes up for sale, unless they are planning to retire or leave farming."	5/8 (GA or SA): Consistent with original finding, respondents indicated this was historically true, but changes to subsidies now make them reconsider.
Categorisation of arable land into 'arable', 'arable grass', 'woodland', 'improved grazing' and 'rough or hill ground'.	Already in model: parameterisable using lookup tables	"Do these types seem reasonable to you?" "What other types of farmers might exist?"	3/3(GA or SA): Agreed. Key informant validation only: all agreed with the types, no further types identified.
Climate change with differential impact by commodity and land use type	Already in model: parameterisable.	"What would you say is the best way to categorise agricultural land, in terms of what it can be used for?"	2/5 (GA or SA): Farmer validation only. The other farmers suggested variations. From this feedback, we surmise that we need not distinguish between grades of arable land in the study site, and that woodland is an appropriate category.
There is social pressure against certain land uses	Already in model: parameterisable in Social Approval ratings.	"How big a climate change would have to occur for the amount of production of your commodities to change? Would this be an increase or decrease in production?"	Varying responses from the five farmers.
Organic farming-social	Already in model:	Key informant ranking of peer approval associated with specific activities.	Inconsistent response, except in relation to organic farming.
	Already in model:	"One reason farmers don't	4/8 (GA or SA): Suitability

approval may limit adoption, but land suitability and perceived economic benefits appear more important	parameterisable in Social Approval rating.	consider adopting organic farming is because they are concerned about what the neighbours would say."	of land and profitability were considered more important by several. However, organic farming was ranked as unpopular with farming peers in the key informant questioning. 6/8 (GA or SA): True in most cases, but innovative farmers may try something new instead.
Changing commodities: farmers most likely to expand current production to increase economies of scale, rather than trying a new commodity.	Already in model: The case-based reasoner will select Land Use, Manager has experience of where these have a higher expected Utility than any untried Land Use.	"When discontinuing a crop or type of stock, farmers are most likely to expand production of another commodity they already produce, rather than trying a new commodity."	

Table 3: Findings and changes to the model that were not checked with respondents, with the reason for not doing so.

Finding	Change to the model	Reason for not checking
Farmers seek advice from 'good farmers', not any neighbouring farmer.	If, when reviewing Land Uses, a Land Manager does not remember a case where they employed a particular Land Use, they may ask neighbouring Land Managers for a case. Land Managers will not ask other Managers they disapprove of for advice, nor give advice to Managers they disapprove of. The first case provided by a neighbour is used to inform the Land Manager's decision. Land Managers have an Advice Strategy, derived from their Subpopulation, which is an algorithm to sort neighbouring Managers in the order they will be asked; e.g. in descending order of the amount of Profit they made last Year. To enable reproduction of the original FEARLUS behaviour, an advice strategy returning an empty list is available, which prevents Managers having any advice to use.	Supported through literature review: finding from Burton (2004). Key informants support this contention, but farmers are not likely to support any finding that they seek advice from peers, due to high cultural levels of independence.
Farmers care only about the good opinion of those they regard as peers.	Social Approval Aspiration is met if the mean Net Approval from Managers not disapproved of meets a threshold obtained per Manager from a distribution at Subpopulation level. The mean Net Approval operates over the set A of neighbouring Managers that the Manager has not disapproved of, and is calculated as the amount of Approval received from A minus the amount of Disapproval received from A,	Social approval is a controversial issue among farmers, for reasons above. See also Burton (2004).

divided by the cardinality of *A*. This was implemented by subclassing from the original CBR class of *Subpopulation*, so original behaviour can be obtained by invoking the model using the original class rather than the subclass.

Farmers do not have to go bankrupt to sell their farms.	Each <i>Land Manager</i> has a probability <i>s</i> of selling up in any given <i>Year</i> , even if not bankrupt. If a random number in the range [0, 1] is less than <i>s</i> , the <i>Manager</i> sells all their <i>Parcels</i> . The probability is obtained from a distribution at <i>Subpopulation</i> level. The original behaviour is obtained by setting this distribution so that it always returns 0.	Relatively obvious, although 'keeping the name on the farm' is a significant motivating factor in staying in farming even when business logic might dictate otherwise.
Farmers do not have an infinite knowledge of other farmers' experiences. Neither are specific personal experiences relevant for decades, as commodity markets change.	Limited <i>Case Base</i> size. Limits on both the number of cases and the amount of time for which a case is remembered can be set for a <i>Land Manager</i> . The <i>Case Base</i> size and time limit are taken from uniform distributions at <i>Subpopulation</i> level.	Obvious.

