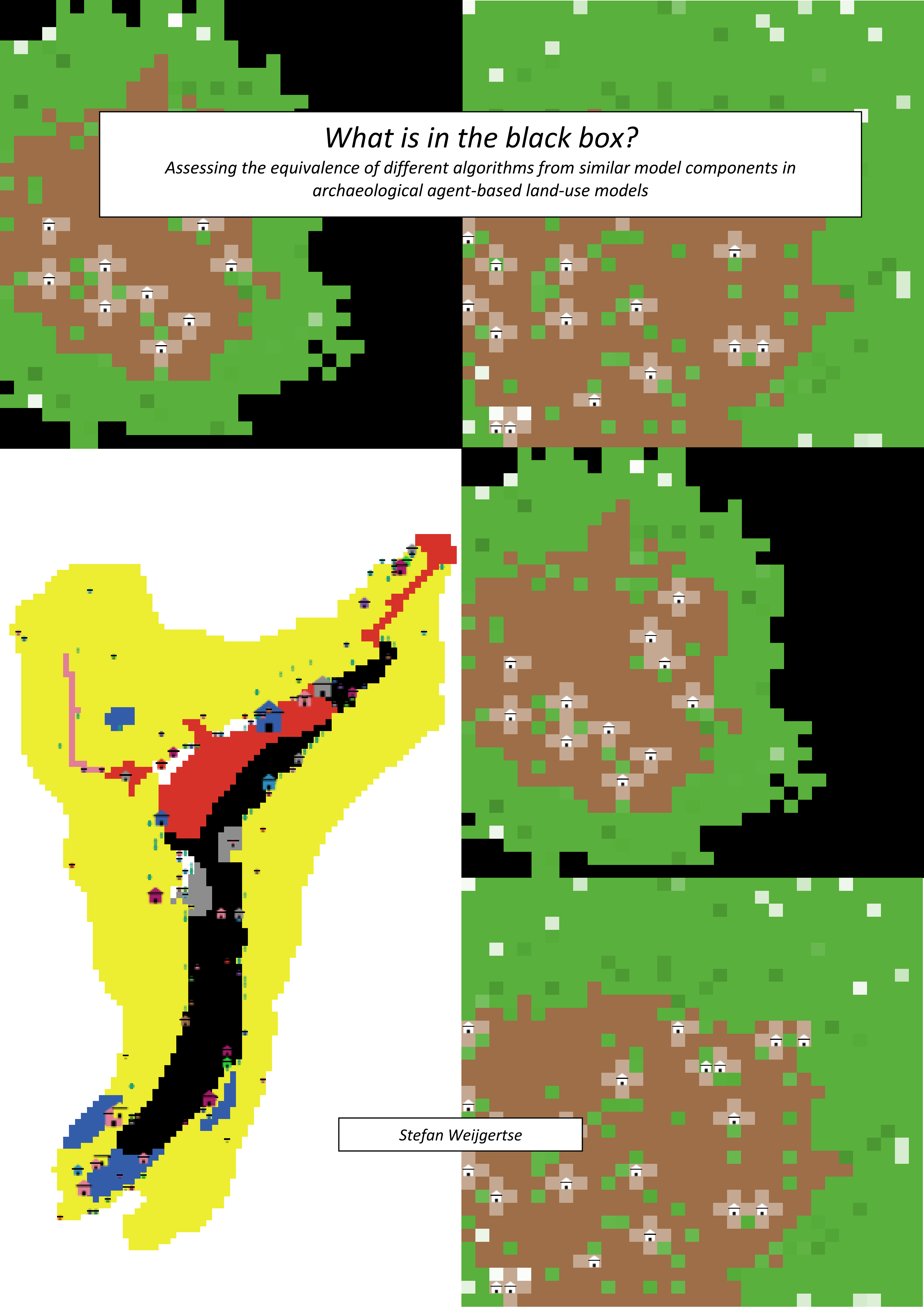


What is in the black box?

Assessing the equivalence of different algorithms from similar model components in archaeological agent-based land-use models

Stefan Weijgertse

Picture on front cover:
Collection of images from a diversity of runs from the Artificial Anasazi model
and ROMFARMS model (own figure).

What is in the black box?

Assessing the equivalence of different algorithms from similar model components in archaeological agent-based land-use models

Stefan Weijgertse, 1497790

Master Thesis Archaeological Science – 1084VTSY

Supervisor: Dr. Lambers

University of Leiden, Faculty of Archaeology

Leiden, 01-07-2020, final version

Table of contents

Acknowledgements

I would like to thank Dr. Karsten Lambers and Dr. Fulco Scherjon for all the help and advice they have given me during the process of this thesis. Without their insights, guidance and thoughts this thesis would not have turned out as it has. Furthermore I would like to thank Dr. Iza Romanowska, who has been a great help in defining the precise research topic and showing where exciting research possibilities lie in the field of archaeological modelling and simulation. Last, I would like to thank my housemates. This thesis has mainly been written during the outbreak of the COVID-19 pandemic, a time that will undoubtedly go down in history as a time where humans were forced to adapt and work as a team while generally living in isolation. This period has generally put a lot of mental and physical stress on people, students and non-students, due to the forced lifestyle changes that were a necessity to prevent contamination with the coronavirus as much as possible. The enthusiasm and positiveness of my housemates during these troubling times has been a wonderful distraction from the (more than ever) lonely hours of thesis writing and performing research.

1. Introduction

This thesis tries to establish whether the algorithms of similar components between different archaeological agent-based land-use models can be considered equivalent to each other. The general lack of transparency in agent-based models makes it difficult to establish whether the code from other models could be considered useful in the development of new models, or even improve existing models. A better insight in the relationships between models and the performance of different algorithms related to similar model components could significantly aid in the overall understanding of agent-based models. It could furthermore help future modellers in their own modelling endeavours. This first chapter introduces the general research problem that this thesis discusses: How the lack of transparency in archaeological agent-based influences our understanding of these models and how it affects model evaluation. The second part discusses the research questions and research aims of this thesis. The third part broadly introduces the applied research methodology and the final part introduces the layout of the thesis and the topics that will be discussed.

This thesis uses agent-based modelling concepts and terminologies that might not be immediately familiar to the general archaeological audience. To make the contents of this research more comprehensible for those readers that are more unfamiliar with agent-based modelling in archaeology, a terminology list where important concepts and terminologies are explained can be found in the first appendix of this thesis.

1.1 The research problem

The principle of Occam's razor states that the simplest and most elegant model holds the highest explanatory power. It is used as a heuristic tool in the process of model development for many sciences, but also in defining which models are the most suitable representation of a real-life phenomenon. Even though Occam's razor is a guiding principle many sciences, its principles do generally not apply to archaeology and must actually be reversed. Archaeologists must assume that the past is complex unless otherwise has been proved, the simplest model of a phenomenon can therefore not generally be assumed to also be the most suitable (van der Leeuw 2004, 121).

A model, or a conceptual model to be more precise, is a simplified representation of a real-world system or concept that has been created for a particular purpose (Lock 2003, 147). A conceptual model can be a physical or schematic representation of a real-world system, although the latter is arguably more often used for academic purposes. Figure 1 is an example of such a schematic conceptual model. It is a simplified representation of a real-world phenomenon, namely the origin and spread of agriculture in different parts of the world. One might debate whether such a model is true, but it is an example which illustrates that the use of models in archaeology can be useful. The model allows a general perception of a complex phenomenon, which makes it usable in archaeological debate. One might argue that the use of models in archaeology is actually inevitable: With the impossibility of directly observing the subject of study, namely the activity of humans in the past, archaeologists are forced to employ simplifications of a complex past (a model) to guide their research and will also produce them as a result of it (Clarke 1972, 3; van der Leeuw 2004, 121-122).

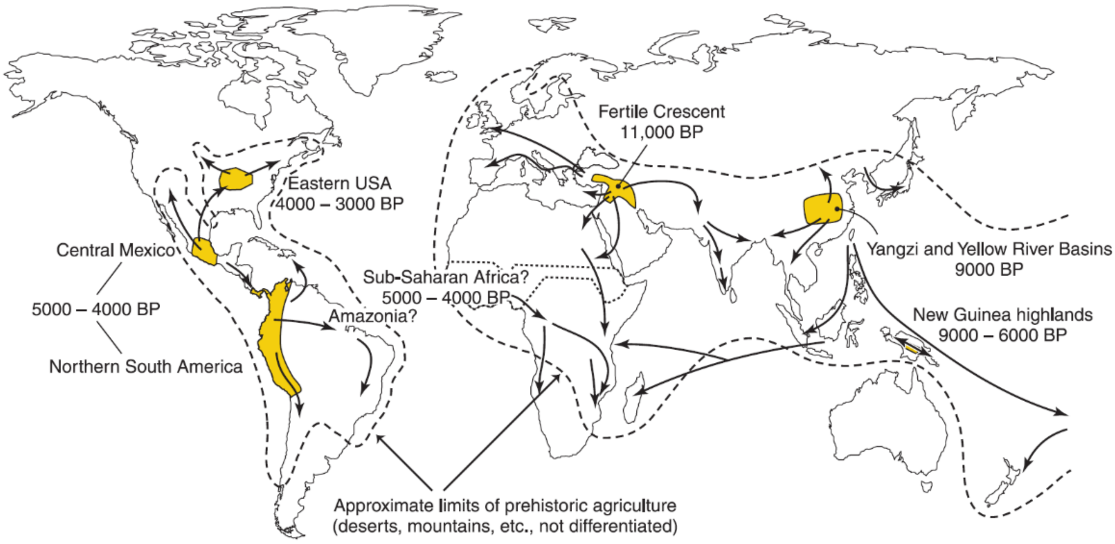


Figure 1: A conceptual model that shows the origins and spread of agriculture during prehistory (Diamond and Bellwood 2003, 597).

One particular class of models that has gained popularity in archaeology over the last decade are agent-based models. These models are a class of computational models (conceptual models implemented as a computer programme) that are employed to study the emerging patterns that are the result of the (inter)actions of autonomous, heterogeneous agents and their environment (Bandini *et al.* 2009, 4; Breitenecker *et al.* 2015, 60-61; Epstein 2006, 5-6; Romanowska *et al.* 2019, 181). An agent-based model consists of autonomous entities that have been programmed with a certain set of behaviour and a formal digital environment in which these agents can execute their behaviour. One could compare it with a videogame where the player does not participate but instead can only observe the behaviour of non-playable characters, whose (inter)actions affect each other and their environment. When an agent-based model is executed for a specific period of timesteps, its properties and behaviour over time can be observed and analysed. If the dimension of time is added to model to observe its behaviour, one speaks of a simulation (Bandini *et al. 2009,* 1; Hartmann 1996, 5; Lake 2015, 8; Romanowska 2015, 170).

Archaeological (agent-based) models and simulations are generally created based on experience, worldview, and ideas about the discipline (van der Leeuw 2004, 122). As a consequence, almost all archaeological agent-based models are black-box models. A black-box model is a model that is primarily built on a foundation of theories, hypotheses, ideas, observations, and knowledge (Breitenecker *et al.* 2015, 56-57). Contrary to white-box models, which are based on proven laws and axioms, the precise workings between input variables and output results are often unknown in black-box models. A black-box model in itself is not necessarily a bad thing and is often created in sciences that do not deal with fundamental truths, but the lack of transparency in these models causes them to have an inherent problem. This problem is that the lack of transparency in the process between input and output makes the quality, use and suitability of these models generally difficult to assess by a (scientific) community. A situation where the opinion of the archaeological community regarding a simulation model is equal to its opinion of the archaeologist that has built it, is certainly to be prevented.

A core component of the “black box” of an agent-based model is the ontology of the model. The ontology of an agent-based model can best be described as its whole of entities, relationships and rules of interaction. The design of the ontology and its dynamics when the factor time is added (i.e. when a simulation is conducted) are computationally expressed in the forms of algorithms (Romanowska *et al.* 2019, 179). Each ontology is unique and specifically suited for a particular agent-based model. However, there certainly is a possibility that there are different agent-based models in comparable research contexts that have semantically similar ontological components. When similarities between the ontological components of different models are observed, the way their algorithms are expressed would likely still vary greatly since black-box models make it difficult to assess how others have implemented solutions to similar challenges.

Model developers generally thus have to find their own solutions or designs of the same entities, relationships and interactions.

A better insight in similar ontological components and their algorithms could however be beneficial. An insight in how the ontologies of different models relate to each other might provide guidance in relation to which models can be considered relevant, whether in terms of inspiration or actual re-use of code, in the development of new models. This insight also allows comparisons and assessments of the performance of algorithms related to these similar components. It might very well be that there are algorithms developed by others that have the potential to produce similar results or even enhance the output quality of existing models. If this were the case for archaeological models, the results of different algorithms could potentially further increase the explanatory power of existing models in relation to the archaeological record.

One of the topics that has never been studied from such a point of view is agricultural subsistence and land-use. Even though agricultural subsistence processes are generally are to be a driving factor behind many emerging phenomena in archaeological theories, no-one has yet investigated and compared how these processes are being designed and algorithmically expressed in archaeological agent-based models that include these processes. There is no clear understanding of which ontological components are generally present in relation to agricultural subsistence processes and how they are given shape.

The design and definition of agricultural subsistence processes in simulations completely depends on the individual effort of the simulation builder(s), with many untransparent black-box simulations and a great variance in the design of agricultural subsistence processes as a result. A better general understanding of how these agricultural subsistence processes have been designed, which principles and ontological components they generally make use of, how these are expressed in computer code and what the effects of different algorithms on the same ontological components are, could potentially aid in future endeavours of simulation design, simulation understanding and simulation validation.

1.2 Research questions and goals

This thesis aims to explore similarity and variety between the ontologies of archaeological agent-based models with an agricultural and land-use component. With regards to those aspects of the ontologies that can be considered similar or general, this thesis furthermore aims to explore how these similar ontological components are algorithmically expressed and whether different algorithms can be considered functionally equivalent to each other.

The primary research question that will be answered in this thesis is as follows:

- *Can different algorithmic expressions of similar components in archaeological agent-based land-use models be considered equivalent to each other?*

This thesis furthermore poses a series of sub questions, which are formulated as follows:

- *Is the most suitable methodological approach towards the research question, which is derived from Axtell et al. (1996), still applicable?*
- *How does the applied research methodology aid in improving the transparency of black-box agent-based models?*

The goal of this thesis is to contribute to the archaeological agent-based modelling community by providing tools and new methodological insights for the assessment of ontological relationships between agent-based models, as well as for the assessment of the functionalities of different algorithms from similar model components. It would furthermore be positive if this research can aid in improving the transparency of black-box models. Higher degrees of transparency make agent-based models and formulating opinions on the models also more accessible to those archaeologists with a less technical background. The focus of the research goals is on archaeological agent-based modelling, but any contribution to modelling practices this research might make to other scientific disciplines is certainly considered a nice bonus.

1.3 Research methodology

The research methodology applied in this thesis is largely based on the methodological approach formulated by Axtell *et al.* (1996). In this study the authors used two different simulation models with a unique cultural transmission component , the SugarScape model from Esptein and Axel (1995) and a cultural convergence simulation from Axelrod (1995), to study how the cultural transmission algorithm from the first simulation influenced the output data from the other simulation. This study has investigated whether the algorithms of overlapping components between two models could be considered functionally equivalent. The authors have done this by implementing the ontological components from SugarScape in the simulation environment from Axelrod's cultural transmission model, a process that the authors called "docking", and consequently compared the output results of the docked model and the original model. The authors called the entirety of this process "model alignment". Statistically similar output results were produced and the authors argued that similar studies had to be performed more often. The main arguments for this statement were that the endeavour increased the understanding of both models, the relationships between the algorithms of the models, the consequences of the appliance of a specific algorithm and the quality of model results (Axtell *et al.* 1996, 136). This specific kind of research has however not been performed very often in its given form, even though the research was very influential for replication and validation studies of agent-based models. In archaeology, studies employing the methodology of model alignment are not known. This makes this research generally explorative. The general aims of the study from Axtell *et al.* match the primary research question and intentions of this thesis, their methodology should therefore be suitable for this research as well.

The study by Axtell *et al.* thus serves as an inspiration and guideline to approach the research questions from the previous paragraph, but the authors do not provide a clear methodology on how to approach the definition of similar components between models. The authors have made use of models that were generated by themselves and that they consequently knew and understood very well. This made a comparison of the models and their ontologies relatively easy, but still very untransparent.

This research aims to establish similarities between the ontologies of two agent-based land-use models and whether relational equivalence (similar behaviour under the same parameter conditions) and distributional equivalence (statistically similar model output) can be established in relation to their overlapping components. The Artificial Anasazi model by Janssen (2009) and the ROMFARMS model by Joyce (2019a; 2019b) are used as case studies. The research methodology follows the same implementation and output analysis phase as Axtell *et al.* (1996), but additionally focuses on developing a suitable methodology to define the ontological similarities between agent-based models. The research methodology therefore consists of three phases:

1. The model comparison phase. In this phase the ontological similarities between Artificial Anasazi and ROMFARMS will be defined.
2. The docking phase: in this phase the ontologically similar components from the Artificial Anasazi model will be implemented in the ROMFARMS model.
3. The verification phase: In this phase it will be assessed whether the docked model behaves as intended and produces statistically similar output results compared to ROMFARMS.

1.4 Thesis layout

This introductory chapter has named the research problem, research questions and how these research questions will be answered in the thesis. The remainder of this thesis can be divided in three parts: Research background, the active research and the discussion and conclusion of the results.

The first part, which consists of the second, third and fourth chapter, elaborates on the background of this research. The second chapter provides a wider context regarding the contribution of agent-based models to archaeology, the practice of developing simulation models and how agricultural subsistence is generally portrayed in archaeological agent-based models. The third chapter elaborates on the choice for the two case studies and what their respective archaeological research contexts are. This chapter also introduces NetLogo, the software in which both case studies have been developed, with the goal to make the technical aspects and terminologies in this research more comprehensible. The fourth chapter focuses on the background of the research methodology. It discusses current practices in relation to the three phases of the research methodology and defines the most suitable workflow for each of these three phases.

The second part focuses on the three phases of the research methodology and comprises the fifth, sixth and seventh chapter. Each chapter relates to a phase of the research methodology. The fifth chapter focuses on the comparison of the case studies, it presents qualitative results regarding their similarities and differences. These results are needed for the following phases of the research because they indicate how similar components from the Artificial Anasazi model can be implemented (docked) in the ROMFARMS environment. The sixth chapter focuses on the process where the similar components from the Artificial Anasazi model are implemented in the ROMFARMS model based on the results of the model comparison. The seventh chapter focuses on the quantitative analysis of the docked model and ROMFARMS and compares their results. This is then used to establish whether the models are equivalent.

The final part consists of the eighth and ninth chapter. The eighth chapter focuses on the discussion of the results related to all phases of the research methodology. The ninth chapter is the concluding chapter, where the research question and sub questions will be answered.

2. Research background

The previous chapter has introduced the reasons for this research and how the research will be performed. This chapter provides a methodological, theoretical and technical background of the research. Its aim is that the reader acquires a better understanding of agent-based models in general, how they are used in archaeology, how they are developed and what the current state is regarding agricultural subsistence and land-use in archaeological agent-based models.

2.1 Agent-based models in archaeology

According to Hartmann (1996, 6) the functional qualities of (agent-based) simulations can be divided in five categories:

1) A technique to study the dynamics of a complex system;
2) A generative tool for heuristics in the development of models, hypotheses or theories;
3) A replacement of real-life experiments;
4) A supportive tool for the efficient implementation of real-life experiments;
5) A tool for teaching and learning.

Archaeological agent-based simulations can potentially fulfil all of these functions, but in research contexts they primarily relate to the first and third function as a result of the impossibility of directly observing behaviour of humans in the past (Romanowska 2019, 180-181).

Agent-Based models employ (among others) elements of complexity science, an academic discipline that investigates the emergence of patterns that cannot be studied by individually researching the components that cause them (Mitchell 2009, 13; Romanowska *et al.* 2019, 179). By modelling individual behaviour in software agents (whether they represent humans or other entities) and their environment from the bottom up, agent-based modelling provides possibilities to empirically approach the individual system components whose dynamics lead to large scale observable patterns (Kohler 2012, 12-13). The output results of agent-based simulations can consequently be compared with the archaeological record. In this comparison it is possible to validate results of a simulation with the archaeological record. The best overview on which archaeological themes are studied with agent-based model can currently be found in Cegielski and Rogers (2016).

A correlation between archaeological data and simulation data, which can be determined with a range of validation techniques, could be considered an indication that an agent-based model has explanatory power. The explanatory power must, however, not be confused with the ability of a model to explain an observed phenomenon. The explanatory power of a model could be used in the interpretation of the archaeological record, but a model itself is never able to explain the archaeological record since it is a simplification of reality by nature. On the other hand, it is also possible that no correlation between output results is established. A lack of fitting data could, however, be just as useful. A lack of fitting data might indicate shortcomings or wrong assumptions in existing models or theories. A lack of correlation between simulation results and the archaeological record thus also provides opportunities to review the quality of existing models and develop new theories or research approaches (Dean *et al.* 2006, 91).

The practice of agent-based modelling and simulation is however not without criticism in archaeology. Examples of criticisms are that unexpected emergent patterns could also be a by-product of the agent-based model's architecture and not per se of the phenomena that they aim to investigate, the lack of transparency and standards in the practice of model building, the lack of refinement in modelling agent-behaviour as a set of rule-based systems and the difficulties in validating agent-based models (Huggett 2004, 83-84; McGlade 2014, 296 297).

The criticism of the lack of transparency is thus in line with the research problem of this thesis. The transparency of agent-based simulations has, however, already significantly increased since Grimm *et al.* (2006; 2010; 2017) have introduced and updated the so-called 'Overview, Design concepts and Details' (ODD) protocol, which provides a framework for the description of agent-based simulations. Grimm *et al.* (2017, 350) argue that, besides a better communication of the simulation model, the ODD also simplifies the process with which the reliability of models can be validated. Even though this protocol could potentially help in the study of simulation model reliability, in archaeology the number of these validation studies has always remained poor due to the complex nature of ABMs, the scale level at which validation should take place, the difficulty in understanding the modelling thought process despite descriptions and the technical know-how required to perform such studies (Axtell *et al.* 1996, 123-124; Kanters 2019, 8).

As a consequence the discipline currently includes many archaeological simulations of which the reliability and validity has never truly been investigated.

2.2 Development of archaeological agent-based models

The process of agent-based model development in archaeology is different and more problematic compared to other disciplines. One of the problematizing factors in archaeological simulation is that the subject of study is not directly observable. Archaeological agent-based models therefore have to rely more on biased or incomplete input datasets and proxy-data during model development. Another factor are epistemological issues regarding whether the models are even applicable on past societies (Romanowska 2015, 171). Romanowska (2015, see figure 2) has formalized a sequence for archaeological simulation development that takes many of these challenging factors into account. She distinguished three primary phases in the simulation development sequence: The conceptual, technical and dissemination phase (Romanowska 2015, 172). Each of these phases internally consists of different steps that all follow each other in sequence. This thesis aims for a comparison between models in relation to the fourth step of the sequence, entities and rules of interaction, because this step refers to the ontology of a model. The docking phase of this research can be compared with the firth step, coding and testing, and the verification phase with the sixth, seventh and eighth step.

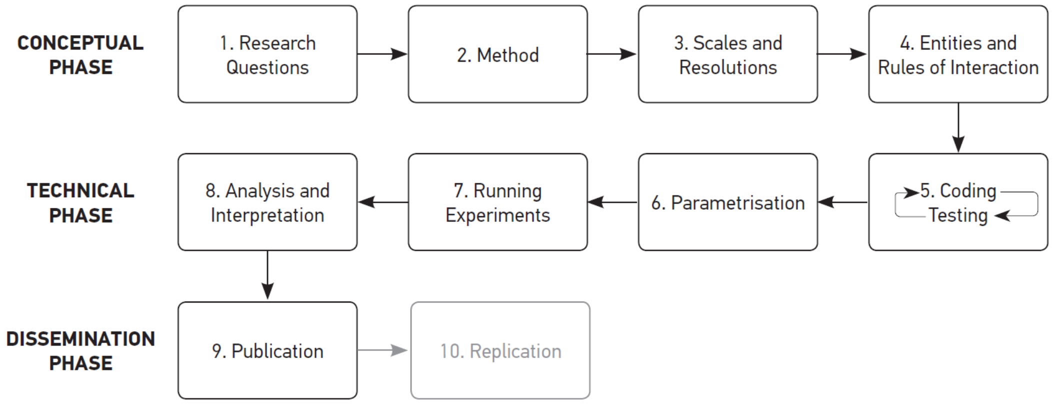


Figure 2: The development sequence for archaeological simulation models (Romanowska 2015, 172).

2.3 Agricultural subsistence in archaeological agent-based models

Across different sciences, the employment of agent-based modelling in land-use research has significantly increased since the early 2000s. Agent-based modelling approaches replaced the earlier use of mathematical formulas, called equation-based modelling, with decision making rules at the level of the individual (agent). This allowed a combination of programmed behaviour of individual entities with feedback from a digital simulation environment (Matthews *et al.* 2007, 1448). The article by Matthews *et al.* (2007) provides an overview of how land-use and agricultural strategies have been modelled throughout different sciences. In their work they mention one of the first examples where agricultural subsistence/land-use processes are employed in archaeological agent-based simulations: the Artificial Anasazi Model by Dean *et al.* (2000), which integrated agricultural subsistence processes with detailed environmental data to explore the settlement pattern dynamics of the Long House valley Anasazi culture from 800 to 1300 AD. This model is one of the most well-known archaeological agent-based models and is generally considered a good example of the contribution that agent-based models can have for archaeology. An elaborate overview of land-use and agricultural subsistence in archaeological agent-based models is, however, almost non-existent. This section nevertheless aims to provide a small introduction to the topic based on models that have been published in the last decade.

The increase of archaeological agent-based simulations in the last decade that has been noted by Lake (2014, 271) can also be seen in the number of agent-based simulations that apply agricultural subsistence in a variety of research contexts over the last decade. Especially in recent years the number of agent-based modelling studies that focus on agriculture has significantly increased (Joyce 2019a, 22). Saqalli *et al.* (2014, 46-47) note that agent-based models are specifically suited in archaeology to model and simulate agricultural and land-use practices since they allow a flexible incorporation of different kinds of social and environmental data, even when this data is (partially) polluted. Executions of the archaeological simulation studies related to agriculture are nonetheless still diverse. Based on recent studies, two different uses of agricultural subsistence are visible in archaeological agent-based models: Models where the functionality of the agricultural subsistence itself is the subject of study and models where the agricultural subsistence is a component in a system that explores different emerging phenomena. When the agricultural subsistence processes itself are the subject of study, the models are employed to increase the understanding of the economic resilience or performance of these processes under certain environmental or demographic circumstances (for example Angourakis *et al.* 2017; Stekerova and Danielisova 2016 and Wilkinson *et al.* 2013). When agricultural subsistence models are a component for exploring other emerging phenomena, models are often designed to investigate the role of agricultural strategies in settlement distributions and population spread/dynamics (for example Baum 2016; Bergin 2016; Heckbert 2013 and Janssen 2009) or their effect on the landscape or ecosystems (Joyce 2019; Riris 2018).

Exploring the literature of the models that have been mentioned as examples for the research contexts above clarifies why a comparison and better understanding of these kinds of models is both a methodological challenge as well as a necessity. Even without exploring the aforementioned simulation models in depth, big differences between the design of the models are visible despite the fact that the majority of them (all except Wilkinson *et al.* (2013) have been written in the NetLogo software. Some models, like for example Heckbert's (2013) and Baum's (2016), incorporate large amounts of environmental (GIS) data to approximate reality. Others, like Angourakis *et al.* (2017), use more abstract environments. Another aspect where differences are clearly visible is the number of adaptable parameters, which also varies across the different models. It is clear that each different research context might require a different approach towards the creation of the model, but it is also understandable that these differences make it difficult for the agent-based model developers in archaeology to understand where these models can be placed in the totality of available models and how the different models relate to each other.

3. Case studies

This chapter introduces the case studies that will be employed for this research. It begins with an explanation of the criteria applied in the choice for the case studies. The third part discusses the two case studies by focusing on their archaeological research contexts, contribution to the field in general and their contribution to the practice of agent-based modelling in archaeology. The chapter closes with an introduction to the NetLogo software, which is the modelling software with which both case studies have been developed.

3.1 Criteria

Two case studies of which the agricultural processes are known have been selected to be compared with each other and assessed for their ontological equivalence. The choice for only two case studies has consciously been made, because this kind of research is relatively explorative and a solid methodology to approach the research question is currently not very well established. It would thus be better to start as small as possible. The two selected models have been selected based on three criteria: The employed agent-based modelling software, the research topic and context, and the general expression of the agricultural processes. Even though a great variety of models have been created in the NetLogo framework, which would technically allow for a large number of possible comparisons, diversity in research topic and ease of accessibility have also been considered as factors in the final choice of the model. The two case studies that have been selected are the replication of the Artificial Anasazi Model by Janssen (2009) and the ROMFARMS model by Joyce (2019a). The fact that the temporal difference between the

	Artificial Anasazi	ROMFARMS
Software	NetLogo (v4.0.2)	NetLogo (v6.0.2)
Research topic	Settlement and population patterns of the Long House Valley Kayenta Anasazi Culture (800 - 1350 AD)	Impact of different agricultural strategies on land and labour in the Lower Rhine Delta (12 BC - 270 AD)
Expression of agricultural component	Agents represent maize agriculturalist households operating in a semi-realistic environment	Agents represent settlements pursuing an agricultural strategy in both semi-realistic and randomly generated environments
Accessibility	OpenABM library	Modelingcommons.org

Table 1: The case studies characteristics alongside the case study criteria.

publication of these models is ten years also allows for the definition of similarities between the models over a ten-year period - twenty when considering that the first publication of the Artificial Anasazi model was by Dean *et al.* in 2000. The characteristics of the case studies relative to the case study criteria are displayed in table 1.

The two case studies portray a visible distinction in terms of research contexts. In terms of expression of the agricultural component the first impression is that the Artificial Anasazi and the ROMFARMS model have a comparable expression despite their 10-year age gap. The choice for the Artificial Anasazi has been made due to the fact that it generally is a relatively well-understood and is an influential model that is generally considered an example of a good application of agent-based modelling in an archaeological context. The ROMFARMS model is chosen because it is one of the most recent applications of agricultural subsistence in archaeological agent-based models, but also because it is designed as a model- and theory building tool despite its reliance on palaeoenvironmental data.

3.2 Artificial Anasazi

The original Artificial Anasazi model was published in 2000 by Dean *et al.* and is considered one of the pioneering and most successful implementations of an agent-based modelling approach towards an archaeological topic (Epstein 2006, 89; Janssen 2009, 1-2). The model explores the spatial and demographic dynamics of the Kayenta Anasazi cultural phenomenon, which was present in the Long House Valley (North Arizona, USA) from 1800 BC until 1350 AD (Axtell *et al.* 2002, 7275; Dean *et al.* 2000, 180; Diamond 2002, 567; Gumerman *et al.* 2003, 436; Janssen 2009, 1). From the period during which these ancestors of contemporary American Pueblo cultures were present in Long House Valley, an extensive archaeological record and accompanying environmental data from 800 AD to 1350 AD has been collected during the late 1970s and 1980s (Dean et al. 1985; Dean and Gumerman 1989). Based on the archaeological and environmental data, static models on the population and spatial settlement dynamics of the Kayenta Anasazi were created in the late 1980s. In these models, it is described that Anasazi people would relocate their settlements and locations for maize cultivation. The models acknowledged that climatic variability in the Long House valley that affected maize growth was a factor (but not a primary factors since would be a deterministic assumption) in this process of relocation (Plot *et al.* 1988, 274-275).

The Artificial Anasazi model is a computational adaptation that combines the static settlement relocation models developed in the 1980s by the Southwestern Anthropological Research Group with Epstein's 1995 SugarScape model (Epstein 2006, 88-89). The SugarScape model was used to implement mechanisms of food consumption agent reproduction and agent death.

The Artificial Anasazi model utilizes environmental data to simulate the climatic circumstances between 800 and 1350 AD. The observed empirical relationship between the climatic circumstances and the dendrochronological data was used to accurately simulate precipitation and the consequent maize yields throughout the different ecological zones of Long House Valley (Diamond 2002, 568). A digital environment (see figure 3) was constructed that accurately simulated precipitation throughout the Long House Valley that was consequently willed with agents, each agent representing a household of 5 persons embedded with monoagricultural and demographic behaviour. The agents would harvest and consume maize throughout the landscape and adapt their subsistence strategies based on maize abundance. The interplay of simulated agent behaviour in a dynamic landscape combined with the transparency of the applied environmental and social data, consequently led to observable patterns in settlement distribution and population dynamics that could be validated against the archaeological record (Diamond 2002, 568).

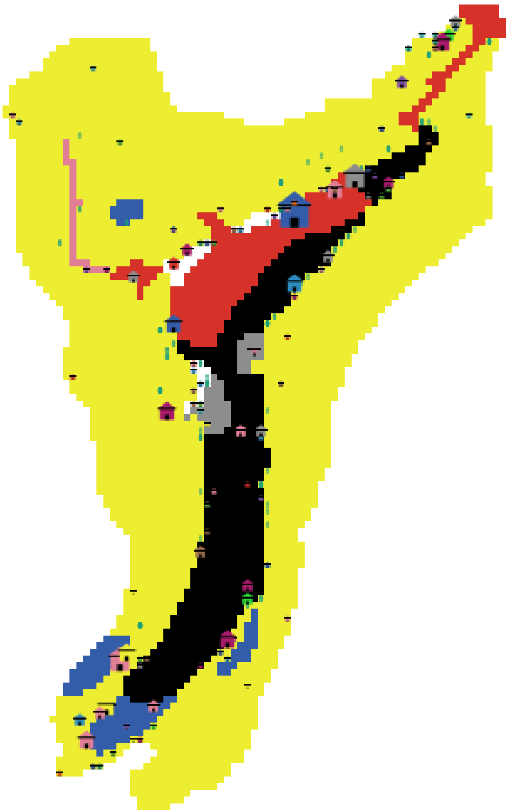


Figure 3: Picture of the simulation environment of the Artificial Anasazi model, the different landscape zones are visualized by different colours (own figure).

The first published version of the Artificial Anasazi model by Dean *et al.* (2000) was successful in simulating comparable general patterns of population growth and decline for the period 800 AD to 1350 AD, even though the population numbers from the simulated population were significantly higher than the estimates from the archaeological record (Dean *et al.* 2000, 191, Diamond 2002, 568). In the second version of the model, a greater degree of heterogeneity was established through alterations to the fertility and reproduction rates of households (Axtell *et al.* 2002, 7277). With these adaptations, the model was successfully able to approximate and reproduce the spatial and demographic patterns observed in the archaeological record, except for the sudden abandonment of Long House Valley around 1350 AD (Axtell *et al.* 2002, 7278).

The original Artificial Anasazi model was written in Ascape, an Integrated Development Environment for the creation of agent-based computational models. The Artificial Anasazi model used in this thesis is a NetLogo replication and slight alteration by Janssen (2009) of the second version of the model, which was published by Axtell *et al.* in 2002. The replication uses the same data as the original model and is made to get a better understanding of the functioning of the original model by attempting to reproduce the results utilizing a different software tool (see figure 4). Janssen concluded that, similar to Axtell *et al.*, the environmental data and agricultural behaviour alone do not generate the pattern of complete abandonment in the valley around 1350 (Janssen 2009, 14). Janssen's analysis furthermore shows that Axtell *et al.*'s adaptation of the demographic and agricultural behaviour rules of the agents in the model does not achieve the closest fit in relation to the archaeological record possible. Instead, an adaptation of the parameters related to the simulated maize carrying capacity of the valley would provide an even closer fit (Janssen 2009, 9). Janssen's Artificial Anasazi replication is written in NetLogo, accompanied by an ODD and accessible via the CoMSES OpenABM library (www.comses.net).

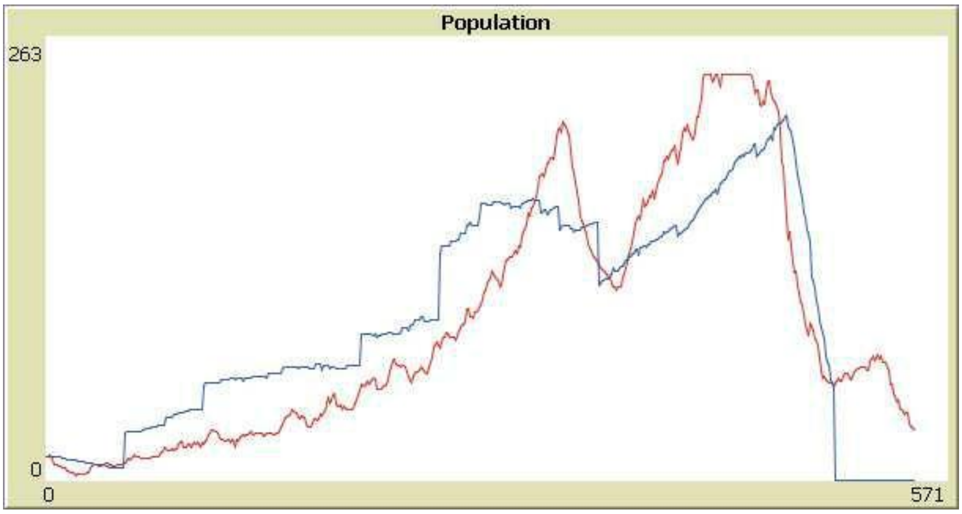


Figure 4: Comparison of the results from the Artificial Anasazi (red line) model with the historical data (blue line) (Janssen 2013, 6).

3.3 ROMFARMS

The ROMFARMS model has been created by Joyce (2019a) from the Free University of Amsterdam. The model explores the impact that different agricultural subsistence strategies would have had on the population and the environment of the Lower Rhine Delta in the Netherlands from 12 BC to 270 AD. The Lower Rhine Delta is the region where the Limes, the northern border of the Roman Empire during this period, was located. It is an explorative model that focuses on the impact of different agricultural subsistence strategies in the Lower Rhine Delta on the environment, aiming to serve as a heuristic tool in the creation of models and theories on agricultural subsistence practices in the Roman Period and other times as well (Joyce 2019a, 22-23). Despite its explorative nature, the ROMFARMS model itself is not completely theory-based since it employs palaeo-environmental data from the Lower Rhine delta.

The ROMFARMS model allows two types of landscape to be examined: a random generated environment and GIS reconstructions of the Lower Rhine Delta between 12 BC and 270 AD (Joyce 2019a, 25). For both these random and reconstructed environments, the ROMFARMS model workflow consists of a series of submodels that together represent a mixed farming subsistence strategy: population dynamics, arable farming, animal husbandry, fuel collection and timber collection to model construction activities (Joyce 2019a, 24). Via this workflow three different scenarios of agricultural strategies were tested by the adaptation of different parameters relating to the submodels and either incorporating or leaving out certain aspects of the submodels. These strategies are subsistence-based agriculture (referred to as intensification) in a random environment, surplus agricultural production (referred to as extensification) in a random environment and both of these strategies in a reconstructed landscape.

Even though the agents in the most recent version of ROMFARMS are not programmed with any forms of internal and external socio-economic behaviour - the agents are assumed to only make economically rational decisions - the three different scenarios did show different dynamics in terms of land-use and labour patterns (Joyce 2019b, 123). The random generated environments (see figure 5) were initially applied to get a better insight and fundamental understanding of the behaviour of the simulation, whose behaviour could subsequently be more critically assessed in a reconstructed landscape. From the analysis of the reconstructed landscapes it was concluded that the Lower Rhine delta

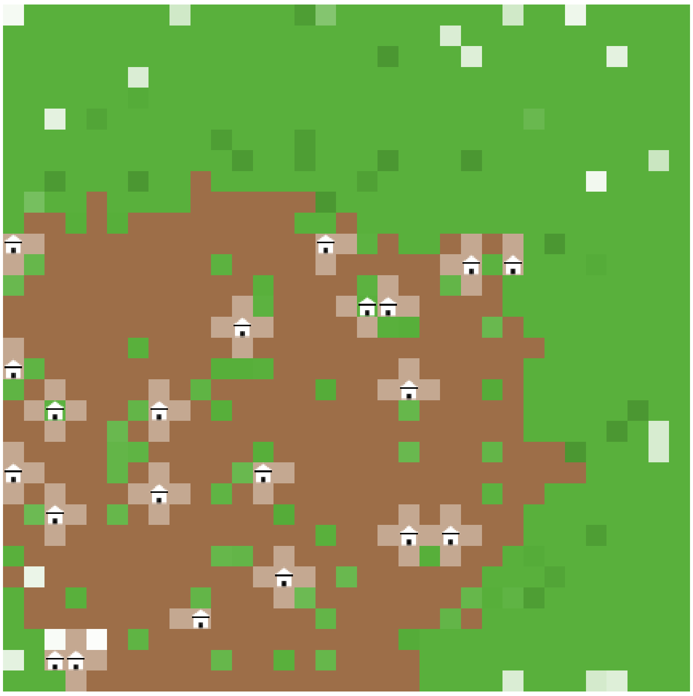


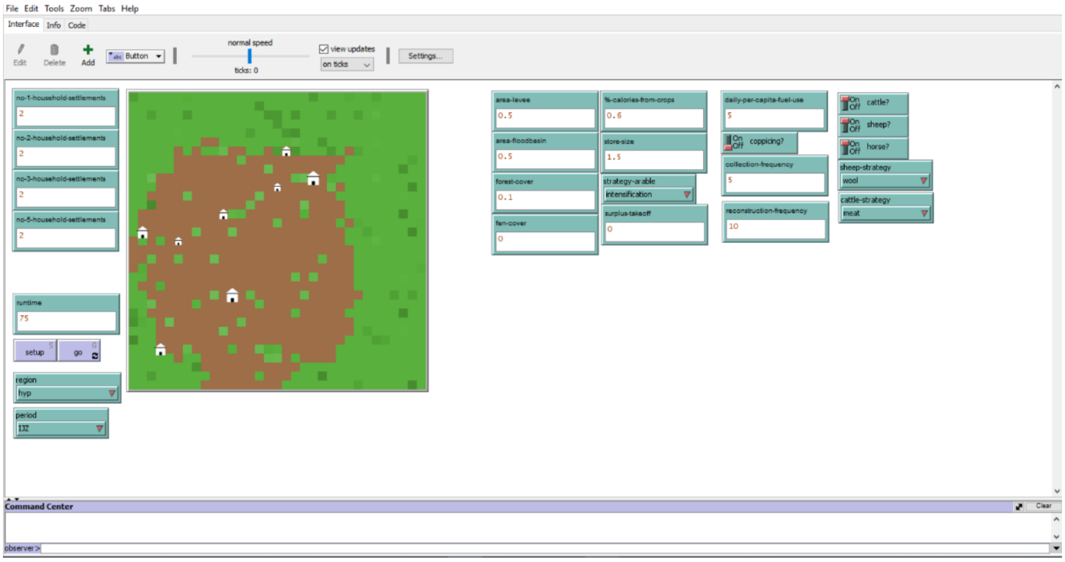
Figure 5: Visualization of the ROMFARMS simulation environment, with the settlements (white houses) engaged in cultivating surrounding areas (own figure).

generally yielded sufficiently available productive land and that arable production was not limited by the landscape, even when settlements pursued an extensive and high space-consuming agricultural strategy (Joyce 2019a, 195). It was furthermore concluded that animal husbandry practices could also not be a limiting factor for a shortage of available land due to the fact that animal husbandry practices, whether intensive of extensive, did not take up significant amounts of space. The only way that the total carrying capacity of the environment could have been crossed would have been by the existence of agricultural settlements with population sizes and density which have not been observed in the archaeological record (Joyce 2019a, 195).

The ROMFARMS model is thus a model to explore the different subsistence strategies and to develop heuristic tools applied in further model- and theory building. In its current form it is a simplistic model, although designed with a high degree of complexity that is the result of its many submodels, in the sense that does not approximate reality-based agriculture in the Roman period due to a significant lack of socio-economic factors related to the functionality and presence of the Roman Empire in the Lower Rhine delta. It is for example a well-established fact that agricultural communities also provided the Roman camps with food. Joyce (2019b, 123) however acknowledges this, and the fact that the model itself is still relevant for explaining and illustrating a sort of null scenario in relation to agricultural subsistence strategies makes it a useful model. The ROMFARMS model, without the GIS data of the reconstructed environments and accompanying model description, can be accessed via modellingcommons (modellingcommons.org).

3.4 The NetLogo environment

Both case studies that will be used in this research have been developed in the NetLogo integrated development environment (IDE). The NetLogo software offers a modelling environment for the creation of agent-based models and their execution and has been developed by Uri Wilensky from Northwestern University's Centre for Connected Learning and Computer-Based modelling. The following section provides a brief introduction to the basic layout and functionality of the NetLogo environment, with the goal that non-specialist readers become familiar with the general terminology that will be used in this thesis.



no-1-household-settlements: 2
no-2-household-settlements: 2
no-3-household-settlements: 2
no-5-household-settlements: 2
runtime: 75
region: hyp
period: IJZ
area-levee: 0.5
area-floodbasin: 0.5
forest-cover: 0.1
fen-cover: 0
%-calories-from-crops: 0.6
store-size: 1.5
strategy-arable: intensification
surplus-takeoff: 0
daily-per-capita-fuel-use: 5
collection-frequency: 5
reconstruction-frequency: 10
sheep-strategy: wool
cattle-strategy: meat

Figure 6: The ROMFARMS simulation environment in NetLogo (own figure).

The NetLogo development environment consists of three different windows in which a model developer can work: the model interface tab, the info tab, and the code tab (see figure 6 and 7). In the interface tab the simulation environment can be initialized, the simulation can be executed and the parameters of the agent-based model can be adjusted for experiments. The info tab can be used by the model developer to explain the background of the model and provide an instruction of its functionality so that it is

More comprehensible for external users. Not all model developers however make use of this info tab, since many prefer to provide a detailed model description in a prescribed framework like the ODD to accompany the model. In the code tab the algorithms that make up the source code of the model can be found. NetLogo uses the programming language Logo for the construction of its algorithms. Logo is a very high-level programming language, meaning that its high level of abstraction makes it relatively easy to read and understand for humans.

```
to births
  set-fertility
  ask persons with [ my-spouse != nobody and sex = "female" and age >= 16 and age < 50 ] [
    if random-float 1.0 < fertility [ ;; generates a random value 1from 0 to 1. If the number is less than the fertility rate assigned to each married woman between 16 and 49 one child is generated. The sex of the
      hatch 1 [
        set breed persons
        set age 0
        set sex one-of ( list "male" "female" )
        set my-house [ my-house ] of myself
        set my-spouse nobody
        set distance-travelled-fuel 0
        set distance-travelled-construction 0
      ]
      set count-children count-children + 1
    ]
  ]
end

to deaths
  set-mortality
  ask persons [
    if random-float 1.0 < mortality [ ;; generates a random value from 0 to 1. If the number is less than the mortality rate assigned to each person, the person dies and is removed from the simulation.
      die
    ]
  ]
end

to marriages
  ask persons with [ sex = "male" and age >= 16 and age < 50 and my-spouse = nobody ] [ ;; if a male aged 16 to 49 is unmarried, they will seek a new spouse. If no unmarried females between 16 and 49 are available
    if any? other persons with [ sex = "female" and age >= 16 and age < 50 and my-spouse = nobody and my-house != [ my-house ] of myself ] [
      ifelse [ count people with [ my-spouse != nobody ] ] of my-house < ( [ no-households ] of my-house * 2 ) [ ;; if there is space for the new couple, the couple move to the male spouse's settlement. If not, th
        set my-spouse one-of other persons with [ sex = "female" and age >= 16 and age < 50 and my-spouse = nobody and my-house != [ my-house ] of myself ]
        ask my-spouse [
          set my-spouse myself
          set my-house [ my-house ] of myself
        ]
      ]
      [
        set my-spouse one-of other persons with [ sex = "female" and age >= 16 and age < 50 and my-spouse = nobody and my-house != [ my-house ] of myself ]
        ask my-spouse [
          set my-spouse myself
        ]
        new-settlement ;; if no space in either of the new spouse's original settlements, the sub-process for creation of a new settlement is called.
```

Figure 7: The ROMFARMS source code in NetLogo, the births, deaths and marriages procedures are visible (own figure).

The functionality and dynamics of the model are being executed in the simulation environment that can be found in the interface tab. The simulation environment generally includes four different kinds of actors: The patch (a static grid cell), agent (referred to as 'turtle' in NetLogo jargon), the link (a relationship between turtles) and the observer (active input from the model operator/developer) the other two.

The behaviour of the different agents is regulated via sets of commands and reporters which can be expressed in the Logo programming language. A command is an algorithm that tells an agent what to do and a reporter is a calculation that makes the agent report a computational value. Many (sets of) commands and reporters are already present in NetLogo and are called primitives, completely self-developed commands are called procedures. In the construction of these procedures, agents might require access to computational values for, for example, sensing and interpreting their environment. The places in which these computational values can consequently be stored are called variables. Variables can belong to an individual agent or patch, or apply to all agents. In the case of the latter the variable is referred to as a global variable. Variables that can be adapted in the interface tab are called parameters.

The final relevant part of the NetLogo environment is the BehaviorSpace tool. This tool allows for the execution of multiple simulations with different (but predefined) parameters. BehaviorSpace allows for the quantitative results of all the executed simulations to be exported in a single table. This table can consequently be used for data analysis with different computational tools like Excel, Python, SPSS and R. The BehaviorSpace tool is a great tool for generating the quantitative output of a model and for conducting experiments, but also lends itself for a sensitivity analysis.

4. The research methodology

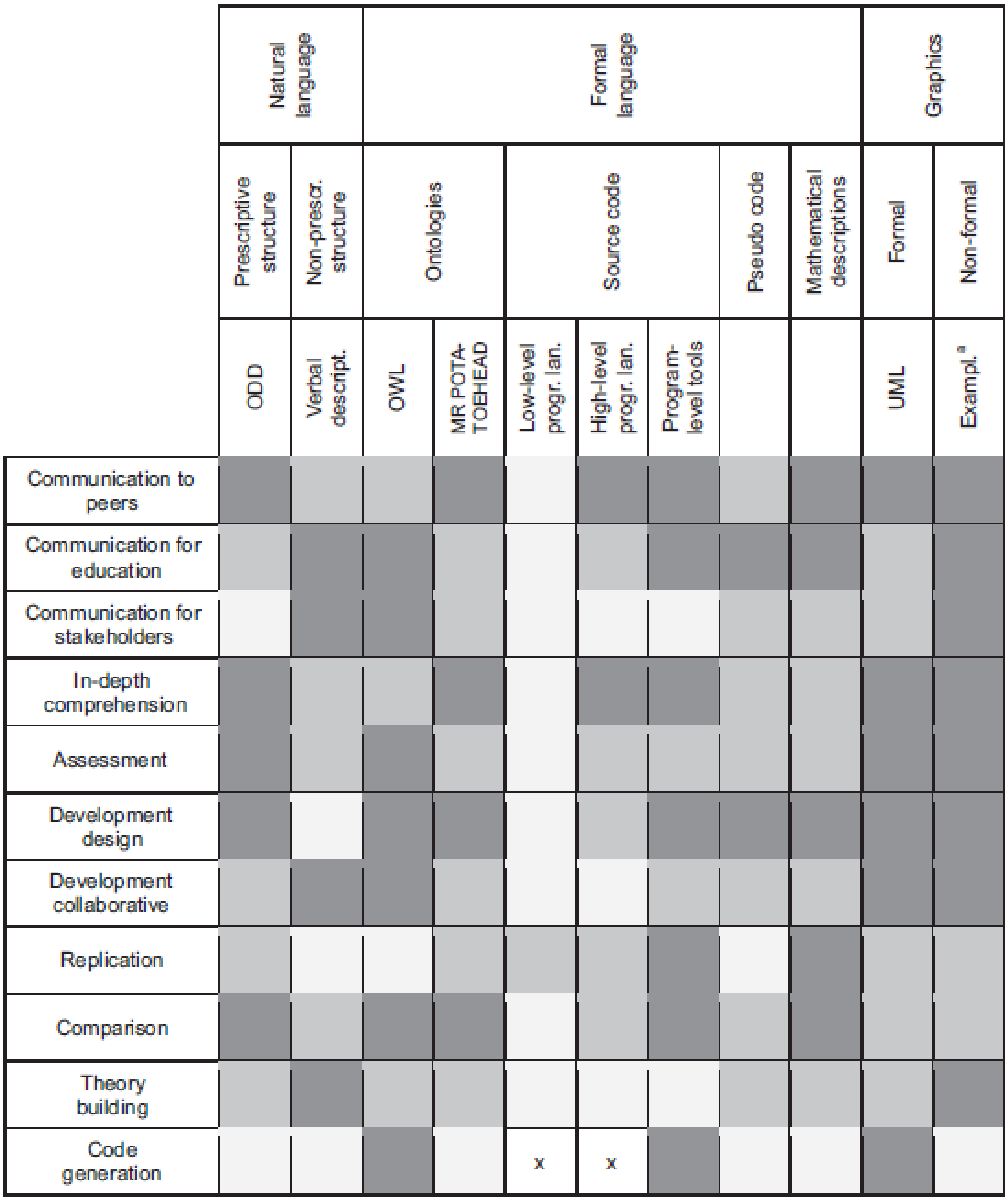
As mentioned in the introduction, this research follows a methodological sequence that it comparable to that of Axtell *et al.* (1996). The methodology from Axtell *et al.* is however not directly transferable since the research dates from 24 years ago and the authors do not provide guidelines for assessment of model applicability, how to efficiently dock models, and how to consequently verify them. This chapter explores current practices in relation to the comparison of agent-based models, docking agent-based models and verifying agent-based models. It is consequently defined how information from these practices can be translated to applicable approaches towards the different methodological phases. Each paragraph relates to one of the three research phases. The first paragraph discusses the comparison of agent-based model ontologies, the second paragraph discusses agent-based model docking and the third paragraph discusses the verification of agent-based models.

4.1 Comparing agent-based models

4.1.1 Current practices

Even though the first chapter of this thesis has illustrated that the nature of agent-based models makes them highly valuable in the study of complex human behaviour, there is still much to be won in terms of model production efficiency and transparency in relation to all phases of the model development sequence (Müller *et al.* 2013a, 156; Schulze *et al.* 2017, 2). The incorporation of simulated human behaviour in social and social-ecological agent-based simulation, with black-box models as a result, requires an explicit need for clear and uniform descriptions of model functionalities to enhance the transparency of agent-based models and their usability by other parties (Janssen *et al.* 2008, 1; Muller *et al.* 2013a, 157). Clear and elaborate model descriptions, especially those employing existing description frameworks, could aid in more efficiently executing the assessment, replication and - relevant for this research - comparison of agent-based models (Müller *et al.* 2013a, 157).

Model descriptions are the most frequently employed tools in the practice of model comparisons, but model descriptions come in different formats and their value in relation to the comparison of agent-based models therefore varies. Müller *et al.* (2013a, see table 2) distinguish three categories of model descriptions: Natural language descriptions,



	Natural language		Formal language							Graphics	
	Prescriptive structure	Non-prescr. structure	Ontologies		Source code			Pseudo code	Mathematical descriptions	Formal	Non-formal
	ODD	Verbal descript.	OWL	MR POTATOEHEAD	Low-level progr. lan.	High-level progr. lan.	Program-level tools			UML	Exampl.[a]
Communication to peers	▓	▒	▒	▓	░	▓	▓	▒	▓	▓	▓
Communication for education	▒	▓	▓	▒	░	▒	▓	▓	▓	▒	▓
Communication for stakeholders	░	▓	▓	▒	░	░	░	▒	▒	▒	▓
In-depth comprehension	▓	▒	▒	▓	░	▓	▓	▒	▒	▓	▓
Assessment	▓	▒	▓	▒	░	▒	▒	▒	▒	▓	▓
Development design	▓	░	▓	▓	░	▒	▓	▓	▓	▓	▓
Development collaborative	▒	▓	▓	▒	░	░	▒	▒	▒	▓	▓
Replication	▒	░	░	▒	▒	▒	▓	░	▓	▒	▒
Comparison	▓	▒	▓	▓	░	▒	▓	▒	▓	▒	▒
Theory building	▒	▓	▒	▒	░	░	░	▒	▒	▒	▓
Code generation	░	░	▓	░	x	x	▓	░	░	▓	░

Table 2: Overview of the different available agent-based model description tools and their suitability for different purpose, the most suitable tools for comparison can be found in the third-lowest row (Müller et al. 2013a, 159).

formal language descriptions and graphical visualizations. The ODD descriptive framework from Grimm *et al.* (2006; 2010; 2017) is the most utilized type of natural language description. Ontologies - a type of data-structure which is not to be confused with the aforementioned definition of an ontology in the first chapter - and program level-tools are the most suitable tools for formal language descriptions (see table 2). Graphical visualizations come in many different forms and are generally employed simultaneously with, or alongside, natural language model descriptions and will therefore not be discussed.

The most employed tool that belongs to the natural language description category is thus the ‘Overview, Design concepts and Details’ (ODD) framework. The ODD was originally developed and published in 2006 by ecologists and intended to be used in the description of ecological agent-based models (Grimm *et al.* 2006, 115; Janssen *et al.* 2008, 3). The ODD framework is a natural language-based model description technique that employs a prescriptive structure for the model description (Müller *et al.* 2013a, 157). The ODD framework focuses on the description of seven elements whose functionalities are crucial for a general understanding of an agent-based model and, even though it to a certain extent limits descriptive freedom, allows for a better overall insight in the agent-based models produced by a modelling community (Grimm *et al.* 2010, 2763).

Even though the first version from 2006 and the updated version from 2010 have been frequently employed and could arguably be considered the minimum baseline for the description of agent-based models, this framework has primarily experienced critique from the social sciences. Due to the ecological origins of the framework Müller *et al.* (2013b, 38-40) argue that the component of human decision-making, a key element in social and social-ecological agent-based models, is not sufficiently represented in the ODD. As a result they have proposed the ODD + D (ODD + Decision) protocol, an extension of the ODD protocol that adds new subcategories and guiding questions related to human decision-making elements. The ODD + D protocol is particular helpful to the social sciences because it allows an increased insight in how agents and their behaviour, some of the key components of a model ontology, are expressed in a model.

The second description technique that Müller *et al.* (2013a) name as being suitable for the purpose of model comparison, is the ontology. In this case the authors do not refer to the totality of entities and rules of interaction in an agent-based model in a conceptual sense. An ontology from a computer science perspective refers to a type of data structure that mathematically describes the components of a conceptual domain, to prevent confusion the ontology of a model it will hence be referred to as ‘computational ontology’. Since a computational ontology is a formal representation of a conceptual domain it can be read and understood by computers, these data structures are therefore often employed in artificial intelligence and knowledge management studies (Livet *et al.* 2010, 4). Different formal languages exist in which these computational ontologies can be written, but the Web Ontology Language (OWL) is currently the most frequently used

(Müller *et al.* 2013a, 157). In relation to agent-based agricultural and land-use models specifically, a framework for the development of land-use models and computational ontologies has been created by Parker *et al.* (2008). This framework is called the "Model Representing Potential Objects That Appear in The Ontology of Human-Environmental Actions & Decisions", or MR POTATOHEAD for short (Parker *et al.* 2008, 3). This framework includes formal expressions of ontological components (classes, properties, qualities, relations, and processes) that are often observed in agent-based land-use models. The MR POTATOHEAD framework allows researchers to pick those components that they deem necessary for their research to simultaneously build conceptual models and their formal ontologies. No studies besides those of the authors, Parker *et al.* (2008) and Parker *et al.* (2012), have however explicitly used the MR POTATOHEAD for model building, computational ontology development, and comparison purposes. The usefulness and impact of the MR POTATOHEAD framework can therefore not critically be assessed, despite its initially promising contribution to the practice of agent-based modelling in the social sciences.

The final useful description technique is the program-level tool. Program-level tools are software platforms that employ libraries of pre-defined algorithms written in high-level programming languages. High-level programming languages have a high degree of abstraction, which makes them easier for humans to read and understand. The NetLogo software is one of these program-level tools: its own programming language called Logo is a very high-level programming language. The fact that the source code of these models is human-readable due to the high levels of abstraction allows models to be compared on the basis of their source code as well.

Müller *et al.* (2013a, 162) note that there is not one type of model description that can completely perform all possible purposes and functionalities, since every type of model description comes equipped with its own limitations. Even though the most suitable types of model descriptions for agent-based model comparison are thus known (ODD, ontology and program-level tool), the way a comparison can be performed still depends on the variety, quality and accessibility of the different type of model descriptions.

Another question must be addressed as well before working towards a suitable methodology. This question is in what aspects archaeological agent-based models with an agricultural or land-use component differ from socio-ecological or land-use models employed in other social sciences. Different points and qualities of these models therefore need to be considered.

In their general overview of applications of agent-based land-use models, Matthews *et al.* (2007) include the Artificial Anasazi model. Matthews *et al.* place the Artificial Anasazi model in line with various socio-ecological models from other sciences that deal with similar issues, but on a contemporary rather than a historical basis. Based on this overview one could thus argue that archaeological agent-based models, especially those with land-use or (settlement) dispersal components like the Artificial Anasazi, could generally be classified as socio-ecological/land-use models in line with those from other social sciences.

The classification by Matthews *et al.* does, however, not insinuate that archaeological agent-based models are similar to those employed in other social sciences. Even though many similarities exist between archaeological agent-based models and those in other social sciences, there are also some inherent differences. As explained in the second chapter, archaeological models differentiate themselves from those in other sciences in relation to their resolution, input data and difficulty to apply models on past societies (Romanowska 2015, 171).

4.1.2 Examples

Another relevant topic is what can be learned from existing descriptive model comparisons. There is a clear need for more studies where the properties of agent-based models are being compared to each other, since the number of such studies that can be used for guidance is negligible. Due to a lack of systematic approaches and standards, most of the studies that present themselves as descriptive agent-based model comparisons are often explorative in their methodology as well (e.g. Cioffi-Revilla, 2017) or rather focus on a cross-comparison of the output results rather than the ontologies of the models themselves (e.g. Adam *et al.* 2017). In these studies, the comparisons are also performed by the developers of the models and not by an external researcher.

One promising example where the properties of two existing agent-based models are being compared to each other is the study from An *et al.* (2014), who compared the land-use models of the Wolong Nature Reserve (China) by An *et al.* (2005) and the Chitwan National Park (Nepal) by Zvoleff and An (2014) based on the ODD framework. The Wolong model was set as a baseline along which the Chitwan model was consequently compared by defining similarities and differences relative to the other model. The authors were able to illustrate that, based on the ODD comparison, the two different land-use models employ similar structural elements and processes (An *et al.* 2014, 741).

Even though the methodological approach in this study seems useful in defining differences and commonalities in model structures, the question remains whether the presence of these structural elements in this research is a result of the methodological approach or the fact that multiple authors performing the comparison were also involved in the development of both models. A certain level of subjectivity in the primary conclusions of this study can therefore not be excluded. This research has been performed on what Cioffi-Revilla (2017, 224) defines as a generic level of comparison. On the specific level, which comprises the ontology and system dynamics (Cioffi-Revilla 2017, 224), no sound conclusions were drawn using the ODD framework.

It could however be argued that such a study is still useful in more accurately defining the areas of an ontology where similar components might be present. Since the ODD + D protocol by Müller *et al.* (2013b) specifically focuses on the description of agent behaviour and the entities related to this behaviour, it also seems possible that the application of this framework allows for a better inclusion of the model ontologies in descriptive model comparisons.

4.1.3 Defining the methodology

The next step is to combine the information from the previous sections in this paragraph to define a suitable methodology for a descriptive comparison with which similar ontological components can be defined. As noted, the best methodology is the one that makes the most efficient use between available materials and three potentially useful tools for the comparison of the models: The ODD, the ontology and the program-level tool. Müller *et al.* (2013a, 162) note that the minimum materials accompanying a model must be a natural language description and the source code of the model. Even though the level on which the comparison must be achieved is that of the ontology, most models do not come equipped with a computational ontology that has all entities, relations and interactions formally defined. The employed case studies in this research also do not come equipped with a computational ontology, so a direct comparison on the level of the ontology is unfortunately not possible. The research from An *et al.* (2014) has however shown that a thorough comparison of both models via the ODD framework has the potential to still define those areas where significant overlap between different models is present. The ODD + D protocol from Müller *et al.* (2013b) provides an additional focus on the agent decision-making processes. This adds extra potential to focus on those (inter)relationships, simulated behaviour and entities that are central in the ontology of a model. This framework is furthermore suited because the comparison does not inherently require the necessity for the presence of an already existing ODD + D.

Table 3 illustrates the differences between the lay-out of the original ODD protocol and the extended protocol. From this table it is relatively easy to make up that the basic structure of the original framework remains largely intact and that most alterations to the framework that have been made are related to the Design Concepts part of the framework. Each category of the framework comes equipped with sub questions whose answers are relevant for its topic. Besides alterations to the labelling and order of some of the categories, Müller *et al.* have also extended the number of questions related to most categories. A total overview of these questions can be found in appendix 2, which has been included because it allows for a better insight in the different topics that each subcategory covers. The ODD + D framework thus provides a framework of relevant categories and their associated sub questions to describe agent-based models that are embedded with some form of human behaviour. All categories can be described using natural language and there are no prescriptions on how this language needs to be structured or formulated.

	Elements of the ODD protocol (Grimm *et al.* 2010)	Elements of the ODD + D protocol (Müller *et al.* 2013b)
Overview	1.1 Purpose	1.1 Purpose
	1.2 Entities, state variables and scales	1.2 Entities, state variables and scales
	1.3 Process overview and scheduling	1.3 Process overview and scheduling
Design Concepts	2.1 Basic Principles	2.1 Theoretical and Empirical Background
	2.2 Emergence	2.2 Individual Decision-Making
	2.3 Adaptation	2.3 Learning
	2.4 Objectives	2.4 Individual Sensing
	2.5 Learning	2.5 Individual Prediction
	2.6 Prediction	2.6 Interaction
	2.7 Sensing	2.7 Collectives
	2.8 Interaction	2.8 Heterogeneity
	2.9 Stochasticity	2.9 Stochasticity
	2.10 Collectives	2.10 Observation
	2.11 Observation	
Details	3.1 Initialization	3.1 Implementation Details
	3.2 Input	3.2 Initialization
	3.3 Submodels	3.3 Input
		3.4 Submodels

Table 3: Overview of the categorical differences between the original ODD framework and the ODD+D framework.

It is for this reason that the ODD + D framework seems to be most suitable as the backbone of the intended model-to-model comparison. The ROMFARMS model does however not come equipped with any model description and the Artificial Anasazi was described prior to the conception of the ODD + D and thus only comes equipped with an ODD description. To overcome this issue, the ODD + Ds of both models will be manually created for this research by consulting a variety of literary sources and the source code of both models. There is sufficient material available to do this. For the ROMFARMS model, the source code and Joyce's PhD dissertation on the model are available. For the Artificial Anasazi model, literature by the developer of the replication (Janssen 2009), the original creators of the model (Axtell *et al.* 2002; Dean *et al.* 2000; Gumerman 2003) and the source code are available. Via these sources, the author should be able provide model descriptions for all categories of the ODD + D framework.