Table 4: Findings from the qualitative field research that did not lead to changes in the model, with explanations.

Finding	Reason for not implementing
Approximately half of the land in the study site is tenanted (rented on a long-term basis) rather than owned.	Tenancing would involve major changes to the land market model in FEARLUS, and there was insufficient time to make these changes within the CAVES project timescales.
There have been very few new entrants to farming in the case study area. It would therefore make sense not to include a "new entrant" in the competition for land parcels being reallocated. This is turn would require that land abandonment be possible, if none of the neighbouring land managers wants control.	Handling land abandonment properly would involve significant changes to the way land uses are represented in FEARLUS. Further, one possible outcome of a shock to the system could be an influx of land managers with different values, priorities and behaviours.
The use of sources of information about land uses other than neighbours.	There were too many sources of information identified by respondents, which varied considerably between respondents, and were accessible to widely varying geographic areas. Further study would be needed to identify principles of information access which could be included in the model.
Farmer ability impacts considerably on farm profitability.	Some aspects of differences that could indirectly reflect ability among farmers are already represented through other changes, such as limits to case storage. We believed the descriptive gain relative to performance in the model would be insufficient for it to be worth implementing.
Including more specific detail on proximity of land and expansion plans to decisions regarding land acquisition.	This would involve representing land managers' plans at a more detailed level of abstraction than FEARLUS is currently designed for.

Including 'well established market for new commodity', 'suitability to current farm set-up' and 'opportunity to benefit from government grants' as factors in decision-making about new commodity uptake.

This would involve significant changes to the algorithm for experimentation during Land Use decision-making. Farm set-up would need to be explicitly represented, and a more precise definition of 'well-established' would be needed. The effort required would be too great given the time available in the CAVES project, relative to the gain in functionality.

Succession is a concern on a number of farms. As labour on farms decreased with mechanisation and profits from farming declined, farmers' children sought employment elsewhere. When the current farmer retires, therefore, there is no guarantee that any of their children will take over.

Succession would require representing farm households in more detail than the current level. Changes made to allow a small probability of a Land Manager selling up could substitute for this to a certain extent.

- 4.4 Most of the checked statements from Table 2 were agreed by the majority of those asked, though given the small sample size, a statistically valid quantitative assessment of the response is not possible. There was strong support for profitability as the prime motivation for changing commodity, and for satisfaction with profit margin as a motivation for not changing—both of which are parameterisable features of the original FEARLUS model. There was strong support for introducing off-farm income to the model. There was also strong support, with clarification, for innovations being introduced by a relatively small number of farms, others adopting after seeing their success; for farmers accepting financial loss with an activity for a period of time before making changes; and for farmers tending to expand their production of existing commodities rather than trying something new if dissatisfied with an activity—all of which are features arising from changes made to the model. Weaker support, with clarification, was given to the introduction of farm-scale fixed costs to the model, and for the existing feature that farmers try to buy neighbouring land when it comes up for sale.



Discussion

- 5.1 The results have shown that we can have confidence in using several of the features in the new FEARLUS model. Support has been given to both pre-existing and added features, and many findings could be implemented using appropriate parameter settings, suggesting that there has been a benefit to adopting the TAPAS approach. Areas of equivocal support in Table 2 occurred around the question of social approval, specifically in relation to organic farming. In general, as noted in Table 3 in relation to a change to the social approval model, this aspect of influence on farm behaviour is contentious as farmers have an independent self-image. Whilst key informants and the literature corroborated that farmers are influenced by norms in their peer group, the results of the checking process suggest that the social approval feature in FEARLUS should be used with caution in the Upper Deeside case study. Two of the questions, focused on parameterisation of the model, also produced no clear agreement: a suggested categorisation of agricultural land, and the possible influence of climate change on farm production.
- 5.2 Not all findings suggesting changes in the model were implemented. The reasons given above in Table 4 are akin to the 'for the sake of simplicity' assumptions, which Moss (1999; 2002) and other agent-based modellers (e.g. Edmonds and Moss 2005) have rightly criticised. Ideally, in those instances where lack of resources prevented changes, a model would have been implemented with changes made according to the finding. However, all models are abstractions from reality; the iterative nature of the interaction between field research and model development^[6] suggests that with unbounded resources, modifications to the model might never stop. Development of FEARLUS has not stopped since it began in 1998, and will continue so long as there are scenarios to which FEARLUS can be applied and researchers interested in doing so. Where changes were not made because they would require too many alterations to the pre-existing code, this suggests that, had we started from scratch, the resulting model would have implemented the feature. With limited resources, the decision to build a new model is a trade-off between the importance of features too difficult to implement in an existing system, and time taken to reimplement features that are also required in the new system. In the case of land tenure, which we would have implemented given time, we did not believe the change to be important enough to abandon the TAPAS approach we had taken and build a new model. One way to approach this issue is to develop more radically modular architectures for social simulation—modularity is widely recognised as having an important role in simulation modelling (Frysjinger 2002; Leavesley et al. 2002)—and we intend to take this approach in future work (Polhill and Gotts 2009).
- 5.3 The process of translating qualitative findings into model behaviour is something of a dark art, as Agar's (2003) article attests. Papers describing the use of qualitative evidence to build an agent-based model tend not to go into much detail about how the qualitative knowledge was translated, other than to first describe the case study and then describe the model. Noting that modellers' intuition is involved in the design process, Taylor (2003, p. 140) nevertheless expresses a desire to be able to annotate the source code of his model with the coded interviews, to make explicit the provenance of the rules. This is something we have tried to address here using the tables above. We are also interested in methods of making the ontological structure of a model more transparent and verifiably consistent with descriptions of it, and this will form the subject of future work (Gotts and Polhill 2009a; Polhill and Gotts 2009).
- 5.4 Indeed, the topic can be expected to be contentious for many qualitative research specialists. Identifying different 'types' of respondents has a substantial history in rural sociology literature (see for example Shucksmith 2002; Bowler et al. 1996; Marsden et al. 1992), and several 'types' were identified in the interviews, which could then be integrated as different subsets of agents within the model. However, the use of algorithms to express qualitative findings implies a degree of certainty and precision that cannot always be achieved through qualitative research. Such uncertainty can be handled by making multiple implementations of the same phenomenon and/or using stochasticity, and then running the model large numbers of times. This has been an approach taken in work using FEARLUS. Another way to address the issue is to follow the qualitative research with more extensive quantitative research with a view to finding out the information needed—common practice in 'mixed methods' social research (Brannen 2005). However, the sheer number of phenomena to investigate and the level of detail and precision required could prove prohibitively expensive. Janssen and Ostrom (2006) cover a number of

challenges for empirically embedding agent-based models.