The ODD + D comparison will consequently be performed by taking the descriptions of the same category for each model and determining differences between the model. The results will consequently be plotted the in one easily readable table. This is similar to the comparison of An *et al.* (2014), but the application of the ODD + D framework instead of the ODD framework should allow for a better comparison of ontological components between the models.

4.2 Docking agent-based models

4.2.1 Current practices

Docking refers to the practice of implementing (components of) a computational model in the simulation environment of another computational model (Carley 2002, 263). The result of docking is a variant of an existing computational model that aims to produce similar results (Arifin *et al.* 2010, 239; Carley 2002, 263). Taking the original core model and its output as the baseline for comparison, the output from the docked model can be compared with this baseline to measure its accuracy and define similarities and differences between the functionalities and outputs of the models. The process of docking and comparing the results of both models is called model alignment, since the models are executable in the same simulation environment (Axtell *et al.* 1996, 128).

Model alignment is however not a replication study. A docked model is not aimed to identically represent the baseline model. In archaeological replication studies like

Janssen (2009) and Kanters (2019), the goal was to identically replicate an agent-model using a different model-building software. These examples are two true replication studies that also aimed to reproduce the results of the same model. The research focus can however also only lie in the assessment of the relative functionality of different (components of) models, like the study of Axtell *et al.* (1996).

4.2.2 Defining the methodology

The number of studies focusing on the functional equivalence of semantically similar algorithms is non-existent in archaeology and the study from Axtell *et al.* (1996) is the only one of such studies that is known that has the same intentions as this research. This research can therefore best be taken as a guideline in the precise methodology for the docking of similar ontological similar components.

Even then, the precise nature of the implementation of components from one model into another relies on the results of the previous methodological phase. The results from the model comparison provide insight in which similar ontological components between the models are present and thus where in the models docking could take place. The best approach relies on how the source of the two models is designed, which makes the implementation tactics from Axtell *et al.* generally have little value since these only focus on the implementation of SugarScape in Axelrod's cultural transmission model. No guidelines or best practices in relation to the docking phase of the research were mentioned. The implementation therefore is a process of trial, error, and improvisation. Similar to the research of Axtell *et al.*, however, it is this process of trial and error that makes a worklog in which all the changes and alterations are noted highly relevant. Not only can the alterations to the model and their time requirements provide an increased level of model transparency, it also allows for the applied methodology to be reviewed when the implementation phase is discussed in the discussion chapter. A discussion of the improvised methodology might aid future modellers in archaeology and possibly in other sciences as well. Besides keeping a detailed work log the need for version control, i.e. consistently updating the model while saving the older versions, is relevant for consulting the changes that have been made to the models.

4.3 Verifying agent-based models

4.3.1 Current practices

Despite model testing and verification being necessary to ensure that a model functions as intended, there is no straightforward manner in which a verification process is to be approached (Ormerod and Rosewell 2006, 133). One reason of the lack of straightforward and predefined approaches in the process of model verification, is that a model cannot be completely verified or unverified. In the case of a verification procedure, models are therefore rather placed on a continuum of verification (Wilensky and Rand 2015, 325). This statement, however, touches upon an epistemological debate that crosses philosophical and computational borders in relation to the terminology of the word 'verification' and will therefore not be elaborated in this thesis. The term 'verification' applied in this thesis rather relates to the verification of computational models, focusing on the relationship between a computational model and its conceptual model. The alignment of models, as is the case in this thesis, is therefore used as an approach that allows the verification of models.

David (2006, 127) lists a set of common techniques and approaches that aid in the verification process of models. In relation to the coding and testing phase of a model, he names structured programming, model embedding and static methods as techniques via which pre-computerized models (i.e. models in the coding and testing phase) can be verified. Structured programming can refer to the modular structuring of the computer code and using an object-oriented programming language to allow a better understanding of the computer code. A helpful practice of model embedding is reusing software components whose functionality is already known and static methods refer to complete walkthroughs of the computer code. In relation to post-computerized models (i.e. computationally finished models that allow for the observation of their output), David names dynamic methods and replication or alignment as suitable verification techniques. Dynamic methods include program testing and sensitivity analyses to investigate whether the model produces the expected output relationships based on how different parameters influence each other.

The research from Axtell *et al.* (1996) has also provided a better understanding how model outputs can be used for their verification. The authors defined three types of quantitative relationships that can be used for the verification of models: 1) Numerical identity, 2) distributional equivalence, and 3) relational equivalence (Axtell *et al.* 1996, 130). Numerical identity refers to a one-on-one relationship between the distributions of model outputs (Axtell *et al.* 1996, 135). A relationship based on numerical identity is however generally not observed in agent-based models due to the incorporation of stochastic elements. A relationship between models based on distributional equivalence refers to the statistical indistinguishability of the distributions of model outputs (Axtell *et al.* 1996, 135). Often employed statistical tests that allow for the investigation distributional equivalence are the Students *t*-test, the Kolmogorov-Smirnov test, and the Mann-Whitney U test (Wilensky and Rand 2015, 324). Of these statistical tests the latter two are easier to employ because they do not imply a normal distribution of the model outputs and can be applied to continuous distributions, which generally allows for statistically testing the behaviour of the models for the duration of an entire simulation run. Relational equivalence implies a functional correlation between models based on changes in model output that are the effect of comparable changes in model inputs. These kinds of relational equivalence can be assessed with a sensitivity analysis.

4.3.3 Defining the methodology

Each model is unique, and model verification is not something that is very straightforward since verification takes place on a poorly defined continuum. A significant aspect in ensuring that a computational model functions as intended already takes place in the programming/implementation phase of the model. Since this research aims for the creation of a docked model, it is necessary to make sure that the utilized code is well-structured and easily understandable for the model implementer and that model implementations primarily rely on the reuse of existing software components. Alterations made to the code or the addition of completely new algorithms must be prevented as much as possible and can therefore only take place if these alterations or additions are necessary. The code needs to be walked through sufficiently, to serve as a first assessment of whether the internal functionality of the model is as intended.

After a docked model has been produced, it is necessary to assess whether its output is as intended. The functionality of the model can be partly observed in the interface of the NetLogo environment by directly observing patch and agent behaviour, but a sensitivity analysis will also be performed to acquire a better understanding of the functionality of the model and the influence that different parameters have on each other. Multiple simulation runs will therefore be performed with different parameters set on their minimum-, middle- and maximum values. This allows for a comparison of how different parameter combinations affect model output and how these relate to already established behaviour of the baseline model, thereby establishing whether relational equivalence between the docked and baseline model is present. Finally, the results of the implemented model and its original counterpart will be compared. The goal is to define whether distributional equivalence exists between the outputs of both models. Which exact outputs to measure depends on the choice of which case study model will be implemented in the other. The data produced will be continuous and a normal distribution between the model outputs is not assumed, this makes the Kolmogorov-Smirnov and Mann-Whitney U test the most suitable candidates to establish whether statistical similarity is observed between the model outputs.

5. The model comparison

This chapter describes the comparison between the Artificial Anasazi model and the ROMFARMS model. The chapter starts with a comparison of the models based on their descriptions in the ODD + D framework. The comparison is then followed by the general results of the comparison. At the end of the chapter a table can be found that provides a general overview of the model comparison. The ODD + Ds that were created for both models can be found in appendices 3 and 4.

5.1 Execution

The way the comparison of the models is executed, is similar to that of An *et al.* (2014). In the following sections, the comparison of the two models in relation to the elements of the ODD + D framework will be discussed. For each element, the Artificial Anasazi model will be described first and the ROMFARMS model will be discussed second. After these descriptions, similarities and differences between the models will be discussed. The comparison from An *et al.* also includes a section where modelling decisions were explained, this is however not possible for this comparison since both models are developed by other researchers.

5.1.1 Purpose

The Artificial Anasazi model is to be used as a tool to explore, explain and reproduce the causes for the settlement and population dynamics observed in the archaeological record of the agricultural communities Long House Valley between 800 and 1400 AD. The model investigates whether the environmental processes and a simple household model based on maize agriculture alone can reproduce the results.

The ROMFARMS model is a tool that generates possible scenarios for the land-use and population dynamics of different multi-faceted agricultural strategies in the Lower Rhine delta in the Netherlands between 15 BC until 270 AD. These scenarios can be used for theory-building and developing heuristics in relation to the interpretation of the archaeological record. It is aimed towards archaeologists studying Roman agriculture, as well as those studying agriculture in general.

There is no observable overlap in both the functional purposes of the models and the spatio-temporal contexts. However, both models investigate population dynamics in a system where agriculture is the only subsistence strategy.

5.1.2 Entities, state variables and scales

The Artificial Anasazi model contains one class of agents: the household. Households decide where to farm and where to settle. Household attributes contain age, collected harvest, a personal prediction of next timestep harvests, fertility age, maize storage, and a nutritional need. The location where a household is situated is the settlement, households can share the same location and thus a settlement. Each patch represents one hectare of landscape zone in the Long House Valley, whose agricultural productivity varies per timestep based on the Palmer Drought Severity Index. A timestep in the model represents one year.

The ROMFARMS model comprises three different classes of agents. The first of these agent classes is the settlement. Its attributes are the number of households, age, inhabitants, livestock, grain storage, labour costs, workforce, nutrition needs, productive land and construction works. The second class is the individual person. Its attributes are sex, age, spouse, hose, distance travelled, number of children, mortality rate, fertility rate, carrying capacity and fuel/timber processing time. The animals are the last agent class and have age, sex, species, fertility rate, survivorship, owner, and presence of milk production as their main attributes. Each patch represents one hectare and belongs to a certain landscape type. It could have an owner, produces biomass, and restores itself after being worked. Every timestep in the model represents one year.

The Artificial Anasazi model does not include agents that represent a single person or animal, but on the level of the household (Anasazi) and the settlement (ROMFARMS) the two models show many overlapping attributes. Despite the fact that a ROMFARMS settlement includes a number of households and the Artificial Anasazi household location is simply the settlement, overlapping variables include age, nutrition needs, resource storage and fertility properties. On a semantic level, however, the ROMFARMS settlement with one household and the household with its settlement from Artificial Anasazi are identical. On the level of a single landscape patch, each patch represents one hectare of a specific landscape type that yields resources/biomass. The last similarity is that each timestep in the model represents one year. The spatio-temporal dimensions of both models are thus identical.

5.1.3 Process overview and scheduling

The Artificial Anasazi is characterized by the following sequence of events per timestep (Janssen 2013, 3):

1. Calculate the harvest for each household;
2. If agent derives not sufficient food from harvest and storage or the age is beyond maximum age of household the agent is removed from the system;
3. Calculate the estimated harvest for next year based on the corn in stock and the actual harvest from the current year;
4. Agents who expect not to harvest the required amount of food next year will move to a new farm location;
5. Find a farm plot;
6. Find a plot to settle nearby;
7. If a household is older than the minimum fission age, there is a probability p_f, that a new household is generated. The new household will derive an endowment of a fraction f_{cs} of the corn stock;
8. Update water sources based on input data;
9. Household ages with one year;

The ROMFARMS model is characterized by the following sequence of events per timestep:

1. New births, deaths, and marriages within each settlement population;
2. Based on these population dynamics the creation of new, or abandonment of old, settlements;
3. Present settlements calculate grain needs, needed land to cultivate these yields, which land can actually be cultivated, total grain yield and the amount of labour needed to do so;
4. The settlements harvest grain and cultivate the land, calculation of total yields;
5. New births and deaths of animals, while also calculating slaughter rates, needed land, required labour, and yields for animal husbandry;
6. The settlements engage in animal husbandry;
7. Calculate fuel requirements and required labour to collect fuel;
8. Collect fuel;

(Every 20 timesteps, perform step 7 and 8 for Timber collection as well)

9. Update all age variables with one year, (partial) regrowth of patch variables;

These two workflows show a number of differences and similarities. The engagement in animal husbandry and fuel/timber collection does not occur in the Artificial Anasazi model, and settlement relocation or movement does not occur in the ROMFARMS model. The Artificial Anasazi model furthermore calculates possible maize yields for the following year. Both models, however, include a harvesting component and population dynamics component. This does not mean that both are completely similar. The ROMFARMS model population dynamics occur at the beginning of the new time step, whereas the population dynamics of the Artificial Anasazi occur across the model sequence. Both models furthermore incorporate marriages and death, but the ROMFARMS model is extended with births of individual persons. At the end of the process, resources are being updated and the age variable is added with one year.

5.1.4 Theoretical and empirical background

The main concept behind the Artificial Anasazi is to explore whether the application of the SugarScape model on an extensive real-world dataset would exhibit the characteristics of this real-world dataset. The model uses extensive palaeoenvironmental data of the Long House Valley in Arizona and households whose characteristics (5 persons, monoagricultural) have been derived from archaeological and ethnographic analogies. Demographics, nutritional needs, and marriage properties of the households were derived from biological anthropological and ethnographic studies. The main model for decision-making is derived from the SugarScape model but the population dynamics were slightly altered for the second version of the model.

There is no specific hypothesis or theory that the ROMFARMS model aims to investigate, it is instead a heuristic and theory-building tool (see the purpose section). Its main assumption is that Roman farms pursued a mixed farming strategy consisting of farming, animal husbandry and fuel/timber collection. This general assumption and the assumed dynamics of this mixed farming strategy and its underlying decision-models come from academic studies and expert judgment, both of which are based on data from (zoo)archaeological, palaeoenvironmental, ethnographic, economic, biological, experimental, and dynamic system model studies.

There are significant differences in the concepts behind the models, which are also related to differences in their purposes. The main theoretical difference is that Artificial Anasazi agents pursue a monoagricultural strategy and the ROMFARMS agents a mixed agricultural strategy. Both models have employed ethnographic/anthropological data and palaeoenvironmental data in the construction of their models, but the ROMFARMS model uses a much more extensive set of data for the creation of its model. This can be explained by its nature of being a theory building tool, but also the fact that the mixed agricultural strategy includes more components and is therefore inherently more complex.

5.1.5 Individual decision-making

The main subject of decision-making in the Artificial Anasazi model is the household, whose objective it is to sustain itself in its nutritional requirements throughout the timesteps. Households make decisions on where to settle and where to harvest based on mathematical calculations of predicted harvest yields for the next year and whether their current location suits these predictions. Predictions can turn out to be erroneous afterwards but follow the same calculations for each settlement. The dynamics of the simulation environment, on which these calculations are based, are influenced by exogenous factors that are identical for each simulation run.

The main subject of decision-making in the ROMFARMS model is the settlement, which can comprise one to five households. The objective of the settlement is sustaining itself in its nutritional and resource requirements as well, despite the fact that these requirements vary per settlement due to differences in their population. Agents make economically rational decisions via the outcomes of mathematical formulas related to the resource and land availability in their environment. Agents do not store information of past experiences and will not predict future yields.

The decisions made in both models are economically rational and based on mathematical programming, but the timing of these decisions and the timestep they apply to is different. The Artificial Anasazi households make economically rational decisions based on their predictions of the following timestep, whereas the ROMFARMS settlements makes economically rational decisions based on the current availability of resources and the pursued agricultural strategy.

5.1.6 Learning

The agents of both models do not engage in any form of individual or collective learning since they do not store information from previous events to influence their behaviour in future timesteps.

5.1.7 Individual sensing

The households in the Artificial Anasazi model engage in sensing when the arable land is assumed to not produce sufficient maize yields for the next year. The household senses for a new plot of arable land and a location to settle within a 1-mile radius.

The settlements in the ROMFARMS model decide after the calculations of resource requirements what the availability of these resources is in their immediate area. This relates to the calculations of grain, meadow, fuel, and timber requirements. In the case of a calculated deficit through exploitation of the current arable area, agents engage in spatial sensing by looking for new resources. The sensing radius is technically speaking unlimited, but the chosen and newly exploited land must always be connected to the currently cultivated area. The sensing for meadow land and biomass for fuel/timber collection also has an unlimited radius and is executed every timestep.

The agents from both models thus have opportunities to engage in spatial sensing. Sensing in the Artificial Anasazi model is however not necessarily executed in each timestep since this only occurs when predicted harvests do not meet the nutritional requirements. The ROMFARMS model, on the other hand, has its agents engage in individual sensing every timestep.

5.1.8 Individual prediction

The households in the Artificial Anasazi model predict the future harvest yields by assuming that it is equal to that of the previous timestep. The ROMFARMS model does not imply any form of prediction by the agents. There are thus no similarities.

5.1.9 Interaction

Even though the Artificial Anasazi model assumes interaction via marriages and the potential presence of two agents on the same plot, this interaction is not explicitly executed in the source code. A marriage is implicitly executed as one household sprouting a new household based on a random probability and multiple households on the same plot do not interact with each other. True interaction is therefore not present.

The ROMFARMS settlements similarly do not engage in any form of interacting, except for the borrowing of surplus grain yields from other settlements in the case of an encountered deficit. This process is however regulated via mathematical calculations and therefore has no social or cooperative consequences for the household in further timesteps.

5.1.10 Collectives

The Artificial Anasazi model does not explicitly model collectives. Households are individually operating agents; it is only when two agents are located on the same patch that a settlement is a collective of households. This has however no further implications.

The ROMFARMS model has collectives of agents that form individual households and comprise the inhabitants of a settlements. Animals can be divided in herds that belong to different settlements. No similarities between the models are thus observed.

5.1.11 Heterogeneity

The households in the Artificial Anasazi model are heterogeneous in terms of their age, location and collected or stored harvest. Their shared characteristics are their nutrition needs, length of corn storage, fertility age, death age, corn stock given to a newly generated household and their maximum distance between a residence and farm. Despite their heterogeneous characteristics, the agents are homogeneous in terms of their decision-making.

All agent classes in the ROMFARMS model are heterogeneous. Settlements are heterogeneous in terms of number of households, inhabitants, livestock, resource yields, storage and worked land for different agricultural purposes. Persons differ in terms of sex, age, house, spouse, fertility rates, mortality rates and number of children. Animals differ in terms of age, sex, species, and owner. Agents are however not heterogeneous in their decision-making and will all follow the same sequences of procedures.

The ROMFARMS model thus includes a more extended set of variables in which agents can differ from each other compared to the Artificial Anasazi model. The agents from both models share heterogeneous characteristics in terms of their age and resources. Despite their heterogeneous characteristics, the agents in both models follow homogeneous decision-making processes.

5.1.12 Stochasticity

The stochastic elements in the Artificial Anasazi model are expressed in the initial conditions of agents and the landscape and the order in which agents perform their behaviour. Furthermore, stochasticity is implemented in relation to the odds that determine the occurrence of settlement marriage/reproduction.

The ROMFARMS model has random probabilities in relation to population dynamics (fertility, mortality, and marriage), it has a randomly generated landscape and the resource contents of the landscape have a randomized additional value that is added to the baseline values.

The Artificial Anasazi model thus primarily starts from a randomized initial state of the model and updates itself in a random manner as well. The ROMFARMS model also has randomized initial conditions in terms of the placement of households with an additional random generated landscape. The updating of agents is randomized, and alongside this are additional random probabilities that affect the population dynamics.

5.1.13 Observation

The Artificial Anasazi model produces data that represents the number of households and its contents at each timestep, which combined shows a pattern of population dynamics and trends. The model can also generate maps that show the locations of settlements throughout the simulation. Both the population development and settlement locations are emerging phenomena that can be compared to the archaeological record.

The ROMFARMS model produces quantitative data that relates to the composition of settlement demographics per timestep, resource consumption and storage levels, livestock constitution, and the labour requirements for executing the mixed agricultural components. Together these data aggregate to a pattern of land-use and resource consumption for different agricultural strategies in various kinds of environments.

The ROMFARMS model is thus able to produce a more elaborate set of data than the Artificial Anasazi model, but both provide outputs regarding the development of the population.

5.1.14 Implementation details

Both models have been implemented in the NetLogo software. The Artificial Anasazi model has been implemented in version 4.0.2 and the ROMFARMS model has been implemented in version 6.0.2.

The Artificial Anasazi model was used in an earlier version of NetLogo, meaning that there could be discrepancies between the intention of the source code and its interpretation by NetLogo's most recent software version. ROMFARMS has been developed in the most recent version, meaning that possible functional discrepancies are not expected.

5.1.15 Initialization

The Artificial Anasazi model starts with 14 households with a random age and random filled storage levels. The landscape patches vary in initial soil quality per simulation rum based on a normal distribution of the soil quality. The ROMFARMS model does not come equipped with any filled in parameters. Initialization is therefore up to the user, without this initialization of parameters the ability to run the simulation is also prevented.

The Artificial Anasazi model is workable from the beginning with the parameter combination that best reflected the archaeological record initialized. The ROMFARMS model is not executable at all at the moment of initialization, the user thus first has to manually enter values for all parameters.

5.1.16 Input data

The Artificial Anasazi model employs five different text (.txt) files which contain different kinds of data: the map that defines all landscape types per patch, the Palmer Drought Severity index (moisture conditions per cell/landscape type), the environment (defines the relationship between a patch and its landscape type and the presence of a water source), the settlement (historical/archaeological data along which results can be compared) and the presence of water sources throughout the landscape.

The ROMFARMS model can be applied to specific sets of GIS data that represents different palaeoenvironments throughout the Lower Rhine delta, but this data is not provided when downloading the model from its online source and is therefore not operable for external parties.

The Artificial Anasazi model thus comes accompanied by several data sources that are being accessed at the moment of initialization, whereas the ROMFARMS model in this state does not apply external data and consequently only generates random landscapes.

5.1.17 Submodels

The Artificial Anasazi model description does not specify a set of submodels, but the description of this section elaborates on the calculations of cell base yields, harvests, and harvest predictions as well as the process for maize consumption, and the search for new settlement locations.

The ROMFARMS model consists of four submodels: 1) the population submodel; 2) the arable farming submodel; 3) the animal husbandry submodel and 4) the fuel/timber collection submodel. The population submodel describes the sequences of birth, death, and marriages. The arable farming submodel describes the calculations of nutritional needs, available land, harvest procedure and grain consumption. The animal husbandry model includes birth and mortality procedures similar to that of the population dynamics model, with added secondary products calculations and slaughter calculations. The fuel and timber collection models are identical in their procedures. It starts with determining needed fuel/timber resources, where these can be collected and consequently sends the required possible workforce to gather the required amounts of fuel and timber.

Of these different submodels, only the arable farming submodel of ROMFARMS shows overlap with the aspects of Artificial Anasazi that are described in more detail. They overlap in the sense that both focus on calculation of achieved harvest and the execution of a consumption process.

5.2 Conclusions from the comparison

When looking at the comparison results, which are visualized in table 4 at the end of this paragraph, it is probably best to first note that both models employ the same spatio-temporal dimensions. One landscape patch represents one hectare and one timestep represents one year in both models. This means that despite differences in the way the simulation environment/landscape is designed, what the landscape represents and what its dynamics are, the spatio-temporal contexts of the agent processes have been designed on a similar scale and can consequently be compared as such. When trying to establish how aspects of a model can be docked in another model, consequences of differences in spatio-temporal dimensions and necessary scale/resolution adaptations would first need to be thoroughly considered before moving to this implementation phase.

In relation to the ontological components of both models, the first thing to note is that there is one entity that is present in both models: the household. Despite the fact that ROMFARMS can include a multiplicity of households per settlement, a settlement with only one household represents exactly the same semantic ontological entity as the household with settlement in the Artificial Anasazi model. A docking process could thus take place on the level of this household. These households share characteristics in relation to their main objective, which is the sustainability of resource consumption, but are also engaged in systems of population dynamics and resource harvesting. Any quantitative comparison can thus primarily be made in relation to the output of these systems. It needs to be noted that the execution of these systems in the Artificial Anasazi model is relatively simplistic compared to the ROMFARMS model. This simplicity definitely limits the amount of similarity between the models. ROMFARMS is more elaborate in relation to the population dynamics and resource harvesting, which is primarily related to the presence of agents that represent persons. This allows for significantly higher degrees of complexity, heterogeneity, and diversity in relation to the population constitution and therefore its dynamics and behaviour.

The ROMFARMS model furthermore includes animal husbandry and fuel/timber collection practices, but the relative simplistic nature of the Artificial Anasazi model prevents any similarities to be found between the models in this regard. When attempting an implementation of the Artificial Anasazi model in ROMFARMS, the relation of the implemented aspects with these submodels would also need to be thoroughly considered.

To summarize the results of the model comparison, it can be concluded that there is significant overlap between the ontologies of the models. From the perspective of the Artificial Anasazi model, whose decision-making system is relatively simplistic, the degree of similarity is high due to the fact that ROMFARMS matches all the semantic concepts that underlie this decision-making system. From the ROMFARMS perspective the degree of similarity is limited, due to the fact that this model is more elaborate and incorporates an extensive set of submodels on whose level the Artificial Anasazi model is not able to fully match. This overlap is however still not insignificant, and the docking of models should be possible for ROMFARMS in Artificial Anasazi and vice-versa.

Element	Category	Artificial Anasazi	ROMFARMS	Similarities
1. Overview	1.1 Purpose	Explore/explain causes for Long House Valley settlement dynamics between 800 and 1400 AD.	Provide scenarios for the impact of different agricultural subsistence strategies on land-use and population dynamics.	Both models investigate the influence of agriculture on population dynamics.
	1.2 Entities, state variables and scales	Agents: Households, Patches: one hectare areas belonging to a certain landscape types, Timestep: One year	Agents: Settlements, persons and animals, Patches: One hectare areas belonging to a certain landscape type, Timestep: one year	Settlement with one household are semantically similar. Area sizes, area representation and dimension of each timestep.
	1.3 Process overview and scheduling	Initiation of agents and environment; households harvesting maize; household movement and demographic factors.	Initiation of agents and environment; population dynamics; arable farming; animal husbandry; fuel and timber collection.	Population dynamics and resource harvesting
2. Design concepts	2.1 Theoretical and Empirical Background	Household properties derived from archaeological, biological anthropological and ethnographical examples. Decision-making system derived from Sugarscape model. Landscape dynamics derived from palaeoenvironmental data and Palmer Drought Severity Index (PDSI) correlations.	Settlement mixed agricultural strategy dynamics and landscape dynamics derived from a combination archaeological, palaeoenvironmental, ethnographic, biological experimental and dynamic system model data.	Reliance on palaeoenvironmental, ethnographic, anthropological data.
	2.2 Individual Decision-Making	Household main actor in decision making, objective is nutritional sustainability. Makes economically rational/optimal decisions based on predicted harvest yields for the next year. Harvest yields depend on exogenous factors.	Settlement is main actor in decision-making, objective is nutritional and resource sustainability. Agents make economically rational/optimal decisions based on calculations of own needs and area productivity.	Objective of nutritional sustainability, economic rationality underlying decision-making.
	2.3 Learning	Agents do not engage in learning.	Agents do not engage in learning.	No engagement in learning
	2.4 Individual Sensing	One-mile radius when searching for new farmland.	Sensing is engaged in relation to the own production area and abroad, with an unlimited radius.	Both models employ some form of sensing, but the scales on which sensing occurs are different.
	2.5 Individual prediction	Agents predict the harvest for the next year by assuming that it is the same as the last year.	Agents do not engage in prediction	No similarities
	2.6 Interaction	Agents do not directly interact. Interactions like marriage/fertility/reproduction are modelled implicitly.	Agents are able to borrow grain from another settlement in the case of a deficit, this has no social or cooperative consequences by creating unbalanced relationships.	No similarities
	2.7 Collectives	Agents can occupy the same patch, but a shared settlement is not explicitly modelled.	Households (two adults and four subadults) and herds (all animals that belong to a settlement). Both are not modelled explicitly.	Collectives only consist on a conceptual level and have not been modelled explicitly.
	2.8 Heterogeneity	Households: age, location and collected and stored harvest.	Settlements: Location, number of households, inhabitants, livestock, resource yields, resource storage and labour forces. Persons: Sex, age, spouse, fertility, mortality, settlement and children. Animals: Sex, species and owner.	Households and settlements share heterogeneous characteristics: Location and collected and stored harvests.
	2.9 Stochasticity	Initial agent locations and landscape conditions, order in which agents are updated	Initial settlement locations and landscape conditions, population dynamics (fertility, mortality and marriage) for persons and animals (fertility, mortality and slaughter)	Random initialization of settlement/household locations and landscape conditions.
	2.10 Observation	Population dynamics and settlement distributions	Settlement demographics and dynamics, resource consumption and storage levels, livestock composition and labour force composition.	Population dynamics and settlement dynamics. Resource storage levels are theoretically also possible.
3. Details	3.1 Implementation Details	Netlogo v4.0.2.	NetLogo v6.0.2	Both implemented in NetLogo, but not in similar versions.
	3.2 Initialisation	14 households with a random age and storage levels.	Model not directly functionable, parameters have to be manually filled in.	No similarities
	3.3 Input Data	Textfiles for map, environmental relations, water sources, PDSI and archaeological/historical data.	No external input data. GIS files were used in original research, but do not come equipped with the model.	No similarities
	3.4 Submodels	No submodels, but elaboration on calculation of base yield, harvest yield and harvest prediction as well as maize consumption process and search for new locations.	Population dynamics, arable farming, animal husbandry, fuel and timber collection.	No similarities

Table 4: Results from the model comparison.

6. The Docking phase

The previous methodological phase provided results on similarities between the ontological components of the Artificial Anasazi model and ROMFARMS. This allows for implementing these components from one model into the other, with a docked (implemented) model as a result. This chapter describes the docking process where components from the Artificial Anasazi model are implemented in the ROMFARMS model. The first section of this chapter describes how the defined similarities and differences are being used to determine how the model-to-model implementation can take place on a conceptual level. The second paragraph consequently focuses on the execution of the model-to-model implementation on a technical level. The third part discusses the docked model, the result of the docking process.

The work log of the alterations and the ODD + D that describes the implemented model can be found in appendices 5 and 6. This work log follows the same layout as Axtell and Axelrod have provided in Axtell *et al.* (1996) since it makes both the alterations to the model, the sequence in which these alterations have taken place and the time intensity of these endeavours clear.

6.1 The implementation from a conceptual perspective

In the previous chapter the areas have been defined in which the Artificial Anasazi and the ROMFARMS model portray sufficient similarity for a possible docking procedure. It was concluded that similarity was found in the spatio-temporal dimensions of both models as well as the agents that represents a settlement with one household. These similarities can form a solid basis for a procedure of model implementation, since it indicates that the spatio-temporal scale on which behaviour occurs is similar and that the properties of agents are directly transferable.

In consultation with the thesis supervisors, it was decided to implement the household agent and its correlating demographic dynamics from the Artificial Anasazi in the ROMFARMS model. The main reason for this was that an implementation of the Artificial Anasazi in ROMFARMS would be able to define whether the already existing algorithms and ontological components of Artificial Anasazi could have been used in the development of ROMFARMS and what influences this would have on the model output. Another reason was the fact that Artificial Anasazi is a model that is already generally well understood by the archaeological modelling community, whereas ROMFARMS is in this sense more of a black-box model. Finally, the Artificial Anasazi model has been developed in an older version of NetLogo. This older version brought a risk that it would function slightly different in a newer version of NetLogo, making the results of a statistical comparison possibly less reliant.

From a conceptual point of view, this kind of implementation implies that the simulation environment of the model should be the regular ROMFARMS environment, with no additions of new landscape components or removals of existing ones. This environment will consequently be inhabited by a set of households whose general characteristics and demography dynamics have been derived from the Artificial Anasazi model. This is a more simplistic execution compared to the households in the ROMFARMS model.

Contrary to the ROMFARMS model, the households from the Artificial Anasazi model have implicitly modelled inhabitants. A household consists of a male, female, and four subadults whose age, nutritional requirements, and contribution to workforces do not change during the simulation. This implicit modelling of a household population implies the removal of the 'persons' agent-class from ROMFARMS. Each household would furthermore act individually and independently, whereas a settlement in ROMFARMS can comprise multiple households that all behave according to the rules determined by the settlement.

This does, however, not mean that the entire suite of programmed behaviour of the Artificial Anasazi household can be fully implemented in ROMFARMS. ROMFARMS employs a series of fundamental assumptions in relation to agent characteristics, agent behaviour and agent-landscape relationships. Altering these fundamental assumptions can have drastic effects on the model performance, even to an extent that the docked model might not sufficiently represent the original and can consequently not be verified. Examples of some the fundamental assumptions and relationships in the model that are different in the Artificial Anasazi model but that may not be altered are the sedentary nature of the household, rules for calorie and grain requirements, the effect of sowing/ploughing on the landscape, resource regrowth and the engagement in animal husbandry and fuel/timber collection. This means that alterations to the Artificial Anasazi household need to ensure that characteristics and dynamics of the removed 'person' agent class are still executable. It is primarily in these fundamental assumptions, which relate to processes that are not present in the Artificial Anasazi model, that modelling decisions need to be made and that solutions need to be found. These decisions and modelling solutions need be expressed as transparent as possible to ensure that verification of the model is possible.

This implementation is a simplification of the ROMFARMS agents and the demography submodel, but the implementation is still assumed to have a significant depth. The household, its inhabitants and their demographic development form the basis of all executed behaviour across the different submodels and any alterations to this core structures will also have an impact on all aspects of the ROMFARMS model. All submodels must therefore be adapted to suit this new core structure.

6.2 The implementation from a technical perspective

This paragraph discusses the technical execution of this implementation[1]. In doing so, it will first discuss the initial state of the ROMFARMS model and preparations before alterations to the ROMFARMS source code were made. It will then discuss alterations to the setup procedure and the different submodels that constitute the workflow of the model as a whole. When referring to a procedure in the source code of the implemented model or ROMFARMS, this procedure is expressed in *italics*. When referring to a variable, this variable is expressed in **bold**.