- 5.5 There is an established body of work using qualitative research in the agent-based modelling process. Here, there are a number of different approaches:
- Use of qualitative evidence for validating agent-based models. This is the approach recommended by Moss and Edmonds (2005) as part of their 'cross-validation' methodology, in which representations, structure and behaviour are validated qualitatively at the micro level, whilst outcomes are validated quantitatively at the macro level. The 'checking' process used here is broadly related to the former of these two activities.
 - Involving stakeholders in the development of a model. D'Aquino et al. (2003) apply the companion modelling approach to land use management issues in Senegal, for example. Ramanath and Gilbert (2004) review various approaches to participatory agent-based modelling in general.
 - Using qualitative representations in an agent-based model. Edmonds and Hales (2004) make this commitment in their model of haggling.
 - Using qualitative evidence to determine the ontology of an agent-based model. For example, Rouchier and Hales (2003) use interviews and field studies as a basis for building an agent-based model of a vegetable wholesale market in Marseille, Agar (2004) has built the Drugtalk model based on interviews with drug addicts, and Taylor (2003) bases a model of e-commerce on interviews.
 - Related to the above is the work done in this article: using qualitative evidence to suggest changes to the ontology of an agent-based model.
- 5.6 Qualitative research has also been used in agent-based modelling of land use change. Manson and Evans (2007) discuss two case studies (one of deforestation, the other of reforestation) in which 'mixed methods' research has been integrated with an agent-based modelling process. Huigen, Overmars and de Groot (2006) use collected oral histories as a basis for configuring their MameLuke model of settlement in the Philippines. Many of the case studies of the 'French school', as Moss (2008) calls proponents of companion modelling, are also related to land use change.
- 5.7 Becu et al. (2008) note that practitioners of companion modelling often start with a model constructed before meeting stakeholders. Here, qualitative research has been used to extend a pre-existing model that was not originally based on such research. In general, since qualitative data collection is iterative in nature (Rubin and Rubin 1995), it should fit well with the TAPAS approach of incremental model building. However, using the terminology of Galán et al. (2009), the danger of using a pre-existing model is that what were 'core' assumptions in the original model become 'accessory' in the new model. For this reason, it was undoubtedly a good idea to include some of the pre-existing features of FEARLUS in the checking exercise. Indeed, in future, it might be better to include all of the features provided by the modelling system, as the process can be used to suggest features that should be disabled.
- 5.8 It is also important that the principles implemented in the model are checked as well as the findings from the qualitative research. This allows evaluation of how far the process of implementing the findings in a model has lost their original accuracy (inevitably there will be some loss due to the abstraction process). Given that agent-based modelling is often predicated on building more realistic representations of social processes than is possible with analytical methods (Moss 1999), validating the *structure* of the model should have a priority at least as high as validating the fit of the model to quantitative data (Moss and Edmonds 2005). Indeed, a poor quantitative fit to empirical data may suggest a change to the structure or algorithms of a model's representation rather than adjustment of parameters.
- 5.9 Standard qualitative research relies on interviews with respondents directly involved in an activity, with supplemental information from key informants. In this case, we found key informants were often better able to give an overview of decision-making processes than the individual farmers themselves, who identified processes specific to their farms. This point relates to the acknowledged weakness of qualitative research, that it does not generalise well. However, rather than indicating that qualitative research may be inappropriate for agent-based modelling, this suggests to us that respondents need to be selected at a level of abstraction from the process that corresponds to the information required to structure the model. If the findings from the research process were intended solely for use in informing the model, data from key informants would have been sufficient. However, the dual use of the data for modelling and qualitative sociological analysis necessitated the inclusion of interviews with the study subjects.
- 5.10 Finally, while including a qualitative researcher in the team should ensure that high-quality data are collected, it raises other issues for the research process. One of the difficulties with interdisciplinary work of this nature is the need for practitioners to be able to communicate their work to colleagues in their own disciplines. This can lead to two results in terms of team work: the team is led by one discipline, and the others serve as 'assisting sciences', resulting in research outputs which are only publishable for the lead discipline; alternately, the researchers work in parallel, addressing the same topic but with different research questions, leading to findings which are difficult to integrate (Wächter 2003). This can create a dilemma in designing studies—if the qualitative field researcher only gathers information with a view to informing model development, the consequent findings may not be suitable for publication in qualitative research literature. However, if the researcher does gather material with a focus on their usual audience, there will be a lot of data that cannot be used by the modellers. The situation needs to be managed carefully to ensure all participants in the research team are making a worthwhile investment of their time.



Conclusion

- 6.1 This paper has shown how the qualitative research process is well-suited to incremental modelling approaches. We have argued that qualitative research offers a means of suggesting and checking the structure of a model, and specifically, changes to it, rather than comparing its statistical accordance with a set of quantitative empirical data, and that this is something of greater importance to agent-based modelling than to analytical approaches. Whilst it has been clear from the work of others that qualitative research with a deep involvement of stakeholders and practitioners in the domain of application is well-suited to agent-based modelling, it is clear that qualitative research with a shallower level of engagement can still be beneficial to the modelling process, in that we are able to evaluate how well features of an existing model apply to a case study, and make and check changes to that model designed to improve such applicability. Qualitative research thus has much to offer those considering empirical agent-based modelling exercises.

Appendix 1: Final question guide for qualitative interviews in the CAVES-Grampian Study

- 7.1 These questions were not necessarily asked in order, or in these particular words. Interviewees often answer questions before they are asked, in the context of other topics.
- 7.2 Introduction
- Review confidentiality of interview and gain permission for recording and transcription.
 - General demographics—size of farm, tenure arrangement, age of farmer, off farm income, age, education, household composition.
 - How long they've been farming, how many people work on the farm, what commodities they produce.
- 7.3 Any major changes in land use they have made;
- What were the changes (5 years, and 20 years); what did they involve?
 - What were the biggest causes of change?
 - Why those commodities particularly, those proportions?
 - What were the major considerations in making the decision: profitability, labour, investment, what they enjoy doing?
 - Have many of other farmers in the area made the same decision(s)?
 - Single Farm Payment—farm responses.
 - Land management contracts and other agri-environmental programmes—farm responses.
 - Planned changes—e.g. organic, diversify, intensify, extensify, niche...
 - Plans for succession.
- 7.4 Patterns of reciprocity
- Other people who work here—are they related to you? Where they get seasonal labour or weekend help?
 - General resource sharing—do they share any equipment, market with, how that got started, how common it is?
 - Machinery ring—involvement.
 - Group membership—e.g. belong to farming organisations.
 - Where do you get your information about farming?
 - Do you interact more or less with your neighbours than you used to?
 - What their neighbours do in terms of sharing resources, meeting together?
 - Types of farmers—what makes a good one; a bad one?
 - Importance of reputation?
 - Groups in the local area—farmers vs. hobby farmers vs. city commuters.
 - Would you say there are cliques—historical, in families, other?
- 7.5 General—major issues in agriculture today
- Why they became a farmer—is it turning out how they expected?
 - What's farming going to look like down the road?
 - Who handles farm sales; who puts up barns?
 - Any problems with the new 'Right to Roam' legislation? ^[7]
 - Recommendations of other people to interview, including key informants.
-