6.2.1 Initial state and preparation of the model

The first step in the implementation process has been a thorough examination of the interface of the ROMFARMS model and the re-structuring of the ROMFARMS source code, with the goal to better understand the internal functionality of the model. As described in section 3.2 (Initialization) of the ROMFARMS ODD + D, the ROMFARMS interface comes equipped with empty parameters values rather than initialized parameters. This means that the model is not immediately functional when the NetLogo file has been opened and that one first has to acquire knowledge on relevant parameter values to set the model up. An example of working values is provided in the form of a .jpg image that accompanies the downloaded model. An extensive set of total employed parameter ranges for the applied experiments and the entire model can be found in Joyce’s (2019a) dissertation. An examination of the source code made clear that ROMFARMS does not have its source

---

[1] To acquire a better understanding of the technical implementation, it is advised to simultaneously consult the docked model that is the result of this implementation.

code clustered per submodel, but rather per level of specificity. ROMFARMS, like the most NetLogo models, includes two major procedures that together allow simulation runs for the model: the *setup* procedure and the *go* procedure. The first is used to generate the simulation environment and the second to execute the simulation. These two procedures consist of a set of commands that specifically apply to the given procedures as well as sub-procedures that function on a deeper level within the workflow of the model. It is in relation to these levels of functional procedural depth that the ROMFARMS source code is clustered.

Level1	Level2	Level 3	Level 4	Level 5
Setup	Setup-landscape	Setup-forest		
	Setup-globals			
	Setup-agents	Config-settlements		
Go	Go-demography	Update-settlements-1		
		Births	Set-fertility	
		Deaths	Set-mortality	
		Marriages	New-settlement	Config-settlements
		Update-settlements-2		
	Go-arable	Calculate-land-arable		
		Calculate-yield-arable		
		Calculate-surplus-arable		
	Go-pastoral	Set-mortality-rates		
		Births-animals		
		Natural-mortality		
		Slaughter		
	Go-fuel	Calc-workforce-fuel	Find-fuel	Take-fuel
	Go-construction	Cald-workforce-construction	Find-timber	Take-timber
	Update-globals			
	Update-patches			

Table 5: Overview of the structuring of the initial ROMFARMS code (by column from left to right) and after re-structuring (per row from top to bottom).

Two major preparatory decisions were made in relation to the interface and the structuring of the source code to increase the user-friendliness of the model for the model-to-model implementation: The replacement of the manual parameter boxes by parameter sliders to ensure immediate model functionality upon initialization and the structuring of the source code so that it would be clustered per submodel. Table 5 provides a schematic overview of the structuring of all the procedures that comprise the ROMFARMS source code and how the code was re-structured. Reading this table per column from left to right illustrates how the procedures that comprise the ROMFARMS code were originally structured, whereas reading this table per row from the top to the bottom illustrates how the structure of the procedures has been altered to ensure a grouping of the source code per submodel.

6.2.2 Setup of the simulation environment

The entire setup of the simulation environment consists of two parts in ROMFARMS, the definition of agents and global variables, and the setup procedure. The first part focuses on the creation of the different agents, agent variables, global variables, and patch variables, whereas the latter focuses on the initialization of the simulation environment.

ROMFARMS comes equipped with three agent breeds: settlements, persons, and animals. The implemented model utilizes the household structure from Artificial Anasazi, which led to the creation of a new household breed that includes all variables that earlier belonged to both the settlement and person breeds. The settlement and person breeds have been removed in the implemented model, and settlements and persons are implicitly modelled in the docked model. Relevant demographical, global and agent variables from Artificial Anasazi, including **DeathAge**, **FertilityAge, FertilityEndsAge**, **HouseholdMaxInitialAge** and **HouseholdMinInitialAge** have been added to implement its demography system. No alterations have been made to the variables of the animal agent variables and the patch variables.

As is visible in table 5 , the setup procedure consists of the *setup-landscape*, *setup-globals* and *setup-agents* procedures. The first of these three procedures, *setup-landscape*, focuses on the initialization of the patches and thus the simulation environment itself. Except for the removal of the initialization of the unavailable GIS-data, no alterations have been made to its source code or its sub-procedure *setup-forest* to ensure the same initialization of random environments and thus also possible agent-environment relationships. The second procedure, *setup-globals*, defines the values of global variables. These global variables are significant in the sense that they primarily include relevant information about environment agent properties, like quantities of landscape yields, household labour times and animal yields and mortality rates. All of these variables define aspects relevant to the initial agent-environment relationship and their variables have thus not been altered.

Three global variables have been added from Artificial Anasazi, namely the aforementioned **DeathAge**, **FertilityAge**, **FertilityEndsAge**, **HouseholdMaxInitialAge** and **HouseholdMinInitialAge**. Most alterations have taken place in relation to the last procedure, *setup-agents*. In ROMFARMS, this procedure lets different patches generate a number of one- or multiple household settlement agents which consequently create their own individual inhabitants. The code of this procedure has been altered to let different patches create the total indicated number of households of a random age (similar to Artificial Anasazi) and consequently configure these settlements via the *config-settlements* sub-procedure.

The major alteration of this procedure has been the definition of required calories by a household. ROMFARMS assumes a household with one adult male, one adult female and four subadults. These are defined as person agents and their caloric requirements will vary as they age. Since persons are modelled implicitly, households in the docked model have a static but similar household composition in the implemented model. This means that calorie requirements stay the same per household throughout the simulation. Furthermore, two new agent variables have been added: **my-plot** and **construction-multiplier**. These variables will be explained in depth when discussing the fuel and timber collection submodels.

The final alteration is one to the *setup* procedure in general. The interfaces of the ROMFARMS and the implemented model have a possibility to define the size of flood basin and levee areas, but it is possible that the sum of the parameter values exceeds the number of available patches in the simulation environment. A stopping command has therefore been implemented in the setup procedure of the docked model to prevent the model from providing an error when this occurs, pure for the user-friendliness of the model.

6.2.3 Go procedure

The *go* procedure, whose workflow is illustrated in figure 8 , can be called upon after the simulation environment has been initialized via the setup button in the interfaces of ROMFARMS and the implemented model. Whereas the setup button calls upon the execution of the *setup* procedure, which has been discussed in the previous section, the go button executes the *go* procedure. This procedure is executed each timestep and thus allows the model to portray its behaviour over a given amount of time.

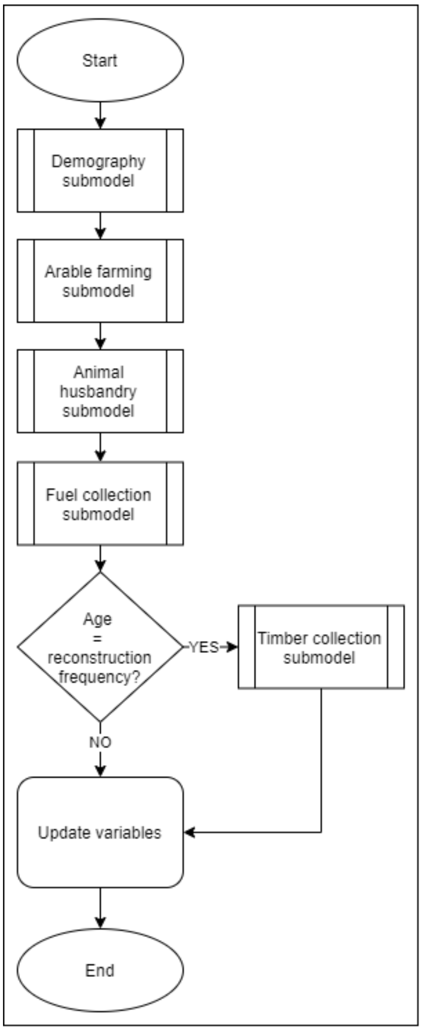


Figure 8: Schematic overview of the workflow in the go-procedure of the docked model (after Joyce 2019a, 24).

The sub-procedures that comprise the *go* procedure are each related to the different submodels indicated in the submodels section of the ROMFARMS ODD + D. They are specified by the label *go-* and the name of the submodel. After the execution of the different submodels, the *go* procedure calls upon the *update-globals* and the *update-patches* procedures to update the global and patch variables after different submodels have executed different kinds of agent behaviour. The general workflow of the *go* procedure has not been altered except for the requirement that allows for the execution of the timber collection model. In the docked model the timber collection submodel is executed once the age of a household, and not the number of a timestep, reaches the value of the **reconstruction-frequency** or a multiplication of it.

6.2.4 Demography submodel

The first submodel that is being called upon in the *go* procedure is the demography submodel. The workflow of this submodel in ROMFARMS is visualized in figure 9. In ROMFARMS this submodel consists of the *update-settlements-1*, *births*, *deaths*, *marriages,* and *update-settlement-2* procedures. The *update-settlements-1* procedure updates the values for calorie, grain, and fuel requirements. Despite its placement in the schematic overview of the workflow after the *births* procedure, the ROMFARMS source code executes the *update-settlements-1* procedure before the *births* procedure. This seems to be a general mistake in the structure of the ROMFARMS code even prior to the re-structuring of the source code, which implies that calorie requirements and the consequential resource consumption of new-born persons and the extra calorie consumption of lactating females is not being calculated for the given timestep.

In the docked model this structuring issue is not present because the *births* procedure is a procedure that specifically focuses on the persons breed, which is not present in the docked model. Since persons are modelled implicitly in the implemented model and the demographical constitution of a household is static, the *births* procedure has been completely removed. The demography submodel thus starts with the execution of the *update-households-1* (the name has been changed "households" to prevent confusion) procedure, which now only focuses on the update of grain store levels based on household consumption since calorie and fuel requirements are static due to the static composition of a household population.

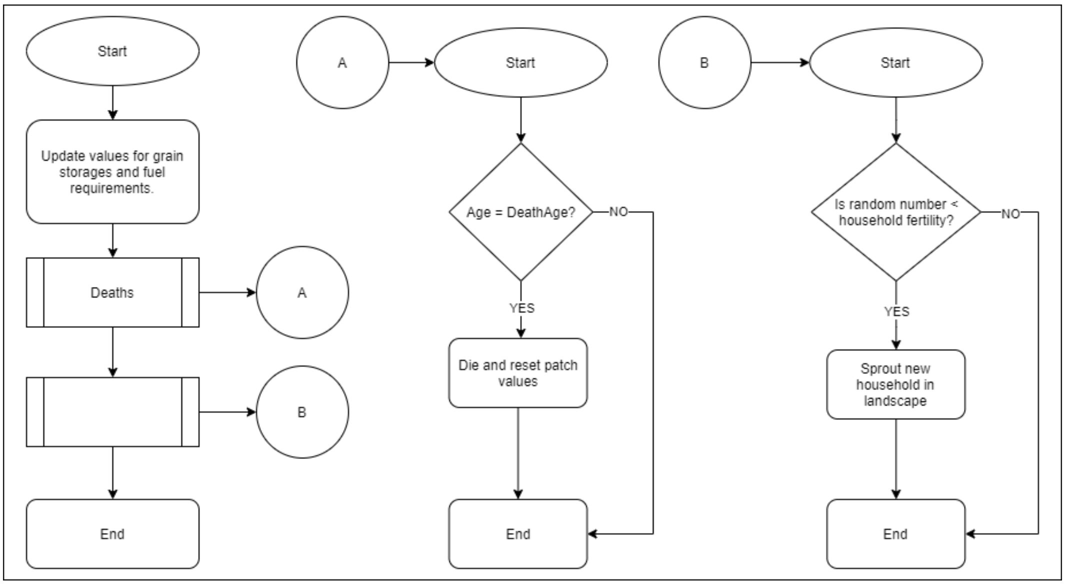


Figure 9: Schematic overview of the workflow in the go-procedure of the docked model (own figure).

The update of the grain stores for all settlements is followed by the *deaths* procedure, which regulates when a household leaves the simulation environment. In ROMFARMS, deaths are being regulated on the person level by first setting mortality rates for each age group and the occurrence of death when a random number between one and zero falls beneath this mortality rate. Artificial Anasazi and the implemented model come accompanied with the **DeathAge** variable in the model interface, a household is removed from the simulation once it has reached this age of death. Once a household has been removed, its associated livestock dies as well and its associated landscape variables are being reset. The algorithms for the death of livestock and the resetting of the landscape originate from the *update-settlements-2* procedure, which is removed in the implemented model since the other commands of this procedure focuses on the transfer of person agents when no adults are present in a settlement. The Artificial Anasazi model furthermore implements a death rule where households die when they are in a nutritional deficit. This rule however contradicts the fundamental rule in ROMFARMS which states that households cannot die by a nutritional deficit. If a household experiences a deficit, its deficit can be compensated with the surplus of another household and otherwise the household is reported to have a famine. Since the compensation for a deficit does not have an impact on the demographical dynamics as a whole, the **deficit-compensation** variable has been added via an on-off switch. This allows to investigate a nutritional deficit under cooperative household conditions (following ROMFARMS rules) or under non-cooperative conditions (following Artificial Anasazi rules).

The *marriages* procedure in ROMFARMS couples a female and male person from different households within the same settlement or between settlements and makes them form their own household. When a settlement does not have any space for more households, the *new-settlement* procedure is called upon to create a new settlement. This procedure provides the household with its basic characteristics and calls upon the *config-settlements* procedure that is also utilized in the *setup-agents* procedure to configurate the settlement. Artificial Anasazi does not directly include a marriage system, but rather implements a system where a new household can fission from another household. In both Artificial Anasazi and the docked model, this is regulated in the *fissioning* procedure that utilizes the **household-fertility** parameter slider in the interface. When the age of a household falls between the values of the **FertilityAge** and **FertilityEndsAge** variables there is a chance, based on a random number between one and zero that must be below the **household-fertility** value, that a new household is created via the *new-household* procedure. Each household has the opportunity to create a new household, since there are four subadults per household. This procedure is similar to the *new-settlement* procedure and also utilizes the *config-settlements* procedure for the configuration of the newly established household.

6.2.5 Arable farming submodel

The arable farming submodel, whose workflow is visualized in figure 10, is executed via the *go-arable* procedure. This procedure consists of the *calculate-land-arable*, *calculate-yield-arable,* and *calculate-surplus-arable* procedures. These three procedures each perform a set of calculations that are fundamental in determining which areas the settlement in ROMFARMS and the households in the implemented model can use for farming, what their yield is and how these yields affect settlement or household grain storages. Artificial Anasazi has its own procedures with calculations for the collection, harvesting and consumption of resources, but these procedures are all related and unique to the agent-environment relationships for that specific model. The parameters and variables used for these calculations are thus specific to that simulation environment and can therefore not be transferred to a different simulation environment.

As a result, the only adjustments that have been made to the source code is related to the calculation of the labour forces required for performing cultivation activities in the simulation environment. Where these calculations in ROMFARMS used variables related to the count of persons in a settlement and their relative age groups, they have been replaced by values representing the static constitution of a household. Variables that relied on the count of adult persons have been replaced by the value ‘2’ (one male and one female adult) or otherwise ‘6’ (adults and subadults).

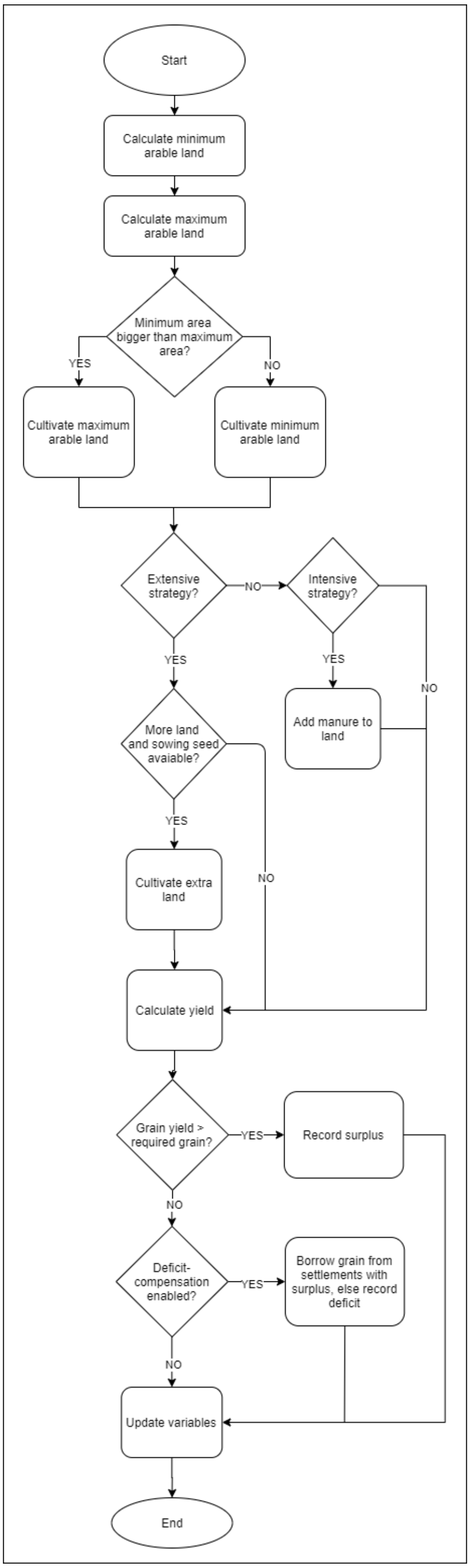


Figure 10: Schematic overview of the workflow from the farming submodel, it is similar between the docked model and ROMFARMS (after Joyce 2019a, 32).

The *calculate-arable-land* procedure focuses on the calculation of the minimum land that needs to be cultivated to sustain each household. The minimum area of land that needs to be cultivated is calculated as followed (formula adapted from ROMFARMS source code):

$$L = \frac{s}{y} + \frac{\left(\frac{s}{y} \cdot r\right)}{y}$$

In this equation L is the minimum area of land that needs to be cultivated (in number of hectare/patches), s is the minimum grain store size (in kg), y is the base yield of grain per hectare and is a global variable defined at 1000 kg/ha. r is the sowing rate of kg per hectare, it is a global variable defined at 200 kg/ha. s represents the minimum grain store size of a household in kg, which is calculated as followed (formula adapted from ROMFARMS source code):

$$s = x\frac{vp}{c}$$

In this equation x is the store size, which is a parameter in the model interface with a range of 1 to 2 with 0.25 increments that allows for the creation of an extra buffer in the minimum grain store. p represents the percentage of calories derived from crops and is a parameter in the model interface with a range from 0 to 1 with 0.1 increments. c is the number of calories in crops and is defined as 3100 kcal/kg. v is the calorie requirement per household and is calculated as followed (Joyce 2019a, 32):

$$v = t\sum(N_{cat}D_{cat})$$

In this equation t stands for the number of days in one year and thus equals 365. N is the number of inhabitants per age category (cat) and D is the calories required per age category (cat). With an average daily calorie expenditure of 3250 kcal for an adult male, 2800 for an adult female and an average of 2585 for a male and female subadults, this brings the value of v to 5.982.350 kcal for one household.

After the calculation of the minimum number of required land, different maximum numbers of cultivatable land are calculated in relation to the available land that fulfils the cultivation criteria, the availability of labour and the availability of sowing seed. If the minimum number of required land exceeds the maximum land, the maximum land is exploited and else the minimum required land is exploited. Based on **Intensification** or **Extensification** arable strategy parameters selected in the model interface, the arable land will be manured to generate extra yields or extended as much as possible to generate extra yields.

The *calculate-yield-arable* procedure calculates the yield for each household, which varies per household. Yields are calculated via the following equation:

$$o = a\big(y + (1 + f)\big)$$

o is the yield of grain in kg, a is the arable land in ha/patches that has been selected in the previous procedure, y is again the base yield of a patch in kg/ha and f is a factor that assumes a random value in a range from - 0.2 to 0.2 to ensure variability in yields between households. If the **Intensification** arable strategy is enabled, extra yields are calculated based on the **nitrogen-content-manure** and **yield-increase-manure** variables. These are predefined global variables respectively set at 6 kg per ton of manure and 15 kg per kg of nitrogen. The *calculate-surplus-arable* procedure calculates whether the combination of yields and existing grain store levels lead to a grain surplus or deficit. In the case of a deficit, a household deficit can be compensated for by other households if the **deficit-compensation** switch in the model interface is enabled. If a deficit is reported, the household is consequently reported as having a famine.

6.2.6 Animal husbandry submodel

The animal husbandry submodel is executed via the *go-pastoral* procedure and consists of the *set-mortality-rates*, *births-animals*, *natural-mortality,* and *slaughter* procedures. The ROMFARMS workflow of the model is visualized in figure 11. The *go-pastoral procedure* fully focuses on the demographic dynamics and the land and labour requirements related to the animal agent breed. Since the animal breed is not a shared ontological component between the Artificial Anasazi and ROMFARMS model and the calculations for the required pasture land of the animals are inherent and fundamental to this simulation environment, the only alterations to the algorithms of these models have been made in relation to the labour requirement calculations. Similar to the arable farming submodel, a required count for the number of adults has been replaced by the value ‘2’ where necessary, since this is the static number of adults per household.

The individual animal agent can belong to three different kinds of species: sheep, cattle, or horses. The presence of these species in the model is regulated via the on-off switches **sheep?**, **cattle?** and **horse?**. Each species can be divided in three age categories, young, immature, or adult, the precise age ranges of these categories vary per species. The *set-mortality-rates* procedure sets the mortality rates for all species and all age categories. The *births-animal* procedure uses a stochastic mechanism, where a random number between 0 and 1 that falls under the animal fertility values defined in the *config-household* procedure causes the creation of a new neonate animal of the male or female

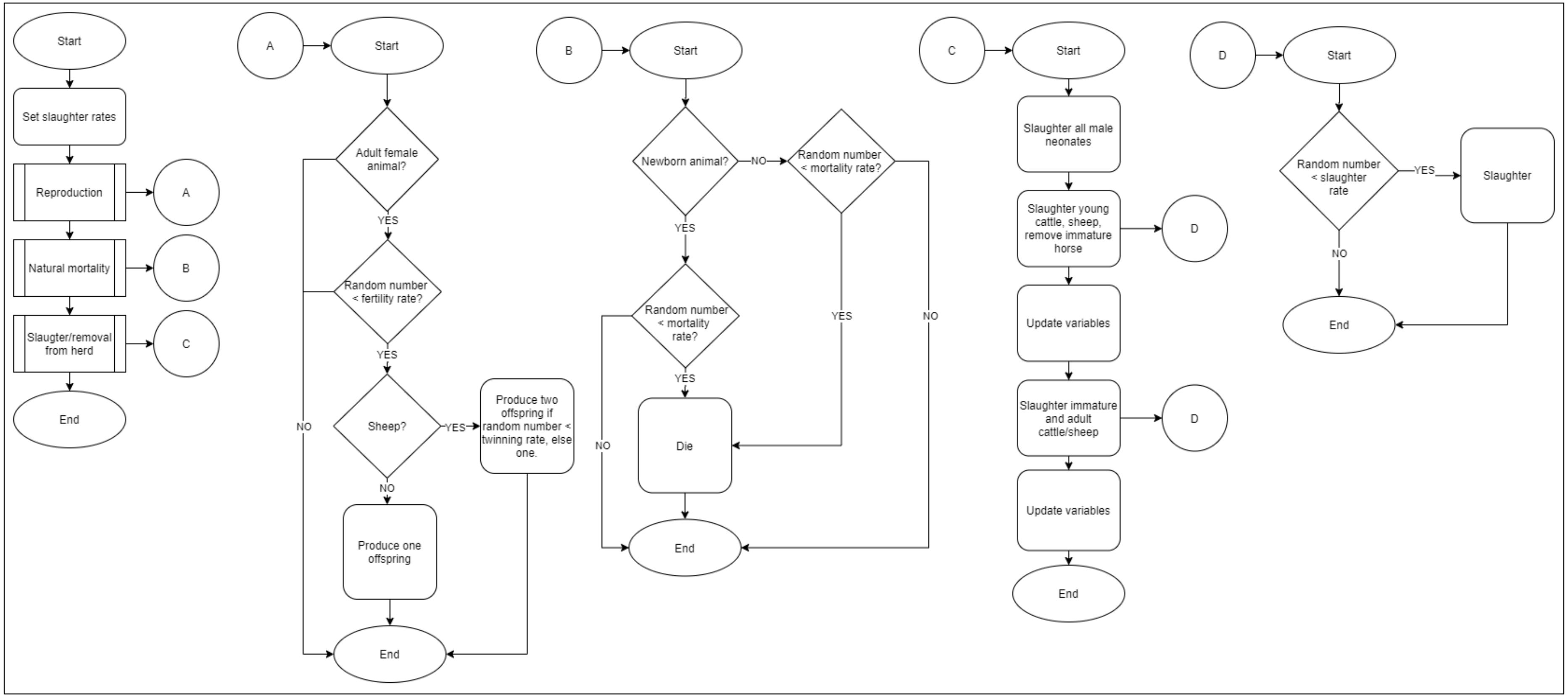


Figure 11: A schematic overview of the workflow of the animal husbandry submodel (after Joyce 2019a, 39)..

sex. The *natural-mortality* procedure does not utilize the mortality rates indicated in one of the previous procedures but has predefined mortality rates for neonate animals and older animals. The *slaughter* procedure uses the mortality rates from the *set-mortality-rates* procedure to define the slaughter of animals via the same stochastic mechanism that regulates the births of the animals. It is also in this procedure that the sum of the resource yields from the animals, which depend on the **sheep-strategy** and **cattle-strategy** values that have been selected from their respective drop-down menus in the model interface, as well as land and labour costs will be calculated. The land-use costs for pasture and meadow land are calculated as followed (Joyce 2019a, 41):

$$a_p = \sum (n_{sp_{cat}} a_{sp_{cat}})$$

$$a_m = \frac{f}{y}$$

In these equations, a_p refers to the area of pasture land in hectare, $n_{sp_{cat}}$ refers to the number of animals per species per age category and $a_{sp_{cat}}$ the area required by these animals per species per age category. f refers to the fodder that all animals require and y is the basic yield of fodder per hectare, which is set at 3000 kg/ha. f is calculated as followed (Joyce 2019a, 41):

$$f = \sum (n_{sp_{cat}} f_{sp_{cat}})$$

Where $n_{sp_{cat}}$ refers to the number of animals per species per age category and $f_{sp_{cat}}$ refers to the fodder required by the animals per species per age category. Fodder can generally be found in meadow areas, but ROMFARMS and thus the implemented model also includes labour costs per household to harvest fodder for those winter months that animals would not be able to go outside. These labour costs are calculated as followed (Joyce 2019a, 42):

$$l = a_m h$$

In this equation l stands for the hours required to harvest the fodder, a_m stands for the meadow area that needs to be worked and its calculation is explained above, h stands for the number of hours required to work one hectare of meadow land and is defined at 16 hours per hectare.

6.2.7 Fuel and timber submodels

The last two submodels that will be discussed are the fuel and timber collection submodels, which are executed via the *go-fuel* and *go-timber* procedures. The inner workings of these two submodels are nearly identical, with the connotation that the fuel collection submodel will be executed each timestep whereas the timber collection submodel will only be performed alongside the values of the **reconstruction-frequency** parameter from the model interface. In ROMFARMS both submodels constitute of three procedures that follow each other up in depth. The *calc-workforce-fuel* procedure calculates the required workforce to collect fuel based on the **fuel-requirement** value of the settlement inhabitants and the composition of its inhabitants. The *find-fuel* procedure is part of the *calc-workforce-fuel* procedure and finds the targets for fuel collection for the workforce. The *take-fuel* procedure is part of the *find-fuel* procedure and ensures that the **biomass** values of the landscape are being transferred to the workforce and removed from the patch variables. The layout and function of the ROMFARMS timber collection submodel is identical, but uses the *calc-workforce-construction*, *find-timber,* and *take-timber* procedures.

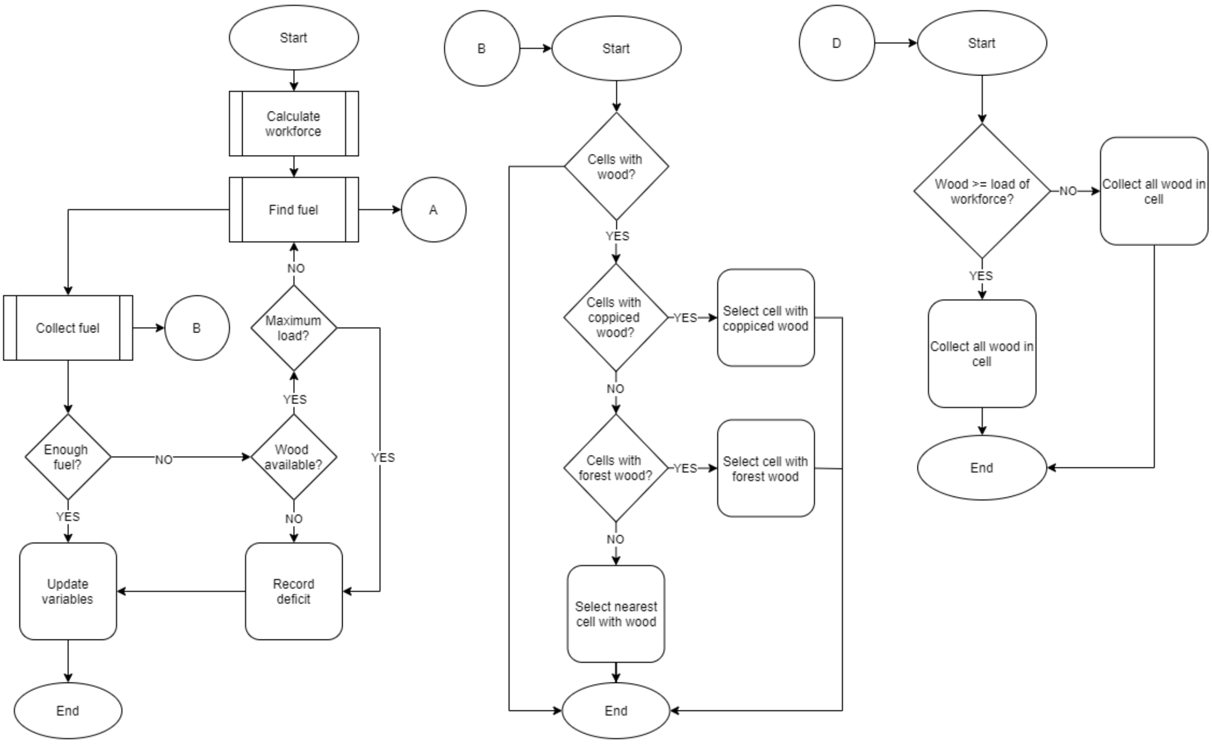


Figure 12: Schematic overview of the workflow from the timber and fuel collection submodels for the docked model (after Joyce 2019a, 46).

Similar to the arable farming and the animal husbandry submodels, the labour force with which fuel can be collected is already determined by the implicit composition of persons in a household. As a result, the *calc-workforce-fuel* and *calc-workforce-construction* procedures have been removed and their code transferred to their respective *go-fuel* and *go-construction* procedures upon further being adjusted to suit a workforce per household. Besides adjusting the algorithms that initially related to the workforce consisting of persons to the household, no further changes have been made to the *find-fuel*, *find-timber*, *take-fuel,* and *take-timber* procedures. When the **coppicing?** variable is set true, some landscape patches include coppiced wood. This wood will first be collected and only if there is none present workforces switch to regular forest wood.

The final and major adjustment has taken place in the *go-construction* procedure. This procedure would initially be called upon once a settlement age has reached the same value as that of the **reconstruction-frequency** parameter from the model interface. After the execution of the timber collection submodel, the procedure would reset the age of the settlement until it reached the same value again. This initially prevented the households in the implemented model from reaching the age of death according to the rules of the Artificial Anasazi model. As a solution a new agent variable called **construction-multiplier** was created. Each household comes equipped with this variable, which increases its value by 1 when the age for construction has been reached. As a result, the timber collection submodel will be executed again once the age of the household has reached the same value as the sum of the reconstruction frequency times the multiplier. This allows the households to continue aging while still engaging in construction activities every required timesteps.

6.2.8 Updating global variables and patches

The *update-globals* and *update-patches* procedures ensure the processing, storing, and updating of relevant information that relates to impact that the simulated agent behaviour has on the simulated landscape as a whole. The *update-globals* procedure includes among others the storage of agent population levels and population growth, and in relation to the different submodels the total labour costs and total resource yields. The *update-patches* procedure regulates the regrowth of arable land that has been cultivated and fuel/timber resources. In relation to the source code of ROMFARMS, small adjustments have been made. The first adjustment is the merge of settlement and person-related commands to apply to the household since the household has taken over the functionality of both the settlement and person-related behaviour. The second adjustment is related to those updated lists and variables that belong to different submodels. The ROMFARMS source code lets the different submodels be run as a whole, whereas in the process of testing the code and functionality of the submodels it was necessary to observe the behaviour of these submodels independently. The implemented model therefore has functionalities that allow the animal husbandry, fuel collection and timber collection models to be disabled when necessary. When updating the global variables, however, this would often leave error messages when running the code because calculations and consequent updates of variables could not take place when submodels were disabled. The docked model still allows to exclude certain submodels from the simulation, so their execution criteria are also applied to the execution of calculating variables that each individual submodel relates to. No further adjustments have been made to the *update-patches* procedure, since this procedure is an integral part of the landscape dynamics simulated in ROMFARMS. Any adjustments to this procedure will significantly impact the possible agent-environment relationship and has therefore not been changed.

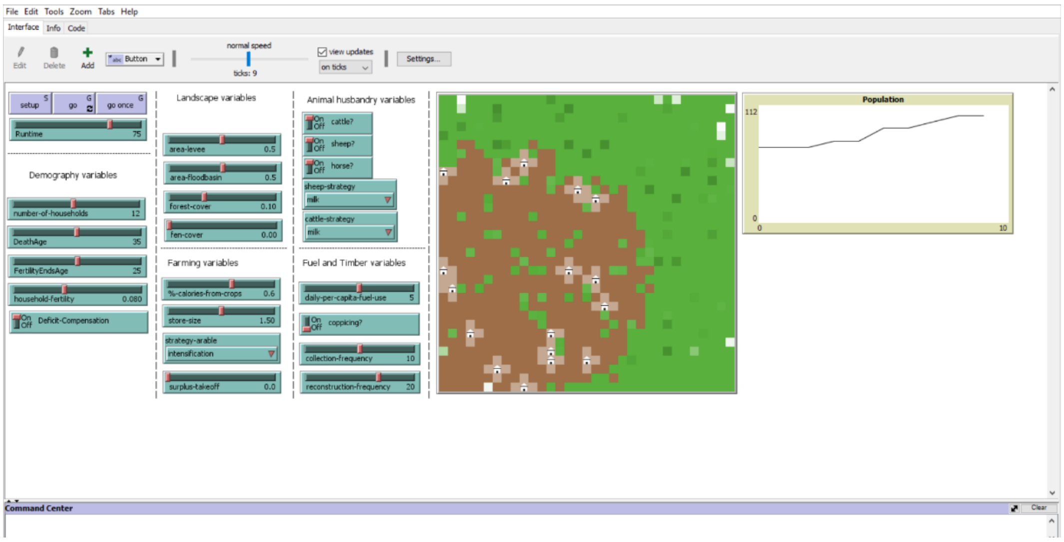
6.2.9 The docked model

The previous sections have discussed the adjustments that have been made to the ROMFARMS model to ensure a functioning model that has integrated ontologically similar entities and their related functionalities from the Artificial Anasazi model. The sections have tried to provide an overview of the alterations that have been made to the algorithms of the original ROMFARMS source code, but have also aimed to provide transparency in the functioning of the model by explicitly explaining the calculations and functionalities underlying some of the algorithms. Not only have the semantically similar ontological entities and their unique functionalities been integrated in the relevant submodels, the other submodels have been adjusted to ensure their functionality with the implementation of these components from the Artificial Anasazi model as well.

The general workflow and structuring of the original ROMFARMS source code was slightly changed as a result of the implementation. Table 6 provides an overview of the new structuring of the procedures present in the implemented model when reading from top row to bottom row. Figure 13 shows the interface of the docked model.

Level1	Level2	Level 3	Level 4	Level 5
Setup	Setup-landscape	Setup-forest		
	Setup-globals			
	Setup-agents	Config-settlements		
Go	Go-demography	Update-settlements-1		
		Deaths		
		Fissioning	New-settlement	Config-settlements
	Go-arable	Calculate-land-arable		
		Calculate-yield-arable		
		Calculate-surplus-arable		
	Go-pastoral	Set-mortality-rates		
		Births-animals		
		Natural-mortality		
		Slaughter		
	Go-fuel	Find-fuel	Take-fuel	
	Go-construction	Find-timber	Take-timber	
	Update-globals			
	Update-patches			

Table 6: Overview of the source code structuring in the implemented model, read from high to low per row (own figure).



normal speed

ticks: 9

☑ view updates

on ticks

Runtime 75

number-of-households 12

DeathAge 35

FertilityEndsAge 25

household-fertility 0.080

On Off Deficit-Compensation

area-levee 0.5

area-floodbasin 0.5

forest-cover 0.10

fen-cover 0.00

%-calories-from-crops 0.6

store-size 1.50

strategy-arable intensification

surplus-takeoff 0.0

On Off cattle?

On Off sheep?

On Off horse?

sheep-strategy milk

cattle-strategy milk

daily-per-capita-fuel-use 5

On Off coppicing?

collection-frequency 10

reconstruction-frequency 20

Figure 13: The interface tab of the docked model (own figure).

7. Verification and equivalence testing

The following chapter focuses on the execution of two types of analysis to aid in the verification of the docked model and to establish the equivalence between the models: a sensitivity analysis and the alignment of model outputs. The sensitivity analysis focuses on the output of the docked model under low-, middle- and high parameter values. The results of this analysis can be compared with known behaviour of the ROMFARMS model, to assess whether relational equivalence exists between the models. The alignment of model results rather focuses on observed statistical relationships between the output results of both models and therefore investigates distributional equivalence. The combination of relational and distributional equivalence will be used to answer the research question of this this thesis. The output data for the sensitivity and statistical analyses has been sorted, plotted, and analysed in the RStudio software. RStudio is an Integrated Development Environment (IDE) built around the R programming language. This programming language has been used because it excels in its capabilities to neatly and efficiently sort, analyse and visualize large amounts of data.

7.1 The sensitivity analysis

Joyce (2019a, 51) mentions that explorative experiments were performed in ROMFARMS to determine the most suitable parameter values, but unfortunately does not provide any elaboration on how these explorative experiments were designed. There is furthermore no elaboration on how often the same sets of parameters have been tested to acquire a general insight in the impact that these parameters would have on the output results of the model. There is thus no direct sensitivity analysis from ROMFARMS with which the results from the docked model can be compared, but in his review of results related to random generated landscapes Joyce makes a few general observations regarding the behaviour of the model that can be used for comparison.

The experiment design for the sensitivity analysis first requires a definition of which parameters to test, which parameters to remain static, what model outputs to measure, and how many times the same set of parameters will be run as a simulation. Since the demography submodel has been significantly altered as a result of the presence of household agents, the impact of the parameters imported from the Artificial Anasazi model (**number-of-households**, **DeathAge**, **FertilityEndsAge** and **Household-fertility**) must be significantly explored. These parameters have a direct impact on population growth or decline and will be observed in relation to the population growth. The farming submodel partially depends on the population development since it determines the number of households that will engage in arable farming, but the **%-calorie-from-crops**, **store-size** and **arable-strategy** parameters also play a role in the total arable land that will be cultivated. The animal husbandry model has not been functionally altered, only in relation to the calculation of workforces. Since no changes have been made in relation to the functionality of the submodel, its functionality will not be investigated in the sensitivity analysis because it is already assumed to be identical. In relation to the fuel and timber collection submodels, the **daily-per-capita-fuel-use**, **collection-frequency**, **coppicing?** and **reconstruction-frequency** parameters will be used to investigate whether there is a threshold for fuel collection that might lead to fuel deficits in households.

The experiments for the sensitivity analysis are divided per submodel. When understanding the behaviour of an earlier executed submodel in the general workflow of the model, the results produced by the next can also be better understood.

7.1.1 Demography submodel

The first experiment focuses on different parameter settings of the demography submodel to get a better understanding of the demographic development of the total number of households under different conditions. Low, middle, and high values were tested while the parameters of other submodels were set to static values, to allow better observation of the dynamics that this submodel produces. The **number-of-households** values were determined at 5, 12 and 25. The **DeathAge** values were determined at 25, 30 and 40. The **FertilityEndsAge** values were determined at 20, 25 and 30 and the **Household-fertility** values at 0.05, 0.1 and 0.15. The values for the age of death might be confusing, but keep in mind that this age reflects the death of a household and not an individual person. This experiment allowed for 81 unique parameter combinations and the fact that each was combination was running for ten times made the total number of simulation runs for this experiment 810. Observations were made in relation to the number of households at each of the 50 timesteps. The applied parameter values can be found in table 7 .

Submodel	Parameter	Tested values
Setup	Runtime	50
	area-levee	0.5
	area-floodbasin	0.5
	forest-cover	0.1
	fen-cover	0
Demography	number-of-households	12
	DeathAge	30-35-40
	FertilityEndsAge	20-25-30
	Household-fertility	0.05-0.1-0.15
	Deficit-compensation	True
Arable farming	%-calories-from-crops	0.6
	Store-size	1.5
	strategy-arable	"None"
	Surplus-Takeoff	0
Animal husbandry	Cattle?	False
	Sheep?	False
	Horse?	False
Fuel collection	daily-per-capita-fuel-use	1
	Coppicing?	False
	Collection-frequency	10
	Reconstruction-frequency	20

Table 7: Parameter combinations used for the sensitivity analysis experiment of the demography submodel.

After all of the simulations were run the data was collected, processed and imported in RStudio, where the data was aggregated by calculating the mean population values of runs with the similar parameter and timestep values. This resulted in the mean demographic results of each unique parameter combination at each related timestep, which are visualized in figure 14.

From these figures the general behaviour of the model can be observed. Even though the final ranges of the output variables vary due to the initial households that were present, the general development of the number of households in relation to the other parameter settings is relatively similar. In all of the simulations, the highest count of households in the last timestep (240 at 25 households, 110 at 12 households and 35 at five households) is achieved by the same set of parameters: the maximum values of the **DeathAge** (=40),

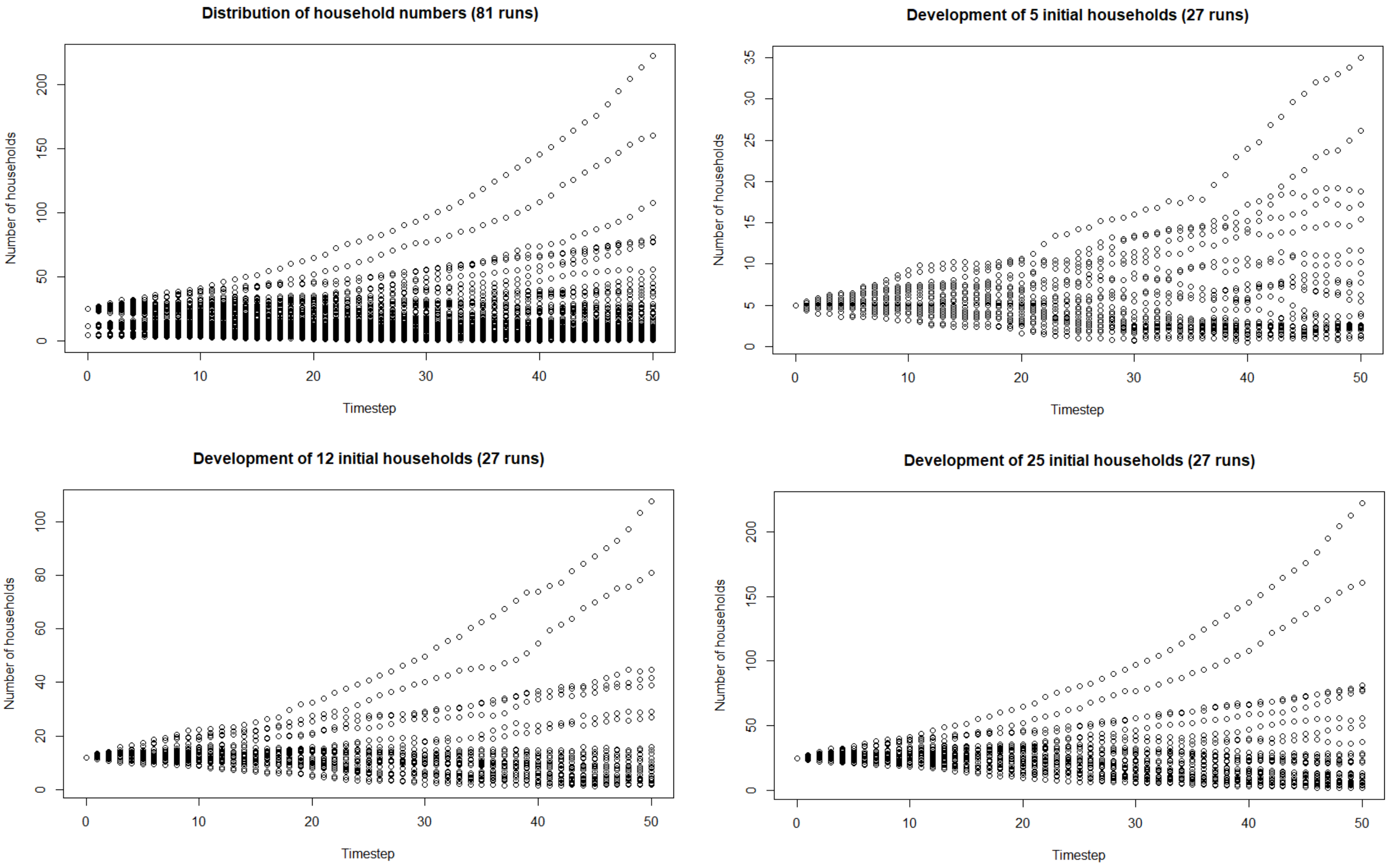


Figure 14: The output of the demography submodel experiment.

FertilityEndsAge (=30) and **household-fertility** (=0.15) parameters. The lowest count of households in the last timestep of a simulation is achieved by the lowest settings of the **FertilityEndsAge** (= 20) and **household-fertility** (= 0.05) parameters. The age of death seems to be of lesser importance in this issue. Middle ranges that deviate slightly (scaled relative to the initial number of households) from the initial number of households at the final timestep are generally determined by middle to low values of **FertilityEndsAge**, middle to high values of **household-fertility** and middle to high ranges of **DeathAge**. The population dynamics of a household are thus primarily determined by the overall time that a household is fertile and the general chance that a new marriage occurs. When the values are set at their lowest, this will generally trigger population decline.

7.1.2 Arable farming submodel

A separate experiment was created for the arable farming submodel. The results from the analysis of the demography submodels were used to determine the set of stable demographic parameters. The chosen parameter setup is visualized in table 8 . Despite still allowing for a degree of variation in population numbers due to the influence of stochastic elements, this parameter setup generally leads to a slightly growing population. Two different arable farming parameters were tested at their low, middle, and high values: **%-calories-from-crops**, **store-size***,* and the three options for arable strategy were tested as well. Each parameter combination was run for a total of ten times, after which all the data was aggregated and sorted in RStudio. The results per individual parameter

Submodel	Parameter	Tested values
Setup	Runtime	50
	area-levee	0.5
	area-floodbasin	0.5
	forest-cover	0.1
	fen-cover	0
Demography	number-of-households	12
	DeathAge	35
	FertilityEndsAge	25
	Household-fertility	0.1
	Deficit-compensation	True
Arable farming	%-calories-from-crops	0.3-0.6-0.9
	Store-size	1.25-1.5-2.0
	strategy-arable	"None" "Intensification" "Extensification"
	Surplus-Takeoff	0
Animal husbandry	Cattle?	False
	Sheep?	False
	Horse?	False
Fuel collection	daily-per-capita-fuel-use	1
	Coppicing?	False
	Collection-frequency	10
	Reconstruction-frequency	20

Table 8: Parameter combinations used for the sensitivity analysis experiment of the arable farming submodel.

will first be discussed in relation to their significant output results. In the case of the arable farming model, the output results are the mean area of arable land, the mean grain surpluses that households produce and the mean grain deficits that households produce.

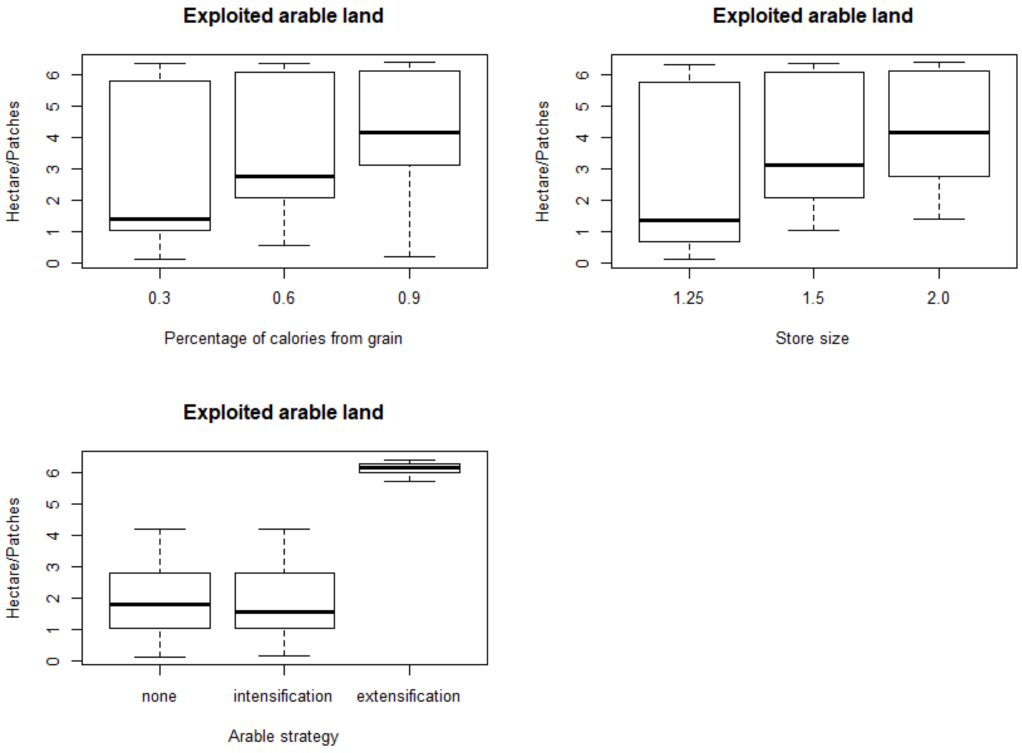


Figure 15: The output of the arable farming submodel in relation to the arable land.

The output results that relate to the mean area of cultivated arable land are visualized in figure 15, which groups the same distributions of output results per parameter set via box-and-whiskers plots. From the combination of the three graphs, the general functioning of the arable farming submodel in relation to the cultivated arable land becomes visible. From the upper two graphs it is visible that an increase in the percentage of calories that is derived from grains or the store size that determines the buffer of sowing seed causes a visible increase in the value of the median value in the total distribution. This implies that the total number of arable land, described in hectares that are equal to one grid cell in the simulation environment, increases when one of these parameters increases as well. The arable strategy graph visualizes how the chosen arable strategy influences the total area that will generally be cultivated. It becomes clear that the extensification strategy, which focuses on expanding the cultivation to its maximum capacity, causes a significant higher exploited area of arable land than the other two. This strategy also explains the high ranges in the third and fourth quartile of the other boxplots. The results in relation to this strategy are therefore as expected and also the main cause for the higher range values in the fourth quartile in the graphs of the other two parameters, since high observed values all related to this arable strategy. The plots of the intensification and lack of arable strategy also produce results that are as expected. The “none” arable strategy focuses on the exploitation of the minimum area to fulfil the nutritional requirements of a household and the intensification strategy focuses on the same purpose but generally has higher yields due to the fact that land is additionally being manured. The fact that the general distribution of arable land between these two arable strategy values is almost identical is therefore not a big surprise. In general it can be concluded that the sensitivity analysis in relation to the areas of arable land does not yield any unexpected results.

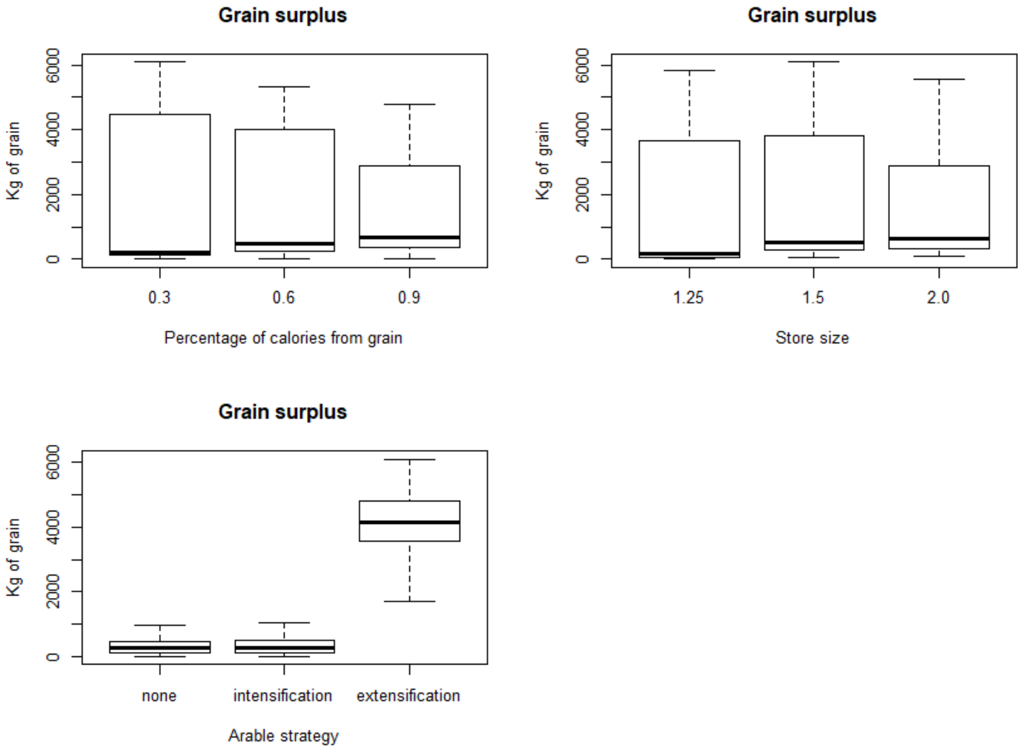


Figure 16: The output of the arable farming submodel in relation to the production of grain surpluses..

The second output variable is that of the grain surplus. The output results of the grain surpluses are visualized in figure 16, the way of visualization is identical to the visualization of the previous arable land output results. The highest grain surpluses seem to be caused by the arable strategy that emphasizes on exploiting the highest level of cultivatable arable land available: extensification. This strategy is again the cause of the relative higher ranges in the third and fourth quartile of the boxplots related to grain consumption in the general diet and the store size of the households. The surplus production between the lack of arable strategy and arable intensification is, comparable to the arable land output results, near identical. From the store size and percentage of the diet that is constituted by grain it can again be concluded that general surplus yields increase when their parameter values increase as well. The underlying reason for this is that an increase in either the store size or the percentage of calories from grain causes proportional increase in the area of cultivated land and surpluses. Based on these parameter settings no significant or surprising results are therefore produced in relation to the produced arable surpluses.

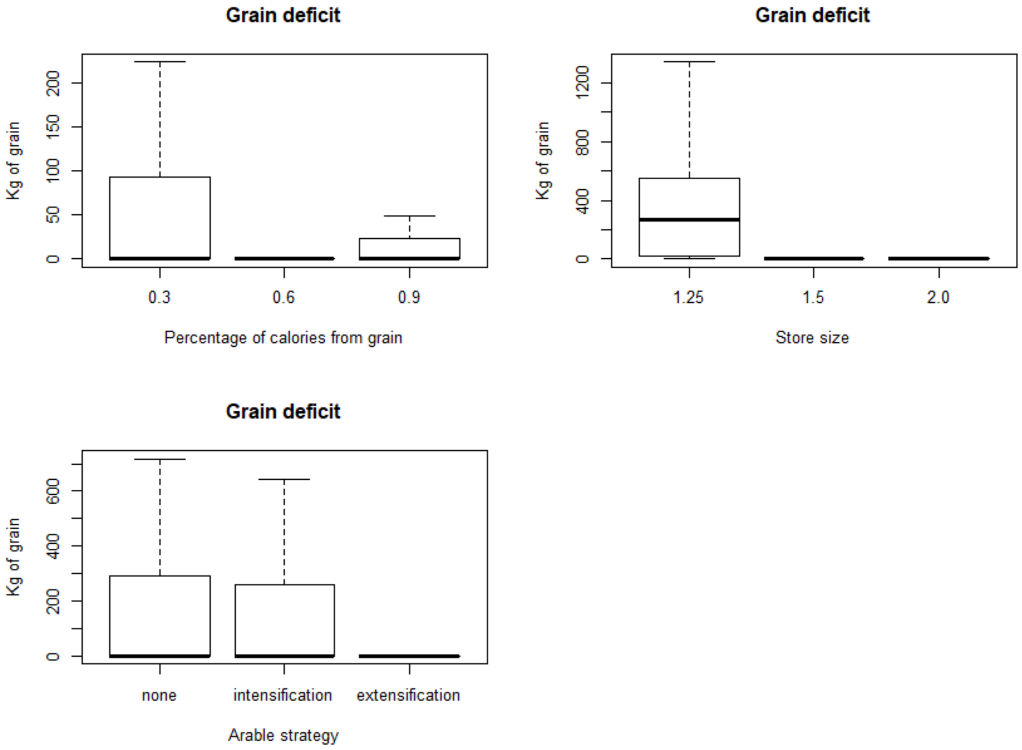


Figure 17: The output of the arable farming submodel in relation to the occurrence of grain deficits.

The final category of output results that is analysed in relation to the arable farming submodel is that of the production of grain deficits (see figure 17) under the different parameter values. The values are visualized in the same manner as the previous categories of results related to the arable farming submodel. When looking at the different boxplots it immediately becomes visible that the only grain deficits are produced when the store size is set to 1.25, for the other two values no grain deficits are produced. It is therefore that the visible outliers that fall outside the maximum upper interquartile range are all produced when grain store sizes are set at 1.25. A quick manual investigation for the causes of these deficits was undertaken after these results were first produced by visually monitoring different agents through all the timesteps of a single simulation run. The factor that caused the households to generate grain deficits under these conditions was that the store size did not provide a sufficient buffer of grain for sowing and cultivating the land during the next timestep. Even though labour and land for cultivation would still be available, the grain to actually sow the and harvest the land was missing. This produced a negative loop of deficits which could not be deviated from inside the general framework of simulation rules. The lowest store size to have absent grain deficits therefore lies between the 1.25 and 1.5 values. This corresponds with Joyce's (2019a, 61) findings that deficits were consistently produced withs store size values below 1.4. Furthermore it can be concluded based on the visualization of the data that the higher the percentage of calories in the diet constitutes of grain, the higher the eventual grain deficits will be. The highest grain deficits are also present when none or an intensification strategy are being pursued. Extensification still allows for the presence of grain deficits in the output results, but the fact that this strategy emphasizes on maximum exploitation of the available area seems to generally compensate for grain deficits.

As a whole, the arable farming submodel behaves as expected and no surprising results were produced by the model output under the different parameter combinations.

7.1.3 Fuel submodel

For the fuel submodel low, middle, and high values were tested for the parameters **daily-per-capita-fuel-use**, **collection-frequency** and **reconstruction-time**. The **coppicing?** parameter was both enabled and disabled. Similar to previous experiments, each unique set of parameter combinations was tested five times, allowing for a total of 270 simulation runs. The total set of parameters can be found in table 9. The set of output results for the fuel collection submodel were the mean amount of collected fuel, the mean travelled distance during fuel collection and the mean required fuel labour. The visualization of the output results is similar to that of the output results from the previous two submodels. The sensitivity analysis for the timber data is derived from this same experiment, since this submodel functions near identical as the fuel collection submodel despite being less often executed.

Submodel	Parameter	Tested values
Setup	Runtime	50
	area-levee	0.5
	area-floodbasin	0.5
	forest-cover	0.1
	fen-cover	0
Demography	number-of-households	12
	DeathAge	35
	FertilityEndsAge	25
	Household-fertility	0.1
	Deficit-compensation	True
Arable farming	%-calories-from-crops	0.5
	Store-size	1.5
	strategy-arable	"None"
	Surplus-Takeoff	0
Animal husbandry	Cattle?	False
	Sheep?	False
	Horse?	False
Fuel collection	daily-per-capita-fuel-use	1-5-10
	Coppicing?	True - False
	Collection-frequency	1-10-20
	Reconstruction-frequency	10-20-30

Table 9: Parameters used for the experiment related to the fuel and timber collection submodels.

The first of the output results related to the fuel collection submodel is the mean fuel store of the households. The results are visualized in figure 18 . From these results it can be observed that the fuel store levels increase when the fuel requirements per households member increase. When fuel has got to be collected every day, which occurs when collection frequency is set at its lowest, the mean storage levels of households is much higher compared to households with a higher collection frequency. The significant scaling of middle and lower values for this parameter clearly indicate that mean storage levels are much lower, which is arguably caused by the fact that mean fuel levels are consistently decreasing. This in return would significantly lower the mean storage values. There is furthermore no difference observable between mean storage levels based on the activation of the coppicing function, which allows for forest patches to yield higher amounts of biomass. Based on these values, the functioning of the model in relation to the storage levels is therefore as expected and correlates to Joyce's (2019a, 78) description of how these parameters influence the output.

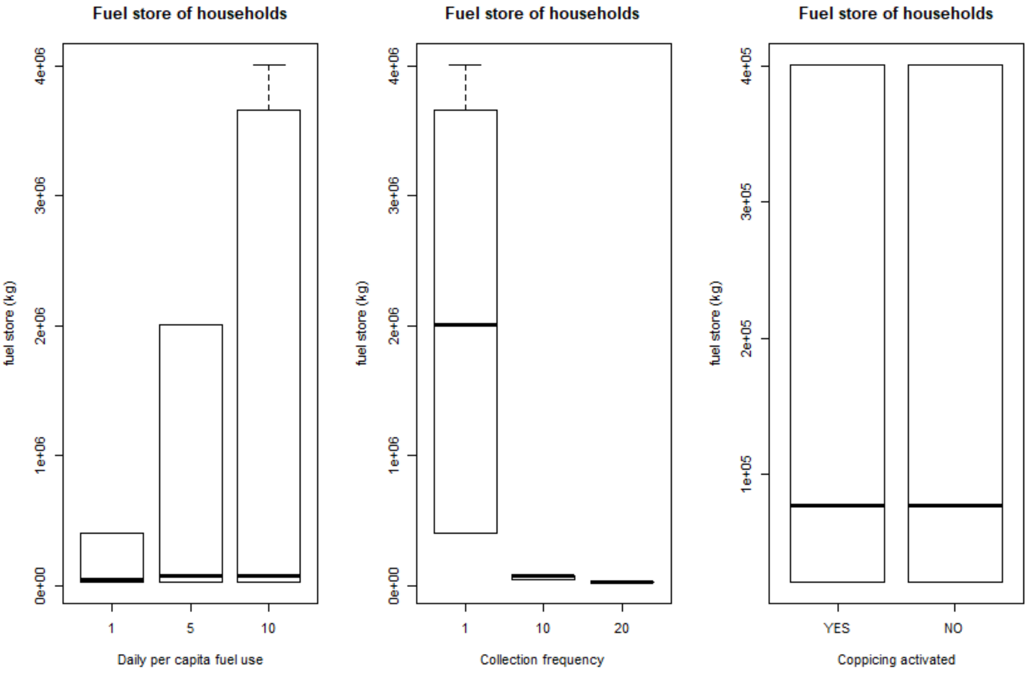


Figure 18: The output of the fuel collection submodel in relation to the fuel storage levels.

Figure 19 describes the mean fuel processing time of the households. It becomes clear that general fuel processing time remains similar when the fuel requirements per household member rise. The processing time is constant and decreases when the collection frequency increases. The pattern is observed for the same reason as with the fuel storage levels, namely that the mean values are affected by the moments that no fuel is processed.

Again, no differences between the activation of the coppicing feature are observed in relation to these results. The results in relation to the fuel processing time thus yielded surprising results, which are primarily being caused by a cruder implementation of the rules for workforce aid in relation to the fuel collecting and processing process.

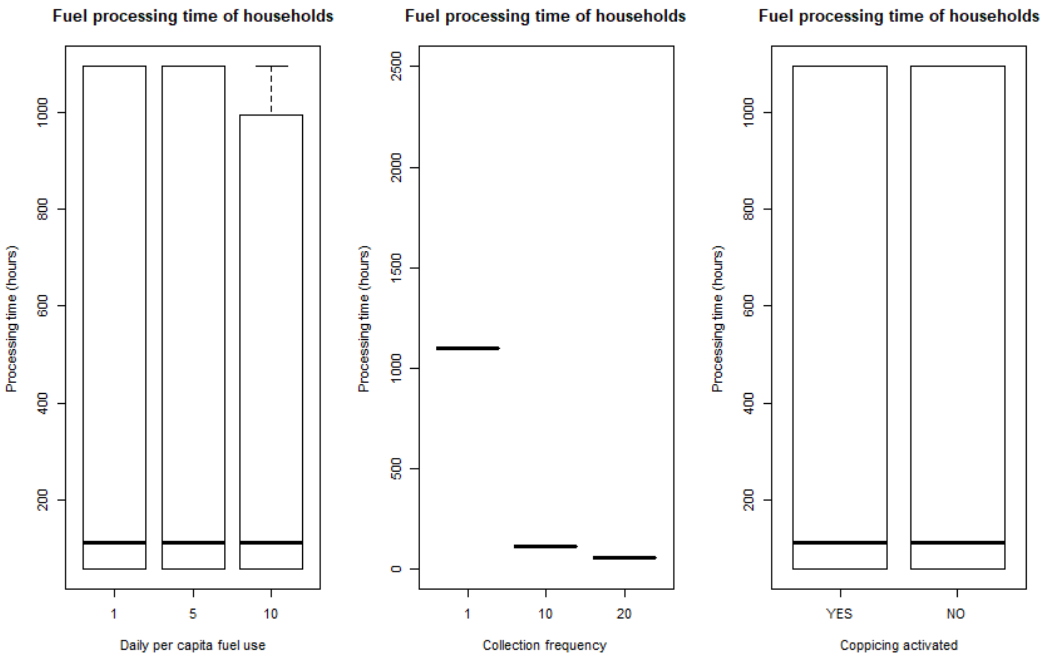


Figure 19:The output of the fuel collection submodel in relation to the fuel processing time.

Figure 20 describes the distribution of the mean distance travelled under the different parameter combinations. One remarkable observation is that the distribution ranges of the third and fourth quartile of the travelled distance are smaller and overall values lower when the fuel use per household member is set at 10 compared to when it is set at 5. This is surprising, and it is assumed that when the fuel use per head is set at 5, this value resembles a point where slightly more than the regular biomass that one hectare/patch produces must be exploited. As a result, more travel occurs since this means that the fuel workforce exploits both unexploited patches as well as those with small amounts of biomass left. When set at 10, the workforce is forced to generally exploit a roughly similar number of patches but collects a much higher amount of biomass from these patches. Average distance travelled, similar to the fuel storage levels, declines when the collection frequency is set higher. Again this is caused by fewer travel days, and it is these empty days that significantly lower the average travelled distances.