Acknowledgements

This work was jointly funded by the Scottish Government Rural and Environment Research and Analysis Directorate, and contract number 12816 under the EU Framework Programme 6 New and Emerging Science and Technology Pathfinder Initiative on Tackling Complexity in Science. The authors thank Innocent Bakam and three anonymous referees for some helpful comments on earlier drafts.

Notes

¹ It should not be inferred from this statement that qualitative research is always the cheaper option...

² Key informants are individuals working in the farming sector who are not farmers themselves (e.g. farm business advisors).

³ The key informant often acts as a 'gate-keeper' to the interviewees, in that his or her recommendation is often important in convincing them to participate.

⁴ The alternative is to use a sampling frame drawn from a telephone directory or other official source, such as government records. Weaknesses of directory-based selection include oversampling of specific types of farms (Burton and Wilson 1999).

⁵ Recording interviews is standard practice for qualitative interviewing, as it enables accurate record keeping of the interview content. Note-taking may occur when the data are particularly sensitive (and the interviewee hesitant to commit to a recording), or if situational variables (e.g. noise) do not allow for recording.

⁶ We were fortunate that in this work the qualitative field researcher and model developer worked in the same building, allowing numerous impromptu meetings and discussions to be had without prior arrangement. Those wishing to conduct a similar process without this luxury will need to allow for a significant level of communication in the research plan.

⁷ The 'Right to Roam' legislation referred to in Appendix 1 is Part 1 of the Land Reform (Scotland) Act, 2003, which stipulates rights of access to land in Scotland for recreation and education purposes. Section 5 Paragraph 1 of Part 1 of the Act states that exercising these rights does not constitute trespass.