.

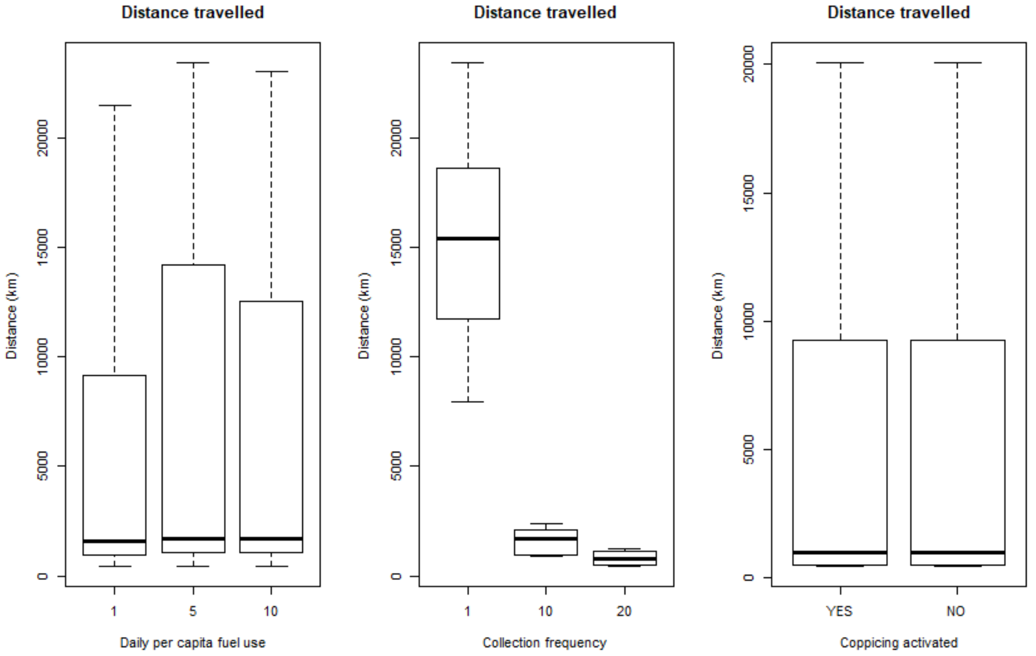


Figure 20:The output of the fuel collection submodel in relation to the distance travelled for fuel collection.

The final output is the average hours of labour, which are visualized in figure 21. One initially surprising result is the drop of required hours when the daily fuel requirements are set at 10 per household member. This is a result of the workforce calculation. Initially the workforce for fuel travel consists of the male and female adults in the household. When fuel requirements become higher, the docked model assumes that all subadults engage in fuel collection. The engagement of all subadults significantly lowers the required fuel labour. Furthermore, no real surprises are visible in this visualization. Where the first column of the collection frequency table gives an oversight of the total average required hours, the fact that labour does not occur daily has a significant impact on the average values at higher collection frequency levels.

The results from this experiment generally overlap with the expectations of the behaviour of this submodel. No real surprises were found.

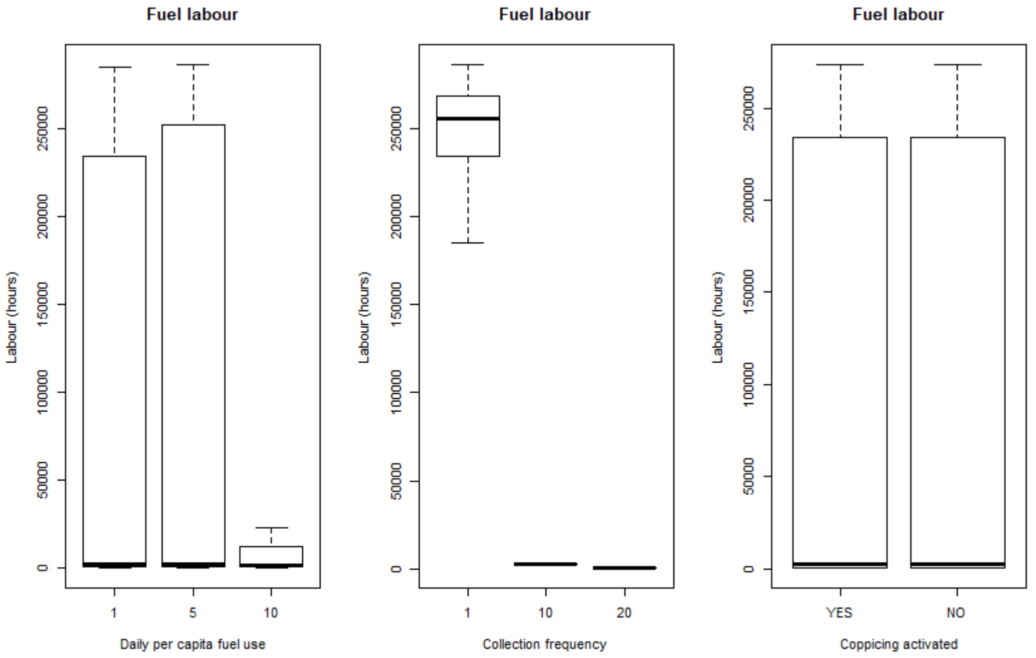


Figure 21:The output of the fuel collection submodel in relation to the required fuel labour.

7.1.4 Timber collection submodel

The experiment for the timber collection submodel was incorporated in the experiment for the fuel collection. The only parameter directly associated with the timber collection submodel is the **reconstruction-frequency** variable. The timber collection submodel uses the same output results compared to the fuel collection submodel: The average timber storage levels, the timber processing time, the travelled collection distance, and the construction labour.

The output result data in relation to the reconstruction frequency is visualized in figure 22 . The first thing to note is that the timber storage levels consistently remain at the level of the construction requirements. A surprising observation is that the distribution ranges for the other output values become wider as the reconstruction frequency increases. After manually monitoring the behaviour of different households, the reasons for this became clear. The wider distribution ranges are primarily caused by two factors. The first is that a higher reconstruction frequency means that there are less households that influence the distribution, since not all households will reach the reconstruction frequency age before the final fiftieth timestep. Lower sample sizes could therefore automatically lead to more spread distributions. The second factor is the fact that a higher reconstruction frequency generally means that households have to travel more to collect their timber. When households have reached a higher reconstruction frequency age, there is generally a chance that there are more households present in the landscape under the applied parameters of the demography submodel in this experiment. The fuel collection of all these households depends on the available biomass as well, and the fact that households engage in timber collection after all households have collected their required fuel might lead to more required travelling distance for timber collection since the closest biomass sources have already been exploited. This leads to higher processing time, travelling distance and labour times.

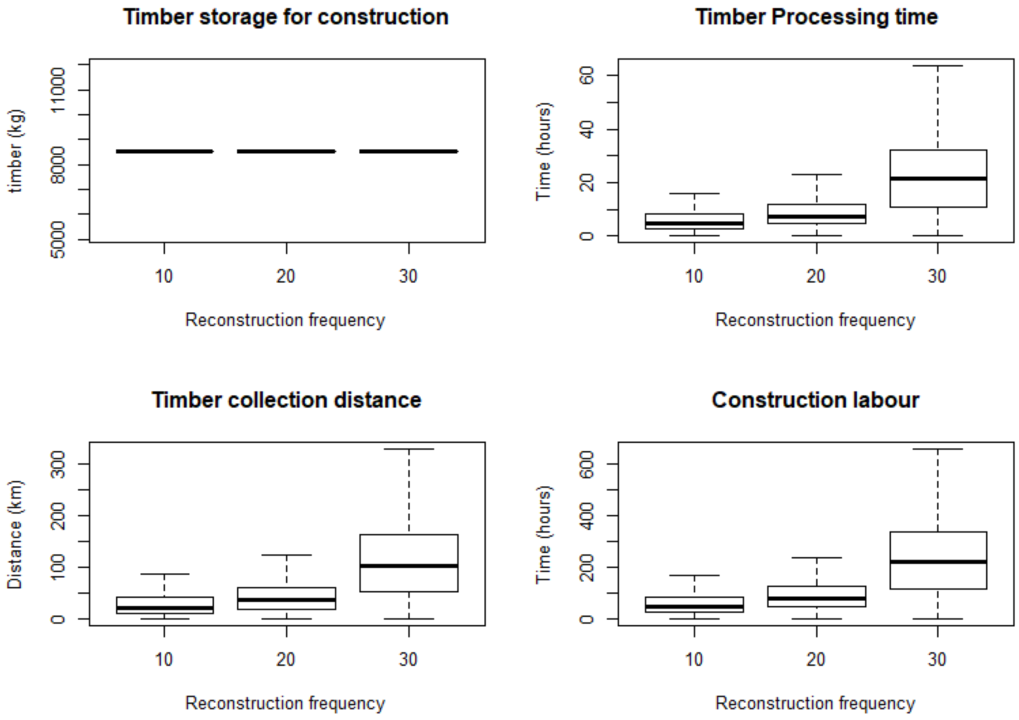


Figure 22: The output of the timber submodel in relation to the same categories that have been discussed for the fuel collection submodel.

Despite the functionality of the timber collection submodel being relatively similar when compared to the fuel collection submodel, the timber collection submodel is much more static in nature. It is generally not called upon every timestep since it depends on the age of the individual household. The model behaves as expected, but previous fuel collection activities by all households need to be considered before interpreting timber collection results.

7.2 Alignment of model outputs

The second analysis focuses on the alignment of the outputs of the ROMFARMS model and the docked model to determine whether distributional equivalence is present between the results. Analyses all submodels except the timber collection will be discussed. The analysis will be based on an experiment where the parameter values for the arable farming, fuel collection and timber collection submodels will be identical between both models where possible. Both ROMFARMS and the docked model allow for the exploration of many parameter combinations. Only the middle ranges of these parameters have been assessed for alignment purposes. The only exceptions to this rule are the household distributions of the ROMFARMS settlements, the docked model household fertility for the demography submodel, and the arable strategy for the arable farming submodel. An overview of the performed experiment can be found in table 10 . The "ROMFARMS-2" experiment only applies to an additional experiment performed in relation to the demography submodel and animal husbandry submodel since it allows

Submodel	Parameter	Docked Model	ROMFARMS-1	ROMFARMS-2
Setup	Runtime		75	75
	area-levee	0.5	0.5	0.5
	area-floodbasin	0.5	0.5	0.5
	forest-cover	0.1	0.1	0.1
	fen-cover	0	0	0
Demography	number-of-households	12	NA	NA
	DeathAge	35	NA	NA
	FertilityEndsAge	25	NA	NA
	Household-fertility	0.05 to 0.1	NA	NA
	Deficit-compensation	True	NA	NA
	no-1-household-settlements	NA	2	12
	no-2-household-settlements	NA	1	0
	no-3-household-settlements	NA	1	0
	no-5-household-settlements	NA	1	0
Arable farming	%-calories-from-crops	0.6	0.6	0.6
	Store-size	1.5	1.5	1.5
	strategy-arable	"Intensification" "Extensification"	"Intensification" "Extensification"	"Intensification" "Extensification"
	Surplus-Takeoff	0	0	0
Animal husbandry	Cattle?	False	False	NA
	Sheep?	False	False	NA
	Horse?	False	False	NA
Fuel collection	daily-per-capita-fuel-use	1	1	NA
	Coppicing?	False	False	NA
	Collection-frequency	10	10	NA
	Reconstruction-frequency	20	20	NA

Table 10: Overview of the parameters used for the comparison of the model outputs.

output comparison on the demography submodel level with different ROMFARMS household to settlement relationships.

The timber model is not always activated during each timestep, because the timesteps the timber model is activated is based on the age of the household and thus constantly varies. The course of the timber model during the situation can therefore not be plotted as an average simulation run with 75 timesteps. Results of the timber collection model are therefore not discussed in singularity. Joyce (2019a, 78-80) faced the same problem and also does not discuss the results of the timber collection submodel individually.

To assess whether the distributions are statistically different or similar from each other, suitable statistical were performed based on the characteristics of the distributions and their employed variables. The output distributions are not assumed to be normally distributed, are unpaired and are measured along a continuous variable (the timestep). Two tests allow for a statistical comparison between the distributions based on unpaired data with continuous variables and no normal distribution, which are the two-sided Kolmogorov-Smirnov (K-S) test and the Mann-Whitney U test.

The K-S test is a goodness of fit test that investigates the probability that two different distributions could have originated from a similar distribution based on the shape of their plotted graphs. It calculates the goodness of fit of two distributions based on the so-called D-value, which is the biggest observed difference between two values for a timestep. The K-S test then measures how often the differences between the observed values correspond with - or deviate from - this D-value. If the difference between the values would adhere to the D-value every timestep, a complete parallel and thus statistically similar distribution would be visible. The K-S test produces a probability-value (the p-value) based on which a hypothesis can be accepted or rejected.

The Mann-Whitney U test assesses whether there are significant differences between the values of two unpaired distributions. This test ranks all of values of two distributions from low to high and assigns a score to each value that represents the count of lower values from the other distribution. The sums of the scores per distribution can be compared to generate a p-value based on which a hypothesis can be accepted or rejected.

All statistical analyses have been performed in RStudio, with an assumed probability value of 5 percent (0.05). When the p-value of a statistical tests exceeds this predefined value the hypothesis that both distributions are similar is accepted, if the p-value does not exceed it the hypothesis that both distributions are similar rejected.

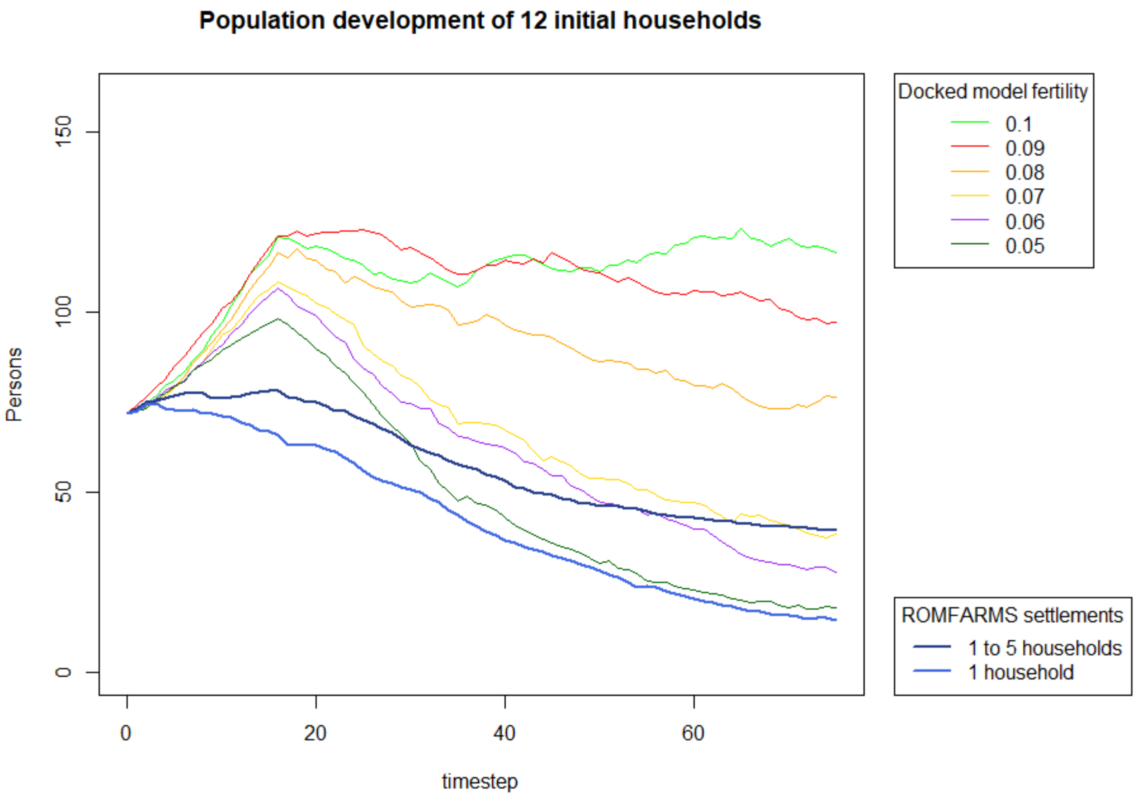
7.2.1 Demography output results

Since ROMFARMS allows for the presence of settlements with one, two, three or five households, the number of twelve initial households in ROMFARMS can be achieved in different ways. Therefore, it was decided to use two different ROMFARMS experiments where the relationship between 12 households and their settlements was different. In the first experiment, which was also used for the statistical analyses of the other submodels, the twelve initial households were spread over one five-household, one three-household, one two-household and two one-household settlements. In the second experiment, the choice was made for twelve one-household settlements since this would be the setup with the closest semantic similarity. The setup of these experiments, labelled ROMFARMS-1 and ROMFARMS-2 is found in table 10 alongside the setup of the docked model. Each individual experiment was run for 30 times to ensure results that adequately describe the patterns that each model produces.

The output measurement for the comparison of the demography submodels is the population count per timestep and is calculated in persons rather than households. The reason for this is that despite the fact that initial households in ROMFARMS start with two adults and four subadults, household compositions may vary over time. The count of persons is therefore a better comparable output result than the number of households. For the docked model, the number of persons is calculated by multiplying the number of households by the static number of six members per household. The goal is to assess the goodness of fit of the output results from the two models.

Both models, however, differ in relation to the parameters that relate to the demography submodel. The choice was made to adhere to an initial household number of twelve, but a variety of parameters was explored to assess which combinations provide the best statistical and visual fit. The difference between the two different household-settlement relationships from the ROMFARMS experiments has already been explained. For the docked model middle values in relation to the **DeathAge** (35) and **FertilityEndsAge** (25) were tested, with varying **household-fertility** values varying between 0.05 and 0.1 with 0.01 intervals (see table 10).

The results from the experiments in relation to both models are visualized in figure 23 . From the plot of the population development it is visible that the three highest employed fertility values provide a growing population whose numbers are higher at the final timestep compared to the initial population. Except for the highest fertility value, all of the populations from the docked model experience a phase of population growth which is then followed by a phase of gradual population decline. The ROMFARMS data shows two different population development curves based on the initial household distribution over the available settlements. In the experiment where multiple households were present in settlements, the total population shows a small period of growth, followed by a period of decline. In the case of only one-household settlements, no growth is observed.



Compared against varied-household distribution		
	K-S Test	Mann-Whitney U
Fertility 0.10	< 2.2E-16	< 2.2E-16
Fertility 0.09	< 2.2E-16	< 2.2E-16
Fertility 0.08	< 2.2E-16	< 2.2E-16
Fertility 0.07	0.00014	0.003
Fertility 0.06	0.002	0.1664
Fertility 0.05	1.20E-06	0.103

Compared agains one-household distribution		
	K-S Test	Mann-Whitney U
Fertility 0.10	< 2.2E-16	< 2.2E-16
Fertility 0.09	< 2.2E-16	< 2.2E-16
Fertility 0.08	< 2.2E-16	< 2.2E-16
Fertility 0.07	7.86E-08	3.97E-10
Fertility 0.06	6.62E-05	2.11E-10
Fertility 0.05	0.002	0.02

Figure 23: The output results of the models in relation to the demography submodel.

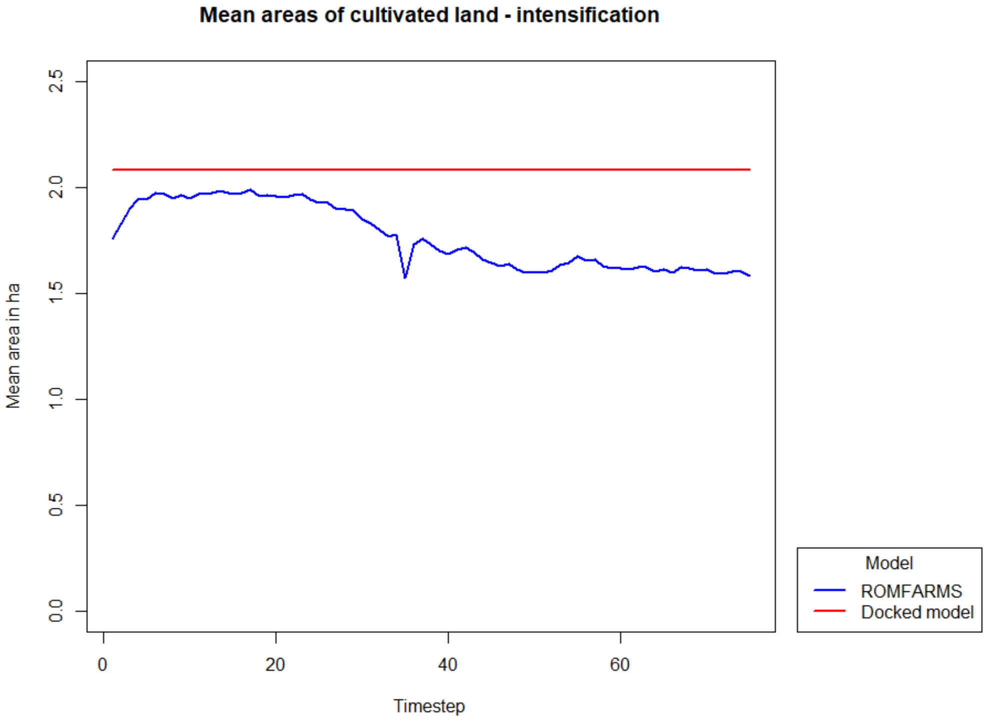
The results of the K-S tests and Mann-Whitney U tests for a statistical comparison in relation to the two different ROMFARMS distributions can also be found in figure 23 . In all of the performed K-S tests, the hypotheses that the distributions are equivalent based on their shape were rejected. The presented p-values are lower than 2.2E-16. This value is a simplified notation of the number 0.00000000000000022, the lowest computable p-value in the R programming language. Since these p-values are lower than 5 percent, the shape of the distributions therefore does not provide any statistical indications that the distributions could have a similar origin.

The K-S tests rejected the hypotheses because ROMFARMS does not produce a comparable peak of population growth and gradual decline, a comparable difference is not observed at any other measurement. From the docked model it is known that the initial population growth is caused by the initial setup of the settlements, that relatively quickly reach the age of fertility, but it was decided to manually investigate why ROMFARMS does not produce at least a little increase in population. A rule was found in ROMFARMS that stated that new settlements can only enter the simulation when their presence does not exceed the settlement density. In random generated landscapes ROMFARMS however naturally assumes a settlement density of 0. As a result, any time a marriage occurs, the male and female starting a new household are removed from the simulation since this household would exceed the settlement density. New households can therefore only move to multiple-household settlements that actually have extra space for a new household. This also explains why the ROMFARMS population levels with multiple-household settlements declines at a lower rate than the distribution with only single-household settlements.

The Mann-Whitney U test in two cases related to the varied-household settlement distributions accepted the hypothesis that the distributions could be considered similar. The hypotheses were accepted with fertility values set at 0.06 and 0.05, producing probability values that exceeded the 5 percent threshold. Both models make use of a high number of possible parameter settings, having a high number of possible outcome distributions that are unique to those parameter settings. However, these statistical tests do indicate that there are thus parameter combinations that can indeed produce statistically similar distribution results.

7.2.2 Arable farming output results

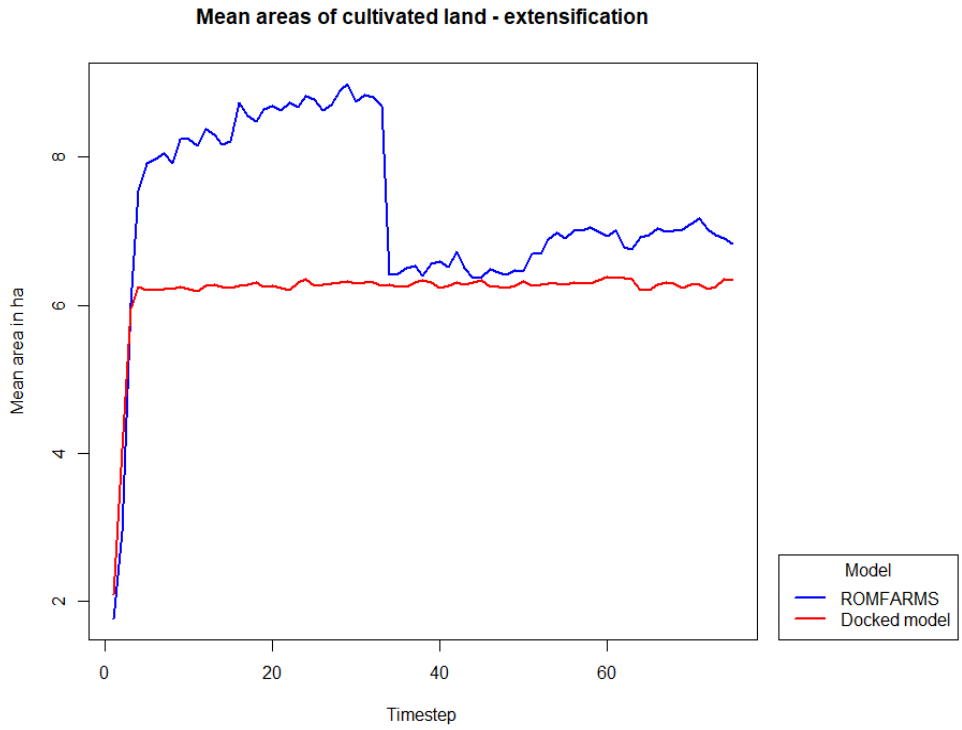
The alignment of the arable farming submodel focused on two output results that were also discussed in the sensitivity analysis: The average area of arable land exploited per household and the size of the generated arable surpluses. Deficits are not taken as output measurements because the middle ranges of the store size, which equals to 1.5, guarantees surplus production for both models. The percentage of calories that constitute of grain are set at 60 percent for both models and both the intensification and extensification arable strategies will be statistically evaluated. The demography submodel parameter settings that produced the strongest statistical similarities between both models were utilized to ensure that the correlation between the populations of both models is as reliable as possible. The experiments per unique parameter combination were performed for 30 times and the results for each timestep have been averaged, similar to the demography submodel. Kolmogorov-Smirnov and Mann-Whitney U tests have also been performed on the produced distributions, where the distribution generated by the ROMFARMS model has been taken as the baseline.



Intensification strategy - arable land		
	K-S Test	**Mann-Whitney U**
Intensification	< 2.2E-16	< 2.2E-16

Figure 24: The output results of the models in relation to the arable land with the intensification strategy

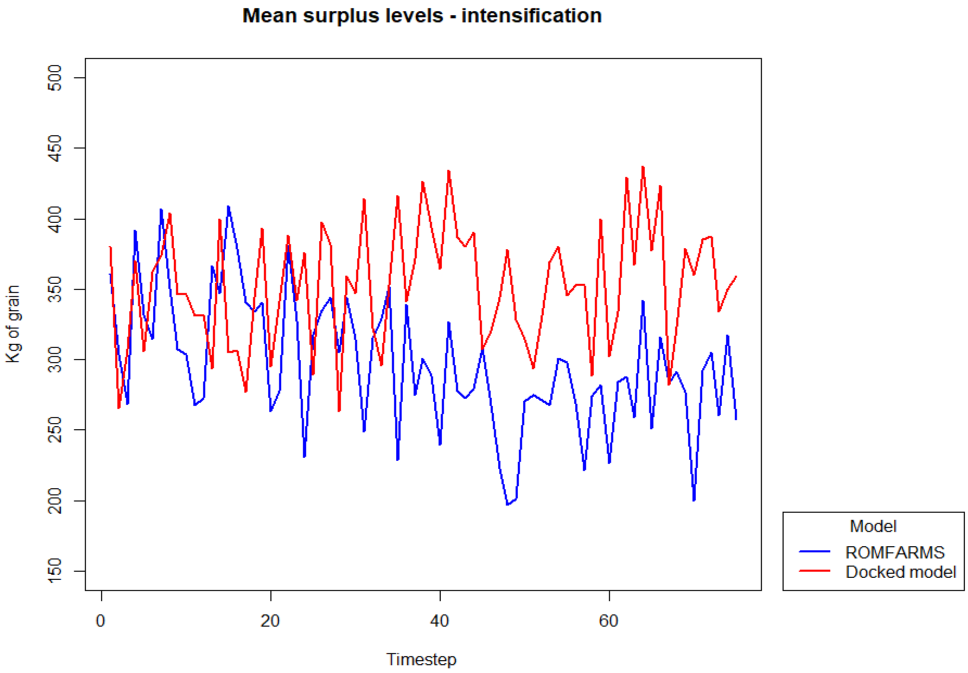
The first output results that will be discussed are the output results related to the cultivated areas of land. These output results can be divided in two categories that each relate to a different agricultural strategy. Figure 24 shows the results of the mean exploited areas of arable land per timestep and the results of the statistical tests related to the intensification strategy. From this figure it is observable that, contrary to ROMFARMS, the total cultivated area per timestep is constant at a value of 2.08 hectare throughout the simulation of the docked model. Due to these constant values the statistical goodness-of-fit tests both rejected the hypotheses that the distributions could be considered equivalent at a p-value of 5 percent. The constant value of cultivated land related to the docked model can most likely be explained by the static calorie requirements of the individual household. The mean cultivated area values of the ROMFARMS model tend to portray the same steady decline and eventual stabilization pattern observed in the results alignment of the demography submodel, although with a slight discontinuity of the pattern around the thirtieth timestep. What causes this slight discontinuation is unclear.



Extensification strategy - arable land		
	K-S Test	Mann-Whitney U
p-value	< 2.2E-16	< 2.2E-16

Figure 25: The output results in relation to the arable land with the extensification strategy.

The results of the average areas of cultivated lands related to the extensification strategy are visualized in figure 25. The distribution related to the docked model shows slight fluctuations between the 6.2 and 6.4 hectare and is almost as constant as the pattern observed at the intensification arable strategy. The ROMFARMS output results, however, vary significantly compared to the results of the intensification strategy. Not only is the exploited arable area consistently higher where it was consistently lower for the intensification strategy, a significant and unexpected drop in cultivated area of around 2 hectare is also observed at the 30th timestep. Whereas a similar drop was visible on a smaller scale for the intensification strategy, it now completely disrupts the pattern. The source code from the arable farming model was consulted, but the code did not contain any bugs that could have led to this discontinuation. The extensification strategy focuses on the maximum exploitable area of land, which is determined based on the lowest value of three possible factors - the available land, available grain, and available workforce. Despite not precisely knowing what has caused this shift, it is clear that at the thirtieth timestep an event occurs that significantly lowers the values of either one, two or all three of these factors. In relation to the statistical tests, the hypothesis that the distributions

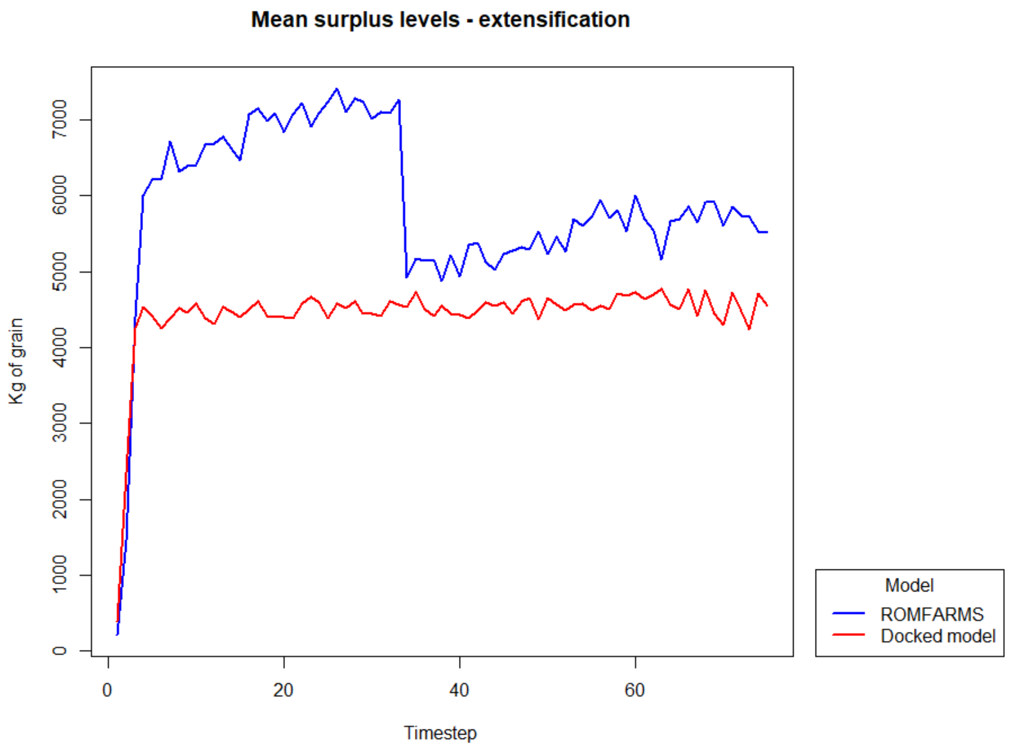


Intensification strategy - grain surplus		
	K-S Test	Mann-Whitney U
p-value	3.42E-08	6.63E-11

Figure 26:The output results of the models in relation to the production of grain surpluses with the intensification strategy

are equivalent are (similar to the intensification strategy) both rejected at a p-value of 5 percent.

The second category of output results related to the arable farming submodel is the produced grain surpluses. As mentioned in the introduction of this submodel, no grain deficits are generated in both models as a result of the store size values applied in this experiment. The produced grain surpluses related to the intensification strategy are visualized in figure 26 . In this visualization it can be observed that the grain surpluses produced under the intensification strategies are fluctuating. The grain surpluses of the docked model fluctuate between 263 kg and 437 kg, whereas the grain surpluses in ROMFARMS fluctuate between 196 and 409 kg. The results of the docked model seem to portray a somewhat upward trend whereas the results of ROMFARMS portray a somewhat downward trend. Similar to previous statistical tests, the hypothesis that the distributions are equivalent are rejected by the Kolmogorov-Smirnov and Mann-Whitney U test at a p-value of 5 percent.



Extensification strategy - grain surplus		
	K-S Test	Mann-Whitney U
p-value	< 2.2E-16	< 2.2E-16

Figure 27: The output results of the models in relation to the production of grain surpluses with the extensification strategy.

The last output result is that of arable surplus production related to the extensification strategy, the results of which are visualized in figure 27 . The general shape of the plot with the average values per timestep is similar to that of the cultivated arable land with the extensification strategy. In this plot the drop of the cultivated area of arable land applies to the produced surplus as well. The highest produced surplus is 7411 kg of grain for the ROMFARMS model and 4779 kg of grain for the docked model. The hypothesis that the distributions are equivalent to each other are also rejected for these distributions.

From the analysis of the output results related to both the average areas of cultivated land and the produced grain surplus levels it can be concluded that no distributionally equivalent results were produced. The docked model produced higher values in relation to the arable land and grain surplus levels at the level of the intensification strategy, whereas the opposite is true for the extensification strategy. The docked model furthermore produces more constant values with marginal fluctuations in its output values, except for the produced grain surpluses at the level of the intensification strategy. One remarkable observation is the sudden drop and discontinuation in the initial upward trend of the extensification arable strategy in relation to both output results. It seems unclear at the moment how the model was able to consistently produce this pattern, the issue might require future examination since it might be a hint of a possible bug in the ROMFARMS source code.

7.2.3 Animal husbandry output results

The animal husbandry submodel has not been discussed in relation to the sensitivity analysis since no ontological overlap related to the animal breed was present between the models and its functionality has therefore not been altered compared to the original model. A similar internal functionality does, however, not mean that the initial input variables of the model are similar. Therefore a small experiment related to the demographic composition will be performed for the alignment of the output results as well. The experiment that has been performed utilizes the same parameters as the arable farming experiment, but with all the different animal species enabled. Figure 28 visualizes the average population development of the cattle, sheep, and horse species over the course of a simulation run with 75 timesteps, similar to previous alignment experiments.

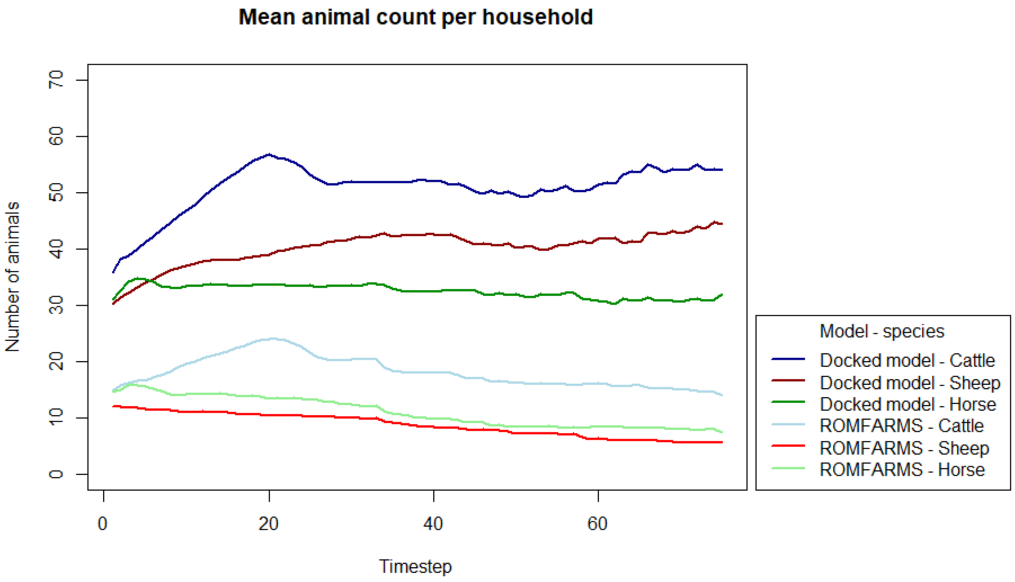


Figure 28:The output results of the models in relation to the development of the herds. The settlement composition related to the one expressed in the ROMFARMS-1 experiment.

As is visible in this figure 28, the number of animals for each species in consistently higher for the docked model compared to the original ROMFARMS. Statistical Kolmogorov-Smirnov and Mann-Whitney U tests furthermore yield p-values of the lowest computable values (2.2E-16) that reject the hypotheses that the distributions per species can be considered statistically similar at a p-value of 5 percent. From a visual perspective, however, the quantitative data yields comparable patterns in relation to the cattle (blue and light blue lines) and the horse (green and light green) lines. The initial period of growth is similar in both of these groups and the horse-related distributions show a comparable period of steady decline of population constitution. Where the cattle population slowly and steadily declines for ROMFARMS, however, the cattle population in the docked model stabilizes and experiences another period of increased population numbers. The only possible reason for this is that the change in calculation of the workforce for cattle herding is automatically a little higher than the ROMFARMS settlement population can suffice for.

The reason for the higher number of animals related to the docked model can, however, be easily explained by the setup of the models. In the docked model, the initial number of cattle, sheep, and horses (30 each) is determined on the household level. In ROMFARMS the initial number of 30 animals per species is determined on the settlement level. Settlements do however include multiple households, meaning that the average numbers of animals per household is per definition smaller. A ROMFARMS scenario where each household starts with 30 animals of each species can only be performed under the conditions that each settlement contains only one household. This experiment was therefore also conducted (see figure 29), and this indeed yielded higher results that were, to some extent, comparable to those of the docked model. The overlap between the development of the cattle and horse populations already visible in the previous experiment is even better observable due to the comparable population numbers. Only between the sheep levels many differences are still observed. The output results of the animal husbandry submodel therefore clearly show that the initial settlement composition has a strong influence on the results when the numbers per household are taken as the rate of measurement. An initial composition with only one-household settlements produces patterns in the data and associated population values that are visually (but not statistically) comparable to the data produced by the docked model. An initial composition of multiple-household settlements implies that more households have to share the same initial number of animals, making the number of animals available per household automatically lower.

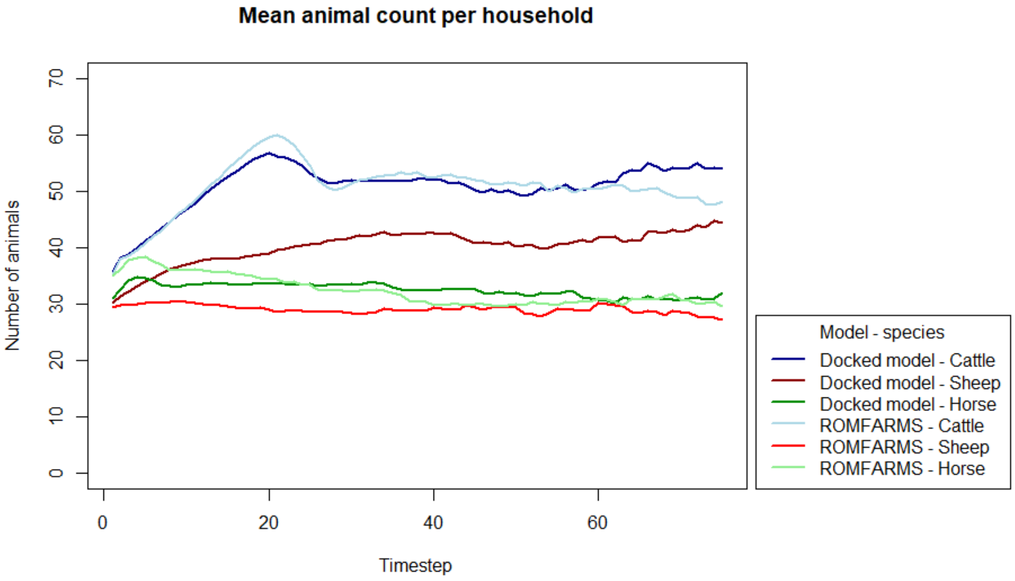


Figure 29: The output results of the models in relation to the development of the herds. The settlement composition related to the one expressed in the ROMFARMS-2 experiment.

7.2.4 The fuel submodel

The final category of model output results that will be discussed relates to the output results of the fuel and timber models. The output results of the fuel model will be discussed alongside the distance travelled and total labour costs, which are two of the four categories that were also discussed in the sensitivity analysis for both submodels.

The other possible output results, the fuel storage levels and the fuel processing time, are deliberately not utilized for the alignment of the output results due to the fact that these develop in the exact same way as the population as a whole. As a result the produced data will, despite differences in numerical values, follow the exact same pattern as the population development discussed in the alignment of the results from the demography submodel.

The distance travelled and the total labour costs, however, are dependent on factors related to the simulation environment since the presence of biomass sources has an influence on both model outputs. The experiment design utilizes the similar population development from the arable farming alignment analyses and the same parameter values as the arable farming submodel with the intensification arable strategy enabled. For the fuel and timber submodel, the daily fuel use per household member is set to five, the collection frequency to 10 and the reconstruction frequency to 20.

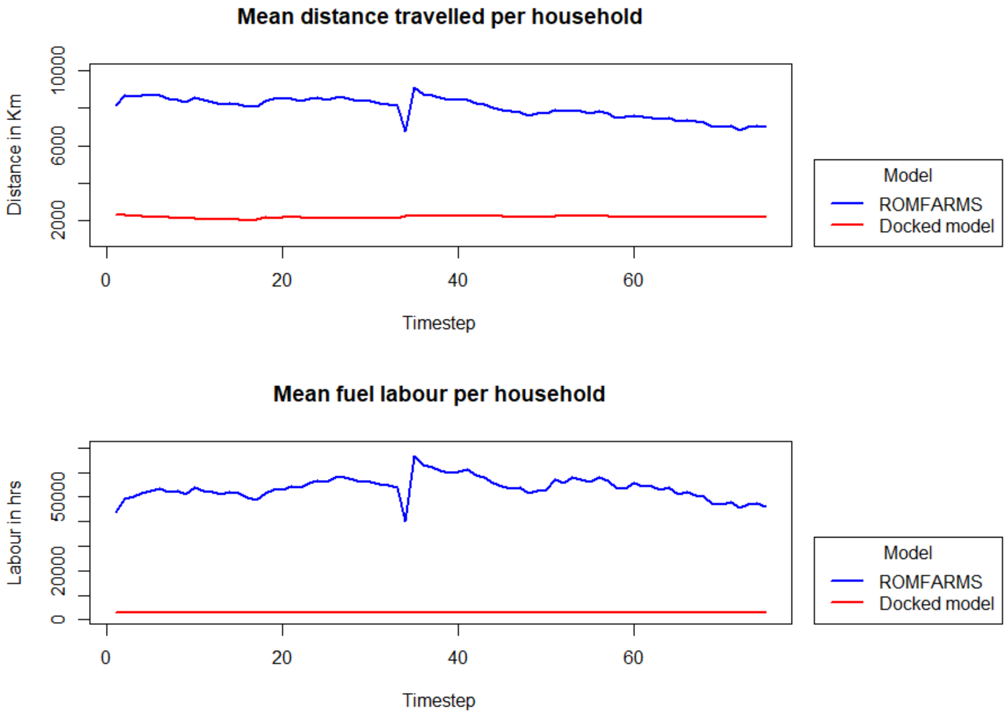


Figure 30: Output results of both models in relation to the distance travelled for fuel collection and the fuel labour.

Figure 30 visualizes the output results of the average distance travelled and the average labour costs per household in relation to the fuel collection submodel. This visualization makes clear that the general distance in kilometre (ca. 2500 km versus 8000 km) and the total hours of labour (ca. 3000 hours versus 53000 hours) for the households in the docked model is much lower in the docked model compared to ROMFARMS. Kolmogorov-Smirnov and Mann-Whitney U tests yield, similar to previously conducted tests on the arable farming and animal husbandry submodels, no indications that the distributions can be considered equivalent at a p-value of 5 percent.

The main differences in relation to the average distance travelled per household does not relate to the way the distance is calculated, which is exactly the same between the two models. The agent class that performs the travelled distance, however, is different. In ROMFARMS, the agents collecting the biomass for the fuel storages are the persons, of which there are many available. Distance is therefore calculated via the distance travelled per person. In the docked model, the biomass for fuel storages is collected by the household agent only. Because the results of the mean travelled distance relate only to the distance that the household agent (and not its implicitly modelled inhabitants) has travelled, produced travelled distance values are lower. The current output results are therefore the result of the intention to not adapt any of the original calculations, changes thus need to be made to the calculations to translate to the level of the person instead of the household. The same holds true for the calculated labour, which is the combination of travelled distance and fuel processing time expressed in hours of labour. As a result of the difference in travel distance calculations, the calculated labour hours per household fall out significantly lower compared to the ROMFARMS results as well.

The issue to solve this is to create a multiplier for the distance travelled by the household and the processed fuel time based on the number of household members that would be engaged in these activities: multiplying values by 2 when the female engages in additional fuel collection and by 6 if all the household members were engaged. Assuming that under the circumstances of a fuel use of 5 kg of biomass per household member the entire household is automatically engaged in the fuel collection and processing activities, all values would be multiplied by 6 for the travelled distance and by 36 for the labour activities (6 times compensation for distance and 6 times compensation for fuel processing). In that hypothetical case the results would reflect a pattern as expressed in figure 31, but it has to be noted that these results are purely for illustrative purposes and merely consist of elevated values that are not derived from real simulation data. It is clear that these values reflect much higher ranges, indicating that the docked model does not provide the level of nuance in the multiplier to achieve comparable ROMFARMS results. The ROMFARMS results, however, reflect the same discontinuity around the thirtieth timestep that was also observed in the results of the arable farming submodel. The reasons for the discontinuity of the patterns in both models in these results still remain unclear.

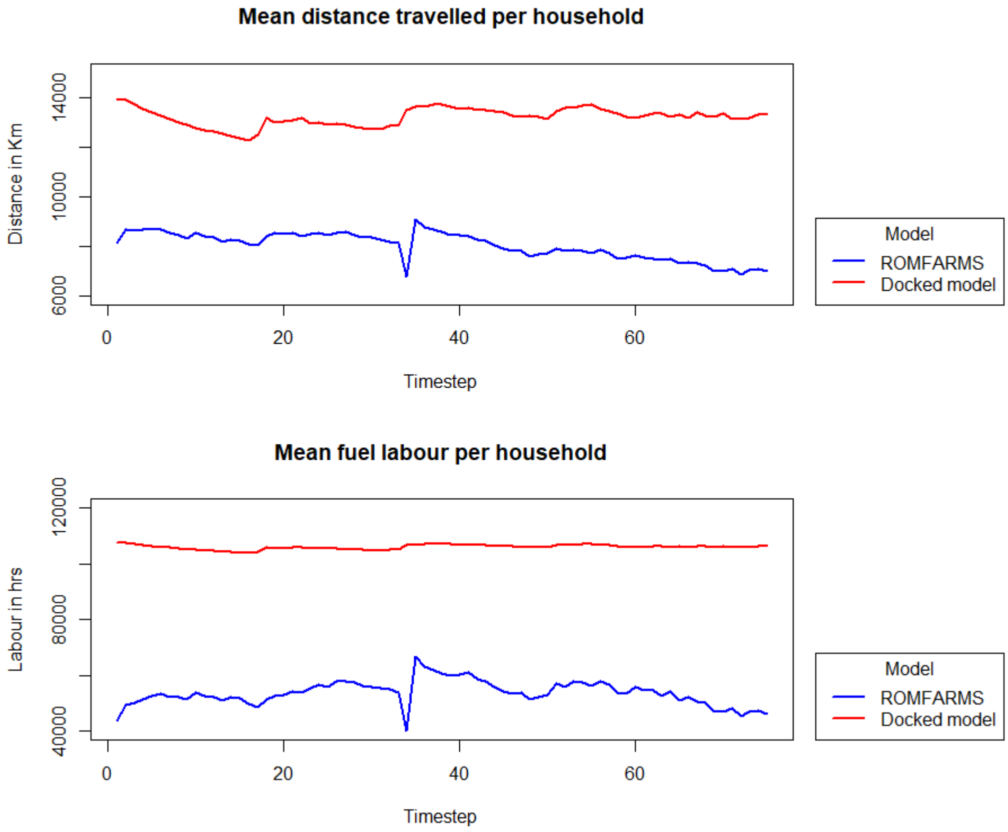


Figure 31: Hypothetical scenario where the results from the fuel collection where altered to the level of the person.

7.3 Results

The majority of the results from the sensitivity analysis are as expected. Initially surprising and unexpected results could be deduced from the source code of the model or by manually observing agent behaviour. In some cases the functionality of the docked model differs from ROMFARMS as a result of the docking process. The demography submodel functions as expected but shows no similarities with ROMFARMS, because this system has been implemented from the Artificial Anasazi model. Another differences that was found relates to a change in the calculation of fuel collection workforces, which has less nuance and therefore causes an abrupt decrease in the travelling distance for labour.

In relation to other aspects the results indicate that the majority of the functionality is as expected. The fact that model output generates expected behaviour arguably means that a high degree of relational equivalence is present in relation to all submodels except the demography submodel. In the case of the animal husbandry submodel no changes have been made to the source code except for the calculation of workforces, meaning that full relational equivalence is present. The relational equivalence can unfortunately not fully be established since Joyce does not include a sensitivity analysis of ROMFARMS.

The results from the model alignment show that distributional equivalence can be achieved between both models on the level of the demography level. This is relevant to the research question since the demography dynamics model from Artificial Anasazi replaced the person-centred demography submodel from ROMFARMS and was the major implementation in relation to the similar ontological components. The implementation of the Artificial Anasazi household structure, however, also affects the other submodel and generates model outputs that are not statistically similar. To achieve possible statistical similarity in relation to the other submodels the calculations used in ROMFARMS, which were intentionally not altered for verification purposes, have to be altered.

When investigating the output results from ROMFARMS' demography submodel, a mistake in the source code was found that prohibited the settling of new settlements in the simulation environment. This was caused by the fact that the settlement density remained at a value of zero for random generated landscapes. To assess whether an alteration of the settlement density affected general population development, a settlement slider was added to ROMFARMS to get a general impression whether this would be the case. These observations indicated that this would certainly be the case. The impressions were that a settlement density of 0.01 generally leads to a decline in population development, a settlement density of 0.0125 leads to a stable population and a settlement density of 0.015 leads to a growing population. Future researchers are thus encouraged to alter the settlement density values when working with random generated environments in ROMFARMS.

A final bug was observed in ROMFARMS that related to the arable farming and fuel collection submodel, where a discontinuation in pattern results was observable around the thirtieth timestep. In the source code of both models no observable bugs were however found. The fact that the discontinuation was visible between different submodels might indicate that there is an overarching structure that causes the disruption. The discontinuation was related to the possible lack of available resources, which is arable land for the farming submodel and biomass for the fuel collection submodel. The current hypothesis for the discontinuation therefore is that the procedure that regulates the regrowth of forest and land resources (the *update-patches* procedure), despite not having visible errors in its source code, reaches a point around the thirtieth timestep that does (temporarily) not allow the full regrowth of resources caused by resource demands of the population.

8. Discussion

This chapter discusses the results of the model comparison, docking and verification. The discussion is divided in three paragraphs that correspond with the each of these methodological phases. Each chapter will discuss the interpretations, limitations and implications of the results and makes suggestions for future research. The results of the model comparison and docking phase are of a methodological nature that does not directly correlate with answering the research question, their results will therefore primarily be discussed from a methodological perspective. The results of the model verification are primarily discussed in relation to the research question.

8.1 Comparing agent-based models

From the descriptive comparison of the two case studies, it was concluded that the highest degree of semantic similarity between the models could be established on the level of the household. This entity was implicitly modelled in ROMFARMS and explicitly modelled as an agent in the Artificial Anasazi model. Other similarities were the spatio-temporal dimensions of the two models, a timestep represented one year and one patch a hectare, and the rational economic thinking of agents. These results cannot be generalized on the level of archaeological agent-based land-use models as a whole. The results relate to a comparison of two specific case studies that are not representative for all archaeological agent-based land-use models. The explorative nature of this research furthermore implies that there are no results from other research to place these results in a wider context in relation to similarities between agent-based land-use models.

The choice for the application of the ODD + D framework for defining the similarity between the models was a conscious choice and it could generally be concluded that the comparison on basis of the ODD + D provided a better insight in the ontologies of the two case studies than the normal ODD would have done. Although the ODD framework would also have been able to identify that the household structures between the models would be semantically and that the spatio-temporal dimensions are identical, aspects like the rational economic thinking, objectives and sensing properties would not have been found if this framework was used. These insights might not have been decisive for the definition of the most relevant similarities between the models, but were certainly helpful in relation to the implementation of the models and better understanding them. These comparison results therefore illustrate that archaeologists should favour the ODD + D

framework over the regular ODD framework when making agent-based model descriptions.
It was also on the basis of this framework that the model descriptions that were used for the comparison were created. The available literature and source code that were used to fill in this framework as objective as possible sufficed to provide answers to all of the guiding questions. This led to the creation of ODD + Ds of the case studies that were of a good enough quality to perform the comparison of the models. The creation of the two descriptions did not take up a significant amount of time. The estimation of time it took to create the ODD + Ds is between 8-10 hours for ROMFARMS and 3-4 hours for Artificial Anasazi.

In agent-based model comparison studies like Cioffi-Revilla (2011) and An *et al.* (2014), the models that were compared were also developed by the same authors. This study however shows that the creation of model descriptions and the comparison of agent-based models is possible even when the researches has developed neither of the models. Having the model developed personally is therefore not a requirement to perform the comparison of these models. The implications of this conclusion are that more researchers should be able to independently perform model comparisons. If more model comparisons were consequentially performed and model descriptions generated or enhanced, the relationships between models and their general characteristics might become better accessible to those archaeologists that have a less technical background. This might enable them to engage in debate regarding model design, model development and epistemological implications of agent-based models in archaeology.

A limiting factor in the approach in relation to the comparison of models is however that the employed methodology in this study relies on the subjective interpretation of model descriptions and the source code. Even though a prescriptive framework and guiding questions are used provide some degree of objectivity, the results can therefore not be perceived as fully objective.
Another limiting factor that is especially relevant in the context of archaeological agricultural and land-use models, is that the ODD + D framework does not allow detailed descriptions of the simulation environment. Especially for this category of models, the relationship between the agent and its environment is a crucial factor for the model output. A better insight in the design of the simulation environment itself has the

potential to generate a better understanding of which aspects of the environments themselves are comparable and what the influences of different simulation environment designs would be on the output of a model. A further extension of the ODD + D protocol that emphasizes on the simulation environment is therefore suggested.

In relation to future research that might employ model comparisons, researchers are encouraged to also perform model comparisons on agent-based models that might not necessarily have similar research contexts. Whereas this research uses case studies that both have an agricultural and land-use component, this methodology should technically also be able to define relationships between models with different research contexts. This might help in improving the perceived transparency of certain black-box models, and in the most positive scenario even lead to new and exciting research opportunities.

8.2 Docking models

The result of the docking process is the docked model, a functional agent-based model that has the household structure from the Artificial Anasazi model and its demography system implemented in the ROMFARMS model. During the development process of the docked model, different verification techniques were used to ensure that the original ROMFARMS source code remained as intact as possible.

Implementing the household agent and demography dynamics system from Artificial Anasazi in ROMFARMS was not a straightforward task. Arguably the most problematizing factor in the docking process was the complexity of the ROMFARMS model. This complexity made the source code of the model initially difficult to understand. The implementation of the household agent and it demographics furthermore touched upon the core structure of the model. As a result, alterations had to be made to all submodels in ROMFARMS to accommodate the functionality of the model with this new agent. The fact that more components of the model had to be altered posed a higher risk that the implementation would be less successful. A critical factor that influences the success of an implementation is thus the number of model components that are affected by the implementation.

However, it is not only the number of affected components that play a role in the success of the implementation. The fact that the implemented agents and demography dynamics from Artificial Anasazi were more simplistic than the existing infrastructure of ROMFARMS meant that significant parts of the code needed to be removed or reduced. Each removal or reduction brought a possibility that the functionality of the model was significantly altered or even prohibited. Based on these experiences it can be concluded that it is generally easier to perform the docking of models when the implementation process involves making a simple model more complex, than when a complex model must be simplified.

The final influential factor that affects the ease with which docking can take place and the success of the implementation, is the user-friendliness of the model. The ROMFARMS model cannot be directly initialized once it has been downloaded, one first has to get insight in the suitable parameter values to set the model up. The source code is furthermore structured in such a way that it is unclear how the submodels have been expressed in it. The ROMFARMS model can therefore not be considered a very user-friendly model to work with. As a result, extra steps had to be taken to acquire a good insight in how the idea of the implementation could be technically executed in the model. The technical implementation would have been much easier to determine if the user-friendliness of the model would have been higher. The success of the implementation is thus also dependent on the user-friendliness of the model since it makes the way the implementation can be technically more clear from the beginning. This consequentially leaves smaller room for mistakes in the implementation design.

Even though Axtell *et al.* (1996) were the first to come up with the term “docking” and their research was the first to perform a docking process, the authors do not elaborate on their own docking process. In archaeology, guidelines or useful insights in the docking process are generally lacking. The relevant factors named above contribute to the ease with which a successful docked model can be achieved. These insights could potentially aid those who are going to use the docking of models for their own research purposes in the future.

The execution of a docking process and the creation of a docked model however also has limitations. Despite the fact that a docked model is created where the demography submodel and other submodels function, the docking process is still a subjective process. The presence of a work log for the model and a chapter in this thesis that explains the alterations to the model in detail does not automatically enhance the overall transparency of the model. Actively working with the ROMFARMS model and creating the docked model made the model certainly more transparent on a personal level, but it is difficult to assess whether the reader of the sixth chapter and the work log also gains a better understanding of the ROMFARMS and docked model due to the subjectivity of the process. The gained transparency of the model on the personal level allows the formulation of well-grounded conclusions regarding the model output from the following methodological phase, but it is thus difficult to assess whether transparency in the docking process also leads to a higher transparency in the ROMFARMS model for the reader.

Future research that employs the docking of models must thus take several factors into account. A good general rule is that the more user-friendly and simplistic the model in which another model will be implemented is, the higher the initial chances of a successful implementation and the easier it is to achieve this. This insight does however not only apply to research that aims to employ the docking of models, it also applies to model development. If there is a need more alignment studies to understand the relationships between models and their explanatory power, ease of use and the level of complexity are critical factors that must be considered in the process of model development to allow more trust in the success of the docking process.

8.3 Model equivalence

In the verification phase, the docked model was assessed in relation to ROMFARMS on two types of model equivalence: relational equivalence and distributional equivalence. An assessment of relational equivalence was done based on a sensitivity analysis of the docked model and a comparison of its results with descriptions of known behaviour of the ROMFARMS model. From this comparison it was concluded that relational equivalence was present between all submodels except the demography model. The distributional equivalence was assessed by performing Kolmogorov-Smirnov and Mann-Whitney U tests on the outputs from the docked model and ROMFARMS under similar parameter conditions. From this comparison distributional equivalence was only observed in relation to the demography submodel.

Particularities were found in the distributions of the ROMFARMS output. It was found that ROMFARMS includes a mechanism that does not allow the settling of new agents in the simulation environment. Furthermore, a disruption of patterns related to the presence of resources (available land and available wood) was observed in the distributions of the arable farming and fuel collection submodel at the thirtieth timestep of the simulation runs.

The main point of model implementation took place in relation to the household structure and the demography dynamics system from Artificial Anasazi. It is in relation to this primary point of model implementation that distributional equivalence was observed between the docked model and ROMFARMS. It can therefore be concluded that on the level of the population dynamics related to the households in both models, equivalence between the different algorithms can be observed. It must however be noted that this only applies on specific parameter combinations between the models. Despite the fact that distributional equivalence is observed, the implementation affected the rest of the model output to such an extent that no distributional equivalence could be established in relation to the other submodels.

To establish the equivalence between the models, the methodology from Axtell *et al.* (1996) was used. Their concept of distributional equivalence and the use of Mann-Whitney U test to investigate the matter of equivalence between models can still produce positive results. Contrary to the practice of model docking, where the authors provide little to no elaboration on the process, their approach to define distributional equivalence can thus still be considered applicable today.

Besides the notion of distributional equivalence, this research also managed to observe some malfunctions in the ROMFARMS model. The comparison of the output from both models allowed these patterns to be observed. The practice of aligning models can therefore not only be considered a relevant methodology to explore whether different models produce similar results, these results could also shed light on imperfections in the models. This allow the models to be improved as well based on the alignment results. Archaeological agent-based models are black-boxes, but this research shows that there are also still possibilities that there are elements in the functionalities of a model that can even be overlooked by the model developers themselves.

This research however does not provide sufficient proof to allow conclusions on whether the concept of relational equivalence is fully applicable. This is not a results of Axtell *et al.*'s research, but rather a lack of sensitivity analysis data related to ROMFARMS. Even though some of the observations in the sensitivity analysis of the docked model indeed matched with Joyce's (2019a) remarks on ROMFARMS' behaviour, the majority of the conclusions in relation to the relational equivalence are based on deductive assumptions and common logic. This is does not provide a fundament on which scientifically grounded conclusions can be drawn. The sensitivity analysis was originally intended to only observe the behaviour of the docked model, and it was only later realised that a sensitivity analysis can also provide insights in the relational equivalence of models. Time constraints however limited the possibility to also perform an elaborate sensitivity analysis of ROMFARMS. It is unfortunate that Joyce does not elaborate on the explorative experiments that were used for the calibration of the parameter ranges and values of the model. Even small descriptions of these experiments might have been useful to make more solid conclusions on the relational equivalence between both models.

This research focuses on the definition of equivalence between the algorithms of archaeological agent-based land-use models, and the docking process took place in relation to an agent-based models that functions as a theory building tool. Even though the methodology of this research can be applied to other models as well, future research is encouraged to also focus on performing the entire process of model alignment on models that actively attempt to interpret observations from the archaeological record. Agent-based models are powerful tools with explanatory power in relation to the interpretation of archaeological record. Model alignment can be used to enhance the explanatory power and quality of agent-based models in relation to archaeological phenomena, so its application should certainly not be limited to theory building and heuristic models.

9. Conclusion

This research has aimed to provide an answer to the question '*Can different algorithmic expressions of similar ontological components in archaeological agent-based land-use models be considered equivalent to each other?'* To answer this research question a research methodology was developed based on the research by Axtell *et al.* (1996). The research methodology was applied using the Artificial Anasazi model by Janssen (2009) and the ROMFARMS model by Joyce (2019a; 2019b) as case studies. First, the models were compared based on the ODD + D descriptions of both models to define similarities between their ontologies. This was followed by the creation of a docked model, where the algorithms from these similar components from the Artificial Anasazi model were implemented in the ROMFARMS model. This allowed the output from the docked model to be compared with the results from the ROMFARMS model to establish whether the different algorithms produce similar results and can thus be considered equivalent.

The results have shown that the algorithms from the similar ontological components between the docked model and ROMFARMS, the household structures and their population dynamics, provided model outputs that were statistically similar under some parameter conditions. It can therefore be concluded that the algorithms of these similar components, provided that the right parameters are used, are equivalent. However, the implementation from the Artificial Anasazi components in the ROMFARMS environment affected the output results of other model components to such an extent that no equivalence was established in relation to these components. To achieve a full model equivalent model, and not just equivalence between certain algorithms, adjustments therefore have to be made to the applied calculations in the docked model.

The first sub question that this research furthermore aimed to answer is '*Is the most suitable methodological approach towards the research question, which is derived from Axtell et al. (1996), still applicable?'* Based on the results of this thesis it can be concluded that the process of model alignment, which includes the docking of a model and comparing the output with its original counterpart, is still applicable. This research however adds a methodological approach for the comparison of agent-based models that adds to Axtell *et al.*'s research. This approach allows the definition of which model components can best be used for the docking and alignment of models.

This research has furthermore added insights in relation to which models are most suitable for docking. It was concluded that user-friendliness and a lower level of complexity are critical factors in the ease with which a successful docked model can be established. Model developers that want to be able to dock and align their models to assess their relationships therefore have to take these factors into account.

The final sub question that this research has aimed to answer is ‘*How does the applied research methodology aid in improving the transparency of black-box agent-based models?’* Based on the results of the research, it can be concluded that the model comparison based on the ODD + D framework and the alignment of the output results of the docked model and ROMFARMS aided in improving the transparency of the case studies.

The descriptive comparison of the two case studies was performed based on the ODD + D framework, even though no ODD + D documents were initially available. These documents were manually created and the comparison provided results in relation so the similarities and differences between the models. Even when descriptive documents are not (completely) present, this research has thus shown that the transparency of black-box models can still be enhanced by external parties. Furthermore, establishing the relationships between (components of) models also enhances the overall transparency of these models.

The alignment of the docked model and ROMFARMS output was also able to uncover some malfunctions in the ROMFARMS model. Comparing the results of docked models and their original counterparts therefore also holds the potential enhance the transparency of models by defining in which areas they could still be improved.

It has to be noted that this research has been explorative to a very high degree. The results regarding the equivalence of the similar components from the case studies can therefore not considered representative for all agent-based land-use models. Only the performance of further research has the potential to establish whether the conclusions in relation to the established equivalence are generally true. The methodological approaches in this research, however, could potentially aid in better understanding different kinds of archaeological agent-based models that are being used in archaeology and how these models relate to each other on the ontological and functional level.