References

- AAMODT A and Plaza E (1994) Case-Based Reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications* 7 (1), 39-59.
- AGAR M (2003) My kingdom for a function: Modeling misadventures of the innumerate. *Journal of Artificial Societies and Social Simulation* 6 (3), 8. <http://jasss.soc.surrey.ac.uk/6/3/8.html>
- AGAR M (2004) Agents living in color: Towards emic agent-based models. *Journal of Artificial Societies and Social Simulation* 8 (1), 4. <http://jasss.soc.surrey.ac.uk/8/1/4.html>
- BARRETEAU O and others (2003) Our companion modelling approach. *Journal of Artificial Societies and Social Simulation* 6 (2), 1. <http://jasss.soc.surrey.ac.uk/6/2/1.html>
- BECU N, Neef A, Schreinemachers P and Sangkapitux C (2008) Participatory computer simulation to support collective decision-making: Potential and limits of stakeholder involvement. *Land Use Policy* 25 (4), 498-509. [doi:10.1016/j.landusepol.2007.11.002]
- BHARWANI S, Bithell M, Downing T E, New M, Washington R and Ziervogel G (2005) Multi-agent modelling of climate outlooks and food security on a community garden scheme in Limpopo, South Africa. *Philosophical Transactions of the Royal Society B* 360, 2183-2194. [doi:10.1098/rstb.2005.1742]
- BOERO R and Squazzoni F (2005) Does empirical embeddedness matter? Methodological issues on agent-based models for analytical social science. *Journal of Artificial Societies and Social Simulation* 8 (4), 6. <http://jasss.soc.surrey.ac.uk/8/4/6.html>
- BOOCH G, Rumbaugh J and Jacobson I (1999) *The Unified Modeling Language User Guide*. Boston: Addison-Wesley.
- BOULANGER P-M and Bréchet T (2005) Models for policy-making in sustainable development: The state of the art and perspectives for research. *Ecological Economics* 55, 337-350. [doi:10.1016/j.ecolecon.2005.07.033]
- BOUSQUET F and Le Page C (2004) Multi-agent simulations and ecosystem management: A review. *Ecological Modelling* 176, 313-332. [doi:10.1016/j.ecolmodel.2004.01.011]
- BOWLER I, Clark G, Crockett A, Ilbery B and Shaw A (1996) The development of alternative farm enterprises: A study of family labour farms in the Northern Pennines of England. *Journal of Rural Studies* 12, 285-295. [doi:10.1016/0743-0167(96)00015-0]
- BRANNEN J (2005) Mixing methods: The entry of qualitative and quantitative approaches into the research process. *International Journal of Social Research Methodology* 8(3), 173-184. [doi:10.1080/13645570500154642]
- BRYMAN A (2008) *Social Research Methods*. 3rd Ed. Oxford: Oxford University Press
- BURTON R and Wilson G A (1999). The Yellow Pages as a sampling frame for farm surveys: Assessing potential bias in agri-environmental research. *Journal of Rural Studies* 15, 91-102. [doi:10.1016/S0743-0167(98)00015-1]
- BURTON R J F (2004) Seeing through the 'good farmer's' eyes: Towards developing an understanding of the social symbolic value of 'productivist' behaviour. *Sociologia Ruralis* 44 (2), 195-215. [doi:10.1111/j.1467-9523.2004.00270.x]
- BURTON R J F, Kuczera C and Schwarz G (2008) Exploring farmers' cultural resistance to voluntary agri-environmental schemes. *Sociologia Ruralis* 48 (1), 16-37. [doi:10.1111/j.1467-9523.2008.00452.x]
- D'AQUINO P, Le Page C, Bousquet F and Bah A (2003) Using self-designed role-playing games and a multi-agent system to empower a local decision-making process for land use management: The SelfCormas experiment in Senegal. *Journal of Artificial Societies and Social Simulation* 6 (3), 5. <http://jasss.soc.surrey.ac.uk/6/3/5.html>
- DENZIN N K and Lincoln Y S (1998) Chapter 1 Introduction. Entering the field of qualitative research. In Denzin N K and Lincoln Y S *Collecting and interpreting qualitative materials*. London: Sage.
- EDMONDS B (2007) Simplicity is not truth-indicative. In Gershenson C, Aerts D and Edmonds B (eds) *Worldviews, Science and Us: Philosophy and Complexity*. Singapore: World Scientific. pp. 65-80 [doi:10.1142/9789812707420_0005]
- EDMONDS B and Hales D (2004) When and why does haggling occur? Some suggestions from a qualitative but computational simulation of negotiation. *Journal of Artificial Societies and Social Simulation* 7 (2), 9. <http://jasss.soc.surrey.ac.uk/7/2/9.html>
- EDMONDS B and Moss S (2005) From KISS to KIDS: An 'anti-simplistic' modelling approach. In Davidsson P, Logan B and Takadama K (eds) *Multi-Agent and Multi-Agent-Based Simulation, Joint Workshop MABS 2004, New York, NY, USA, July 19, 2004, Revised Selected Papers. Lecture Notes in Artificial Intelligence* 3415, pp. 130-114. [doi:10.1007/978-3-540-32243-6_11]
- FRENKEN K (2006) Technological innovation and complexity theory. *Economics of Innovation and New Technology* 15 (2), 137-155. [doi:10.1080/10438590500141453]
- FRYSINGER S (2002) Integrative environmental modelling. In Clarke K C, Parks B O and Crane M P (eds.) *Geographic Information Systems and Environmental Modeling*. Upper Saddle River, NJ: Prentice Hall. pp. 211-222.

- GALÁN J M, Izquierdo L R, Izquierdo S S, Santos J I, del Olmo R, López-Paredes A and Edmonds B (2009) Errors and artefacts in agent-based modelling. *Journal of Artificial Societies and Social Simulation* 12 (1), 1. <http://jasss.soc.surrey.ac.uk/12/1/1.html>
- GOTTS N M and Polhill J G (2007) Using collective rewards and social interactions to control agricultural pollution: Explorations with FEARLUS-W. In Amblard F (ed.) *Interdisciplinary Approaches to the Simulation of Social Phenomena: 4th Conference of the European Social Simulation Association, Toulouse, France, 10-14 September 2007*. pp. 253-262.
- GOTTS N M and Polhill J G (2009a) Narrative scenarios, mediating formalisms, and the agent-based simulation of land use change. In Squazzoni F (ed) *Epistemological Aspects of Computer Simulation in the Social Sciences: Second International Workshop, EPOS 2006, Brescia, Italy, October 2006, Revised Selected and Invited Papers. Lecture Notes in Artificial Intelligence* 5466. Berlin: Springer. pp. 99-116. [doi:10.1007/978-3-642-01109-2_8]
- GOTTS N M and Polhill J G (2009b) When and how to imitate your neighbours: Lessons from and for FEARLUS. *Journal of Artificial Societies and Social Simulation* 12 (3), 2. <http://jasss.soc.surrey.ac.uk/12/3/2.html>
- GOTTS N M, Polhill J G and Adam W A (2003) Simulation and analysis in agent-based modelling of land use change. *First Conference of the European Social Simulation Association, September 18-21 2003, Groningen, The Netherlands*. http://www.uni-koblenz.de/~essa/ESSA2003/gotts_polhill_adam-rev.pdf
- GOTTS N M, Polhill J G and Law A N R (2003) Aspiration levels in a land use simulation. *Cybernetics and Systems* 34, 663-683. [doi:10.1080/716100277]
- GOTTS N M, Polhill J G, Law A N R and Izquierdo L R (2003) Dynamics of imitation in a land use simulation. *Proceedings of the 2nd International Symposium on Imitation in Animals and Artefacts, University of Wales, Aberystwyth, 7-11 April 2003*. pp. 39-46.
- HARE M and Deadman P (2004) Further towards a taxonomy of agent-based simulation models in environmental management. *Mathematics and Computers in Simulation* 64, 25-40. [doi:10.1016/S0378-4754(03)00118-6]
- HISSCHEMÖLLER M, Tol R S J and Vellinga P (2001) The relevance of participatory approaches in integrated environmental assessment. *Integrated Assessment* 2, 57-72. [doi:10.1023/A:1011501219195]
- HUIGEN M (2004) First principles of the MameLuke multi-actor modelling framework for land use change, illustrated with a Philippine case study. *Journal of Environmental Management* 72, 5-21. [doi:10.1016/j.jenvman.2004.01.010]
- HUIGEN M G A, Overmars K P and de Groot W T (2006) Multiactor modeling of settling decisions and behavior in the San Marino watershed, the Philippines: A first application with the MameLuke framework. *Ecology and Society* 11 (2), 33. <http://www.ecologyandsociety.org/vol11/iss2/art33/>
- JANSSEN M A and Ostrom E (2006) Empirically based, agent-based models. *Ecology and Society* 11(2), 37. <http://www.ecologyandsociety.org/vol11/iss2/art37/>
- LEAVESLEY G H, Markstrom S L, Restrepo P J and Viger R J (2002) A modular approach to addressing model design, scale, and parameter estimation issues in distributed hydrological modelling. *Hydrological Processes* 16 (2), 173-187. [doi:10.1002/hyp.344]
- LEWIS J A, Henry S M, Kafura D G and Schulman R S (1991) An empirical study of the object-oriented paradigm and software reuse. *ACM SIGPLAN Notices* 26 (11), 184-196. [doi:10.1145/118014.117969]
- MANSON S M and Evans T (2007) Agent-based modeling of deforestation in Southern Yucatán, Mexico, and reforestation in the Midwest United States. *Proceedings of the National Academy of Sciences* 104 (52), 20678-20683. [doi:10.1073/pnas.0705802104]
- MARSDEN T, Munton R and Ward N (1992). Incorporating social trajectories into uneven agrarian development: Farm businesses in upland and lowland Britain. *Sociologia Ruralis* 32, 408-430. [doi:10.1111/j.1467-9523.1992.tb00940.x]
- MASON J (2002) *Qualitative Researching*. Second Ed. London: Sage.
- MOSS S (1999) Relevance, realism and rigour: A third way for social and economic research. *Manchester Metropolitan University, Centre for Policy Modelling Report No. 99-56*. <http://cfpm.org/cpmrep56.html>
- MOSS S (2002) Agent based modelling for integrated assessment. *Integrated Assessment* 3 (1), 63-77. [doi:10.1076/iaij.3.1.63.7407]
- MOSS S (2008) Alternative approaches to the empirical validation of agent-based models. *Journal of Artificial Societies and Social Simulation* 11 (1), 5. <http://jasss.soc.surrey.ac.uk/11/1/5.html>
- MOSS S and Edmonds B (2005) Sociology and simulation: Statistical and qualitative cross-validation. *American Journal of Sociology* 110 (4), 1095-1131. [doi:10.1086/427320]
- PARKER D C and Meretsky V (2004) Measuring pattern outcomes in an agent-based model of edge-effect externalities using spatial metrics. *Agriculture, Ecosystems and Environment* 101, 233-250. [doi:10.1016/j.agee.2003.09.007]
- POLHILL G, Gimona A, Davies B, Gotts N, McCracken D, Catalan A, Dury G and Decors D (2008) A model for exploring human-natural system interactions in the context of agricultural biodiversity. *Fifth Conference of the European Social Simulation Association, University of Brescia, Italy, 1-5 September 2008*.
- POLHILL G, Izquierdo L and Gotts N (2008) *FEARLUS 1-1-5 User Guide*. <http://www.macaulay.ac.uk/fearlus/FEARLUS-model1-1-5-UG.pdf>
- POLHILL J G and Gotts N M (2009) Ontologies for transparent integrated human-natural system modelling. *Landscape Ecology* 24 (9), 1255-1267. [doi:10.1007/s10980-009-9381-5]

- POLHILL J G, Gotts N M and Law A N R (2001) Imitative versus nonimitative strategies in a land use simulation. *Cybernetics and Systems* 32 (1-2), 285-307. [doi:10.1080/019697201300001885]
- POLHILL J G, Parker D C and Gotts N M (2005) Introducing land markets to an agent based model of land use change: A design. In Troitzsch K G (ed.) *Representing Social Reality: Pre- Proceedings of the Third Conference of the European Social Simulation Association, Koblenz, September 5-9, 2005*. Koblenz, Verlag Dietmar Fölbach. pp. 150-157.
- POLHILL J G, Parker D C and Gotts N M (2008) Effects of land markets on competition between innovators and imitators in land use: Results from FEARLUS-ELMM. In Edmonds B, Hernandez Iglesias C and Troitzsch K G (eds.) *Social Simulation: Technologies, Advances and New Discoveries*. IGI Global. pp. 81-97. [doi:10.4018/978-1-59904-522-1.ch007]
- RAMANATH A M and Gilbert N (2004) The design of participatory agent-based social simulations. *Journal of Artificial Societies and Social Simulation* 7 (4), 1. <http://jasss.soc.surrey.ac.uk/7/4/1.html>
- ROBINSON D T, Brown D G, Parker D C, Schreinemachers P, Janssen M A, Huigen M, Wittmer H, Gotts, N, Promburom P, Irwin E, Berger T, Gatzweiler F and Barnaud C (2007) Comparison of empirical methods for building agent-based models in land use science. *Journal of Land Use Science* 2 (1), 31-55. [doi:10.1080/17474230701201349]
- ROUCHIER J and Hales D (2003) How to be loyal, rich and have fun too: The fun is yet to come. *First European Social Simulation Association Conference, Groningen, the Netherlands, 18-21 September 2003*. <http://www.uni-koblenz.de/~essa/ESSA2003/proceedings.htm>
- RUBIN H J and Rubin I S (1995) *Qualitative interviewing. The art of hearing data*. London: Sage.
- RYKIEL JR. E J (1996) Testing ecological models: the meaning of validation. *Ecological Modelling* 90 (3), 229-244. [doi:10.1016/0304-3800(95)00152-2]
- SHUCKSMITH M (1993) Farm household behaviour and the transition to post-productivism. *Journal of Agricultural Economics* 44 (3), 466-478. [doi:10.1111/j.1477-9552.1993.tb00288.x]
- SHUCKSMITH M and Herrmann V (2002) Future Changes in British Agriculture: Projecting divergent farm household behaviour. *Journal of Agricultural Economics* 53, 37-50. [doi:10.1111/j.1477-9552.2002.tb00004.x]
- SILVERMAN D (2001) *Interpreting qualitative data. Methods for analysing talk, text and interaction. 2nd Ed.* London: Sage.
- SUTHERLAND L (2010). Environmental grants and regulations in strategic farm business decision-making: A case study of attitudinal behaviour in Scotland. *Land Use Policy* 27 (2), 415-423. [doi:10.1016/j.landusepol.2009.06.003]
- TAYLOR R I (2003) Agent-based modelling incorporating qualitative and quantitative methods: A case study investigating the impact of e-commerce upon the value change. *Ph. D. Thesis, Centre for Policy Modelling, Manchester Metropolitan University. CPM Report No. CPM-03-137*. <http://cfpm.org/cpmrep137.html>
- WÄCHTER M (2003) The "Social-Ecological Research" Program. In Tress B, Tress G, van der Valk A and Fry G (eds) *Interdisciplinary and transdisciplinary Landscape studies: Potential and Limitations*. DELTA: Wageningen.