Abstract

The majority of agent-based models in archaeology, if not all, are black box models. The internal processes of these simulation models from input value to output results are generally difficult to understand and even more difficult to assess. The assessment of these models is this difficult because creating archaeological agent-based models is a very individual/team-specific operation, each developer or developing team might come up with different solutions for the same modelling challenges. Code or concepts might be transferable between models, but the fact that the produced models are black-box models makes it difficult to determine if, where and how useful materials can be used for the development of other models. Model descriptions only are generally not enough to solve this problem.

This thesis investigates the black-box modelling problem from the perspective of archaeological agent-based agricultural/land-use models. It aims to explore whether the different technical executions of similarities in the ontology (the total framework of entities and rules of interaction) between two archaeological land-use models can be considered equivalent to each other. The two applied case studies are the Artificial Anasazi replication by Janssen (2009) and ROMFARMS by Joyce (2019a; 2019b). Similarities between the ontologies of both models were defined via a descriptive comparison based on the ODD + D framework. The similar ontological components, the household structure and its corresponding demography dynamics, from the Artificial Anasazi model were consequently implemented in the ROMFARMS model. This allowed for a comparison of the outputs from the docked model and the original ROMFARMS model to establish the level of equivalence between the models.

Statistical similarity was observed between the outputs of the demography dynamics systems from the two models under certain parameter conditions, meaning that the similar ontological components could be considered equivalent. The implementation of the household structure from Artificial Anasazi however significantly affected the input values for the calculations from the ROMFARMS model to such a degree that no further equivalence was established in relation to its other submodels.

This thesis therefore contributes to a better understanding of similarities between archaeological agent-based agricultural and land-use models from both a qualitative and quantitative perspective. It furthermore employs a methodology that can be applied to other agent-based models as well.

Internet sources

http://modelingcommons.org/browse/one_model/5687#model_tabs_browse_info, accessed on 12 February 2020.

https://www.comses.net/codebases/2222/releases/1.1.0/, accessed on 12 February 2020.

Bibliography

An, L., M.A. Lindeman, J. Qi, A. Shortridge and J. Liu, 2005. Exploring complexity in a human environment system: An agent-based spatial model for multidisciplinary and multiscale integration. *Annals of the Association of American Geographers* 95(1), 54-79. https://doi.org/10.1111/j.1467-8306.2005.00450.x

An, L., A Zvoleff, J. Liu and W. Axinn, 2014. Agent-Based Modeling in Coupled Human and Natural Systems (CHANS): Lessons from a Comparative Analysis. *Annals of the Association of American Geographers* 104(4), 723-745. https://doi.org/10.1080/00045608.2014.910085

Angourakis, A., M. Salpeteur, V. Martínez Ferreras, J.M. Gurt Esparraguera, 2017. The Nice Musical Chairs Model: Exploring the Role of Competition and Cooperation Between Farming and Herding in the Formation of Land Use Patterns in Arid Afro-Eurasia. *Journal of Archaeological Method and Theory* 24, 1177, 1202. https://doi.org/10.1007/s10816-016-9309-8

Axelrod, R., 1995. The Convergence and Stability of Cultures: Local Convergence and Global Polarization. *Santa Fe Institute working paper* 95-03-028.

Axtell, R., R. Axelrod, J.M. Epstein and M.D. Cohen, 1996. Aligning Simulation Models: A Case Study and Results. *Computational and Mathematical Organization Theory* 1(2), 123-141. https://doi.org/10.1007/BF01299065

Axtell, R.L., J.M. Epstein, J.S. Dean, G.J. Gumerman, A.C. Swedlund, J. Harburger, S. Chakravarty, R. Hammond, J. Parker and M. Parker, 2002. Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley. *PNAS* 99(3), 7275-7279. https://doi.org/10.1073/pnas.092080799

Bandini, S., S. Manzoni and G. Vizzari, 2009. Agent Based Modeling and Simulation: An Informatics Perspective. *Journal of Artificial Societies and Social Simulation* 12(4)4, 1-16. http://jasss.soc.surrey.ac.uk/12/4/4.html

Baum, T., 2016. Simulating Land Use of Prehistoric Wetland Settlements: Did Excessive Resource Use Necessitate a Highly Dynamic Settlement System?, in J.A Barceló and F. Del Castillo (eds.) *Simulating Prehistoric and Ancient Worlds.* Dordrecht: Springer, 255-280.

Bergin, S.M., 2016. Mechanisms and Models of Agropastoral Spread During the Neolithic in the West Mediterranean: The Cardial Spread Model. PhD Dissertation Arizona State University.

Breitenecker, F., M. Bicher and G. Wurzer, 2015. Agent-Based Simulation in Archaeology: A Characterization, in G. Wurzer, K. Kowarik and H. Reschreiter (eds.), *Agent-based Modeling and Simulation in Archaeology.* Berlin: Springer, 53-75.

Carley, K.M., 2002. Computational organizational science and organizational engineering. *Simulation Modelling Practice and Theory* 10, 253-269. https://doi.org/10.1016/S1569-190X(02)00119-3

Cioffi-Revilla, C., 2011. Comparing Agent-Based Computational Simulation Models in Cross-Cultural Research. *Cross Cultural Research* 45(2), 208-230. https://doi.org/10.1177/1069397110393894

Cegielski, W.H., and J.D. Rogers, 2016. Rethinking the role of Agent-Based Modeling in Archaeology. *Journal of Anthropological Archaeology* 41, 283-298. http://dx.doi.org/10.1016/j.jaa.2016.01.009

Clarke, D.L., 1972. Models and paradigms in contemporary archaeology, in D.L. Clarke (ed.) *Models in Archaeology.* London and New York: Routledge, 1-60.

David, N., 2006. Validation and Verification in Social Simulation: Patterns and Clarification of Terminology, in F. Squazzoni (ed.), *Epistemological Aspects of Computer Simulation in the Social Sciences*. Berlin: Springer, 117-129.

Dean, J.S., R.C. Euler, G.J. Gumerman, F. Plog, R.H. Hevly and T.N.V. Karlstrom, 1985. Human Behavior, Demography, and Palaeoenvironment on the Colorado Plateaus. *American Antiquity* 50(3), 537-554. https://doi.org/10.2307/280320

Dean, J.S., G.J. Gumerman, J.M. Epstein, R.L. Axtell, A.C. Swedlund, M.T. Parker and S. McCarroll, 2000. Understanding Anasazi Culture Change through Agent-Based Modeling, in T.A. Kohler and G.J. Gumerman (eds.) *Primate Societies: Agent-Based Modeling of Social and Spatial Processes.* New York: Oxford University Press, 179-206.

Dean, J.S., G.J. Gumerman, J.M. Epstein, R.L. Axtell, A.C. Swedlund, M.T. Parker and S. McCarroll, 2006. Understanding Anasazi Culture Change through Agent-Based Modeling, in in J.M. Epstein (ed.) *Generative Social Science: Studies in Agent-Based Computational Modeling.* New Jersey: Princeton University Press, 90-116.

Diamond, J., 2002. Life with the artificial Anasazi. *Nature* 419(6907), 567-568. https://doi.org/10.1038/419567a

Diamond, J. and P. Bellwood, 2003. Farmers and Their Languages: The First Expansions. *Science* 300, 597-603. https://doi.org/10.1126/science.1078208

Epstein, J.M., 2006. Agent-Based Computational Models and Generative Social Science, in J.M. Epstein (ed.) *Generative Social Science: Studies in Agent-Based Computational Modeling.* New Jersey: Princeton University Press, 4-46.

Epstein, J.M., and R. Axtell, 1995. *Growing Artificial Societies: Social Science from the Bottom Up.* Washington DC: The Brookings Institution.

Grimm, V., U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. Heinz, G. Huse, A. Huth, U.J. Jepsen, C. Jørgensen, W.M. Mooij, B. Müller, G. Pe'er, C. Piou, S.F. Railsback, A.M. Robbins, M.M. Robbins, E. Rossmanith, N. Rüger, E. Strand, S. Souissi, R.A. Stillman, R. Vabø, U. Visser, D.L. DeAngelis, 2006. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling* 198, 115–126. https://doi.org/10.1016/j.ecolmodel.2006.04.023

Grimm, V., U. Berger, D.L. DeAngelis, J. G. Polhill, J. Giske and S.F. Railsback, 2010. The ODD protocol: A review and first update. *Ecological Modelling* 221, 2760-2768. https://doi.org/10.1016/j.ecolmodel.2010.08.019

Grimm, V., G. Polhill and J. Touza, 2017. Documenting Social Simulation Models: The ODD Protocol as Standard, in B. Edmonds and R. Meyer (eds.), *Simulating Social Complexity.* Berlin: Springer, 349-365.

Gumerman, G.J. and J.S. Dean, 1989. Prehistoric Cooperation and Competition in the Western Anasazi Area, in L.S. Cordell and G.J. Gumerman (eds.), *Dynamics of Southwest Prehistory.* Tuscaloosa: University of Alabama Press, 99-148.

Gumerman, G.J., A. Swedlund, J.S. Dean and J.M. Epstein, 2003. The Evolution of Social Behavior in the Prehistoric American Southwest. *Artificial Life* 9, 435-444. https://doi.org/10.1162/106454603322694861

Hartmann, S., 1996. The World as a Process: Simulations in the Natural and Social Sciences, in R. Hegselmann, U. Mueller and K.G. Troitzsch (eds.) *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View.* Berlin: Springer, 77-100

Heckbert, S., 2013. MayaSim: An Agent-Based Model of the Ancient Maya Social-Ecological System. *Journal of Artificial Societies and Social Simulation* 16(4)11, 1-12. https://doi.org/10.18564/jasss.2305

Huggett, J., 2004. Archaeology and the new technological fetishism. *Archaeologia e Calcolatori* 15, 81-92.

Janssen, M., L. Na'ia Alessa, M. Barton, S. Bergin and A. Lee, 2008. Towards a Community Framework for Agent-Based Modelling. *Journal of Artificial Societies and Social Simulation* 11(2)6, 1-13. http://jasss.soc.surrey.ac.uk/11/2/6.html

Janssen, M., 2009. Understanding Artificial Anasazi. *Journal of Artificial Societies and Social Simulation* 12(4)13, 1-17. http://jasss.soc.surrey.ac.uk/12/4/13.html

Janssen, M., 2013. *ODD of Netlogo implementation of Artificial Anasazi.* Unpublished model description.

Joyce, J., 2019a. *Farming along the limes: Using agent-based modelling to investigate possibilities for subsistence and surplus-based agricultural production in the Lower Rhine delta between 12BCE and 270CE.* Vrije Universiteit Amsterdam Research Portal.

Joyce, J., 2019b. Modelling Agricultural Strategies in the Dutch Roman Limes via Agent-Based Modelling (ROMFARMS), in P. Verhagen, J. Joyce and M.R. Groenhuijzen (eds.), *Finding the Limits of the Limes: Modelling Demography, Economy and Transport on the Edge of the Roman Empire*. Dordrecht: Springer, 109-128.

Kanters, H., 2019. *Replication of agent-based models in archaeology: A case study using Brughmans and Poblome's MERCURY model.* Leiden University Repository.

Kohler, T.A., 2012. Complex Systems and Archaeology. *Santa Fe Institute working paper* 2011-06-023.

Lake, M.W., 2014. Trends in Archaeological Simulation. *Journal of Archaeological Method and Theory* 21, 258-287. https://doi.org/10.1007/s10816-013-9188-1

Lake, M.W., 2015. Explaining the Past with ABM: On Modelling Philosophy, in in G. Wurzer, K. Kowarik and H. Reschreiter (eds.), *Agent-based Modeling and Simulation in Archaeology.* Berlin: Springer, 3-35.

Leeuw, S. van der, 2004. Why Model? *Cybernetics and Systems* 35(2-3), 117-128. https://doi.org/10.1080/01969720490426803

Livet, P., J. Müller, D. Phan and L. Sanders, 2010. Ontology, a Mediator for Agent-Based Modeling in Social Science. *Journal of Artificial Societies and Social Simulation* 13(1)3, 1-14. http://dx.doi.org/10.18564/jasss.1538

Matthews, R.B., N.G. Gilbert, A. Roach, J.G. Polhill and N.M. Gotts, 2007. Agent-based land-use models: a review of applications. *Landscape Ecology* 22, 1447-1459. https://doi.org/10.1007/s10980-007-9135-1

Mitchell, M., 2009. *Complexity: A Guided Tour.* New York: Oxford University Press.

McGlade, J., 2014. Simulating as Narrative: Contingency, Dialogics, and the Modeling Conundrum. *Journal of Archaeological Method and Theory* 21, 288-305. https://doi.org/10.1007/s10816-014-9201-3

Müller, B., S. Balbi, C..M. Buchmann, L. de Sousa, G. Dressler, J. Groeneveld, C.J. Klassert, Q. Bao Le, J.D.A. Millington, H. Nolzen, D.C. Parker, J.G. Polhill, M. Schlüter, J. Schulze, N. Schwarz, Z. Sun, P. Taillandier and H. Weiise, 2013a. Standardised and transparent model descriptions for agent-based models: Current status and prospects. *Environmental Modelling & Software* 55, 156-163. https://doi.org/10.1016/j.envsoft.2014.01.029

Müller, B., F. Bohn, G. Dressler, J. Groeneveld, C. Klassert, R. Martin, M. Schlüter, J. Schulze, H. Weise, N. Schwarz, 2013b. Describing human decisions in agent-based models - ODD + D, an extension of the ODD protocol. *Environmental Modelling & Software* 48, 37-48. https://doi.org/10.1016/j.envsoft.2013.06.003

Niaz Arifin, S.M., G.J. Davis and Y. Zhou, 2010. Verification & validation by docking: a case study of agent-based models of Anopheles gambiae. *Proceedings of the 2010 Summer Computer Simulation Conference*, 236-243.

Ormerod, P., and B. Rosewell, 2006. Validation and Verification of Agent-Based Models in the Social Sciences, in F. Squazzoni (ed.), *Epistemological Aspects of Computer Simulation in the Social Sciences*. Berlin: Springer, 130-139.

Parker, D.C., D.G. Brown, J.G. Polhill, P.J. Deadman and S.M. Manson, 2008. Illustrating a new 'conceptual design pattern' for agent-based models and land use via five case studies: The MR POTATOHEAD framework, in A. Lopez Paredes, C. Hernandez Iglesias (eds.), *Agent-based modelling in natural resource management.* Valladolid: Unversidad de Valladolid, 23-51.

Parker, D.C., D.G. Brown, T. Filatova, R. Riolo, D.T. Robinson and S. Sun, 2012. Do Land Markets Matter? A Modeling Ontology and Experimental Design to Test the Effects of Land Markets for an Agent-Based Model of Ex-Urban Residential Land-Use Change, in A. Heppenstall, A. Crooks, L.M. See and M. Batty (eds.) *Agent-Based Models of Geographical Systems.* Dordrecht: Springer.

Plog, F., G.J. Gumerman, R.C. Euler, J.S. Dean, R.H. Hevly and T.N.V. Karlstrom, 1988. Anasazi adaptive strategies: the model, predictions and results, in G.J. Gumerman (ed.) *The Anasazi in a Changing Environment*. Cambridge (UK): Cambridge University Press, 230-276.

Riris, P., 2018. Assessing the impact and legacy of swidden farming in neotropical interfluvial environments through exploratory modelling of post-contact Piaroa land use (Upper Orinoco, Venezuela). *The Holocene* 28(6), 945-954. https://doi.org/10.1177/0959683617752857

Romanowska, I., 2015. So You Think You Can Model? A Guide To Building and Evaluating Archaeological Simulation Models of Dispersals. *Human Biology* 87(3), 169-192. https://doi.org/10.13110/humanbiology.87.3.0169

Romanowska, I., S.A. Crabtree, K. Harris and B. Davies, 2019. Agent-Based Modeling for Archaeologists: Part 1 of 3. *Advances in Archaeological Practice* 7(2), 178-184. https://doi.org/10.1017/aap.2019.6

Saqalli, M., A. Salavert, S. Bréhard, R. Bendrey, J. Vigne and A. Tresset, 2014. Revisiting and modelling the woodland farming system of the early Neolithic Linear Pottery Culture (LBK), 5600-4900 BC. *Vegetation history and archaeobotany* 23(1), 37-50. https://doi.org/10.1007/s00334-014-0436-4

Schulze, J., B. Müller, J. Groeneveld and V. Grimm, 2017. Agent-Based Modelling of Social-Ecological Systems: Achievements, Challenges, and a way forward. *Journal of Artificial Societies and Social Simulation* 20 (2), 1-31. https://doi.org/10.18564/jasss.3423

Stekerova, K., and A. Danielisova, 2016. Economic Sustainability in Relation to Demographic Decline of Celtic Agglomerations in Central Europe: Multiple Scenario Approach, in J.A Barceló and F. Del Castillo (eds.) *Simulating Prehistoric and Ancient Worlds.* Dordrecht: Springer, 335-358.

Wilensky, U., and W. Rand, 2015. *An Introduction To Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo.* Cambridge (MA): The MIT Press.

Wilkinson, T.J., M. Gibson and M. Widell, 2013. *Models of Mesopotamian Landscapes: How small-scale processes contributed to the growth of early civilizations.* Oxford: Bar Publishing (BAR International Series 2552).

Zvoleff, A., and L. An, 2014. Analyzing human-landscape interactions: tools that integrate. *Environmental Management* 53(1), 94-111. https://doi.org/10.1007/s00267-012-0009-1

List of figures

List of tables

Appendix 1: Terminology list

Term	Definition
Agent / Turtle	Heterogeneous and autonomous beings that follow instructions that have been expressed as computer code. They can be embedded with qualities and behaviours to make them represent real-life or hypothetical entities.
Agent-based model	A class of computational models that allows the observation of a system's behaviour based on the interactions of its individual components (agents).
Algorithm (Computer science)	A sequence of instructions, expressed in a formal language, that can be read by a computer.
Black-box model	A model whose inner workings between input and output are not known, since it relies on a foundation of theories, hypotheses, ideas, observations, and knowledge rather than proven laws and axioms.
Conceptual model	A simplified representation of a real-world system (concept) that has been created for a specific purpose. A conceptual model can be a physical or schematic representation of a real-world system.
Computational model	A conceptual model that has been implemented as a mathematical expression. It can be studied with computational resources.
Distributional equivalence	Statistical similarity between the quantitative output results of different models.
Docking	Implementing (components of) a computational model in the simulation environment of another computational model.
Formal language	A language whose structure and semantics are exact and thus not open for interpretation. Computer languages are formal languages.
High-level programming language	A computer language that has been written with a high degree of abstraction, which makes it easier to understand for humans.
Integrated Development Environment	A software program that assists a software developer (or model builder) with the development of computer software.
Low-level programming language	A computer language with a low level of abstraction, it is closer to machine code and therefore more difficult to understand for humans.
Model alignment	Having different computational models executed in the same simulation environment, and comparing their results.
Ontology (Model development)	The total framework of entities, relationships and rules of interaction within a domain or conceptual model.
Ontology (Computer science)	A data structure that formally defines the total framework of entities, relationships and rules of interaction within a domain or conceptual model.
Patches	Grid cells that constitute the digital simulation environment of an agent-based model. They can be embedded with qualities to make them represent real-life, hypothetical or abstract environments.
Relational equivalence	Similarity in model behaviour/output when the input values are similarly altered.
Semantic similarity (model development)	The closeness of entities, relationships, rules of interaction and processes (ontological components) between models based on their meaning or representation.
Simulation	The execution of a modelled process or system over time, which allows its behaviour to be analysed.
Simulation model	A computational model that can be used for simulation. Agent-based models are simulation models.
Source code	A human-readable collection of computer code.
Stochastic	With an element of randomness
Validation	The process of assessing whether a model is successful in reproducing the data of an observed phenomenon.
Verification	The process of assessing whether a model functions as intended.
White-box model	A model whose inner workings between input and output are known, this makes the model highly transparent.

Appendix 2: ODD + D overview from Müller *et al.* (2013b)

New guiding questions compared to to the original ODD protocol from Grimm *et al.* (2010) are expressed in **bold**.

Structural elements		**Guiding questions**
I) Overview	I.i Purpose	I.i.a What is the purpose of the study?
		I.i.b For whom is the model designed?
	I.ii Entities, state variables and scales	I.ii.a What kinds of entities are in the model?
		I.ii.b By what attributes (i.e. state variables and parameters) are these entities characterised?
		I.ii.c What are the exogenous factors/drivers of the model?
		I.ii.d If applicable, how is space included in the model?
		I.ii.e What are the temporal and spatial resolutions and extents of the model?
	I.iii Process overview and scheduling	I.iii.a What entity does what, and in what order?
II) Design Concepts	II.i Theoretical and Empirical Background	II.i.a Which general concepts, theories or hypotheses are underlying the model's design at the system level or at the level(s) of the submodel(s) (apart from the decision model)? What is the link to complexity and the purpose of the model?
		II.i.b On what assumptions is/are the agents' decision model(s) based?
		II.i.c Why is/are certain decision model(s) chosen?
		II.i.d If the model/submodel (e.g. the decision model) is based on empirical data, where do the data come from?
		II.i.e At which level of aggregation were the data available?
	II.ii Individual Decision-Making	**II.ii.a What are the subjects and objects of the decision-making? On which level of aggregation is decision-making modelled? Are multiple levels of decision making included?**
		II.ii.b What is the basic rationality behind agent decision-making in the model? Do agents pursue an explicit objective or have other success criteria?
		II.ii.c How do agents make their decisions?
		II.ii.d Do the agents adapt their behaviour to changing endogenous and exogenous state variables? And if yes, how?
		II.ii.e Do social norms or cultural values play a role in the decision-making process?
		II.ii.f Do spatial aspects play a role in the decision process?
		II.ii.g Do temporal aspects play a role in the decision process?
		II.ii.h To which extent and how is uncertainty included in the agents' decision rules?
	II.iii Learning	II.iii.a Is individual learning included in the decision process? How do individuals change their decision rules over time as consequence of their experience?
		II.iii.b Is collective learning implemented in the model?
	II.iv Individual Sensing	II.iv.a What endogenous and exogenous state variables are individuals assumed to sense and consider in their decisions? **Is the sensing process erroneous?**
		II.iv.b What state variables of which other individuals can an individual perceive? **Is the sensing process erroneous?**
		II.iv.c What is the spatial scale of sensing?
		II.iv.d Are the mechanisms by which agents obtain information modelled explicitly, or are individuals simply assumed to know these variables?
		II.iv.e Are the costs for cognition and the costs for gathering information explicitly included in the model?
	II.v Individual Prediction	II.v.a Which data do the agents use to predict future conditions?
		II.v.b What internal models are agents assumed to use to estimate future conditions or consequences of their decisions?
		II.v.c Might agents be erroneous in the prediction process, and how is it implemented?
	II.vi Interaction	II.vi.a Are interactions among agents and entities assumed as direct or indirect?
		II.vi.b On what do the interactions depend?
		II.vi.c If the interactions involve communication, how are such communications represented?
		II.vi.d If a coordination network exists, how does it affect the agent behaviour? Is the structure of the network imposed or emergent?
	II.vii Collectives	II.vii.a Do the individuals form or belong to aggregations that affect and are affected by the individuals? Are these aggregations imposed by the modeller or do they emerge during the simulation?
		II.vii.b How are collectives represented?
	II.viii Heterogeneity	**II.viii.a Are the agents heterogeneous? If yes, which state variables and/or processes differ between the agents?**
		II.viii.b Are the agents heterogeneous in their decision-making? If yes, which decision models or decision objects differ between the agents?
	II.ix Stochasticity	II.ix.a What processes (including initialisation) are modelled by assuming they are random or partly random?
	II.x Observation	II.x.a What data are collected from the ABM for testing, understanding and analysing it, and how and when are they collected?
		II.x.b What key results, outputs or characteristics of the model are emerging from the individuals? (Emergence)
III) Details	III.i Implementation Details	**III.i.a How has the model been implemented?**
		III.i.b Is the model accessible, and if so where?
	III.ii Initialisation	III.ii.a What is the initial state of the model world, i.e. at time t ¼ 0 of a simulation run?
		III.ii.b Is the initialisation always the same, or is it allowed to vary among simulations?
		III.ii.c Are the initial values chosen arbitrarily or based on data?
	III.iii Input Data	III.iii.a Does the model use input from external sources such as data files or other models to represent processes that change over time?
	III.iv Submodels	III.iv.a What, in detail, are the submodels that represent the processes listed in 'Process overview and scheduling'?
		III.iv.b What are the model parameters, their dimensions and reference values?
		III.iv.c How were the submodels designed or chosen, and how were they parameterised and then tested?

Appendix 3: ODD + D of Netlogo implementation of Artificial Anasazi

1/17/2013, additions to original in red by Stefan Weijgertse on 7 April 2020

The model description follows the ODD +D protocol for describing individual- and agent-based models (~~Grimm et al. 2006~~ Müller et al. 2013) and consists of ~~seven~~ all seventeen elements. ~~The first three elements provide an overview, the fourth element explains general concepts underlying the model's design, and the remaining three elements provide details.~~

The model described is the Artificial Anasazi model as reported in Dean et al. (2000), Axtell et al. (2002) and Gumerman et al. (2003). This ODD + D is employed for a comparison with the ODD + D (Weijgertse 2020) of the ROMFARMS model from Joyce (2019). The writer of this ODD + D has not developed the Artificial Anasazi NetLogo implementation model but has utilized the original ODD (Janssen 2013), literature (Janssen 2009; Axtell et al. 2002; Dean et al. 2000; Gumerman et al. 2003) and the source code of the model to create the ODD + D.

Overview

1.1 Purpose

The purpose of this model is to explore the causes of the indirectly observed population dynamics in the Long house Valley in Arizona between 800 and 1400 AD. Does the environmental variability itself explain the abandonment of the Anasazi of the Long House Valley? Does a simple model of household rules on choosing locations for farms and settlements can reproduce the archaeological records of occupation in Long House Valley?

This model is a replication and is aimed to get a better understanding of the original Artificial Anasazi model and attempts to reproduce its results. Any archaeologist interested in the original model by Dean et al. (2000) and Axtell et al. (2002) can also use it to get a better understanding of the original model.

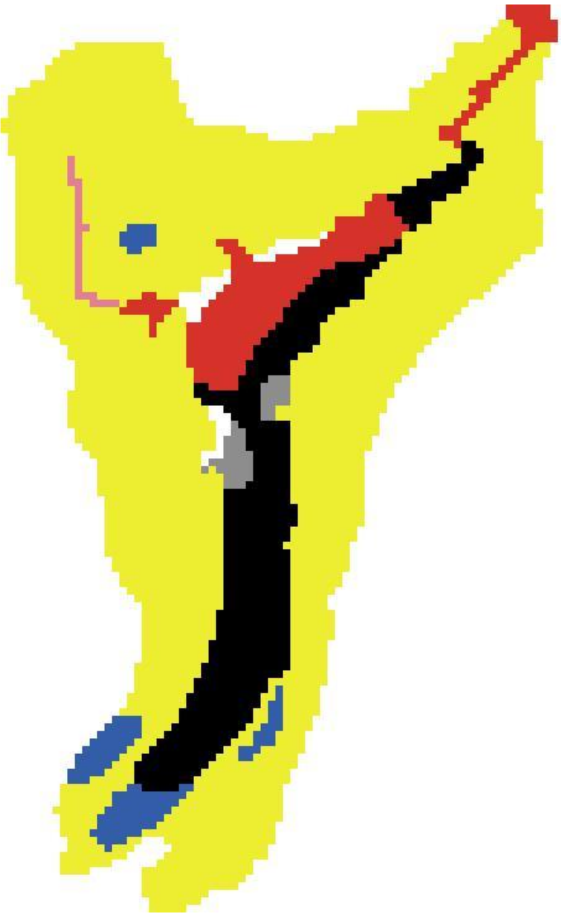


Figure 1: Different zones of land cover:

1.2 Entities, State variables and scales

Each agent represents a household of 5 persons. Each household makes annual decisions on where to farm and where to settle.

The attributes of a household:

- age,
- harvest during the last year
- estimated amount of food for current year
- fertilityAge (when household splits off a new household)
- stock of corn left over from previous years

- nutritionneed, which is the amount of food a household needs a year.

Each cell represents a 100mx100m space. Each cell is within one of the different zones of land: General Valley Floor, North Valley Floor, Midvalley Floor, Arable Uplands, Uplands Nonarable, Kinbiko Canyon or Dunes. These zones have agricultural productivity that is determined by the Palmer Drought Severity Index (PDSI) as discussed later.

Space is determined by a set of text files representing the map, precipitation, settlements, environmental zones and water sources. They are explained in more detail at the 'Input data' section.

Black: General Valley Floor
Red: North Valley Floor
White: Mid and North Dunes
Gray: Midvalley Floor
Yellow: Nonarable Uplands
Blue: Arable Uplands
Pink: Kinbiko Canyon

A household also has a physical place to live, which is called a settlement. Such a settlement can be occupied by multiple households. Every timestep that these households act represents a year.

1.3 Process overview and scheduling

Every year the following sequence of calculations is performed:
1. calculate the harvest for each household
2. if agent derives not sufficient food from harvest and storage or the age is beyond maximum age of household the agent is removed from the system
3. calculate the estimated harvest for next year based on corn in stock and actual harvest from current year
4.agents who expect not to derive the required amount of food next year will move to a new farm location.
5. find a farm plot
6. find a plot to settle nearby
7. If a household is older than the minimum fission age, there is a probability p_f, that a new household is generated. The new household will derive an endowment of a fraction f_{cs} of the corn stock. A household dies when it has reached the age of 30.
8. Update water sources based on input data
9. Household ages with one year.

Design concepts

2.1 Theoretical and Empirical Background

The main agent in this model is the household, which consists of five people. This number is based on archaeological and ethnographical analyses. Biological anthropological and ethnographic studies were furthermore used to derive agent demographics, nutritional requirements and marriage characteristics. All of these factors play a role in the movement, abandonment and creation of households and thus the complex dynamics that this model aims to investigate.

The main model for decision-making is related to the movement of households throughout the environment. There is no empirical data underlying this movement system, but social data or theory underlying this data are also not mentioned in Dean et al. (2000). It seems to stem from the assumption for settlement placement, namely that a settlement needs to be as close to water as possible.

2.2 Individual Decision-making

The main subject in the decision-making is the household, the land patches and their maize yields are the main object. The objective for all households is to suffice in their nutritional requirements. Agents make decisions on where to settle and where to harvest based on the harvest of their previous year and the surplus that they might have built up (a generated surplus collected in one year cannot be consumed anymore after two years). The dynamic simulated environment provides exogenous factors that will cause the agents to change their settlement location if needed and the space (location of water) has an influence on the location of new farming grounds and settlements. Social norms and cultural values do not play a role in the decision-making process

2.3 Learning

The households do not engage in any form of learning.

2.4 Individual Sensing

When the current location of a household does not suffice, the household senses for a new farming plot and location to settle. It chooses the most productive agricultural patch for harvesting within a one-mile radius and settles the household in a less productive zone. The agents thus search for the patch with the highest maize yield and patches close to a water source. These mechanisms are explicitly modelled and cannot be erroneous.

2.5 Individual prediction

The households predict their harvest rates for the next timestep based on their harvesting results from the previous timestep. Since the true yields are dependent on the precipitation from the PDSI index, the agents can be erroneous in their predictions

2.6 Interaction

Agents do not interact directly. Indirectly they interact by occupying potential farm plot. Agents are also marrying but this is implemented implicitly. In fact a household can derive an offspring at a certain agents, representing that the daughter leaves the house and starts a new household (immediately with 3 children that require 160 kg corn per year)

2.7 Collectives

Multiple households are able to settle on the same landscape patch, and thus forming a shared settlement. This is however an implied collective that has no function in the overall structure of the modelled system.

2.8 Heterogeneity

Agents are heterogeneous in terms of their age, location and collected or stored harvest. Their shared characteristics are their nutrition needs, length of corn storage, fertility age, death age, corn stock given to a newly generated household and their maximum distance between a residence and farm. Despite their heterogeneous characteristics, the agents are homogeneous in terms of their underlying decision-making processes.

2.9 Stochasticity

Initial conditions of cells and agents, and the order in which agents are updated.

2.10 Observation

The model produces an overview from the number of households at each timestep, simulating the trend of population dynamics. The model can also generate maps illustrating the locations where most households would have settled over specific periods of time.

The key result of the model is the overview of the number of households at each timestep, which match and even approximate the archaeological record better (figure 2c) than Axtell et al. (2002, figure 2a and 2b).

~~Adaptation. Agents adjust the location of farming and housing if they expect that current location is not sufficient next year.~~

~~Fitness. Fitness of agent is determined by the amount of harvest derived and amount of corn in storage. An agent is fit or not. If not sufficient food is available, it is removed from the system.~~

~~Prediction. The prediction of harvest next year is the same as the experienced harvest from this year.~~

Details

3.1 Implementation details

The Ascape 5.0.1 implementation available at http://sourceforge.net/projects/ascape/ is used to replicate the model in Net Logo 4.0.2. Unfortunately, the Ascape implementation is not well documented, and not well commented. The model is available via https://www.comses.net/codebases/2222/releases/1.1.0/. ~~We are however able to replicate the basic features of the model as discussed below. If errors are found in the implementation, please let me know: Marco.Janssen@asu.edu~~

~~Below we show the simulated population levels as published in Axtell et al. (2002) (Figure 2a), the default results of the model in Ascape 5.0.1 (Figure 2b) when the parameters are set to those described in Axtell et al. (2002), and the Net Logo implementation for the same parameter setting and the best of 100 runs using the L2 norm (Figure 2c).~~

3.2 Initialization

The initial number of households is 14. Each household is initialized by setting a household age from the uniform distribution [0, 29] and by setting the value of the corn stocks by the uniform distribution [2000, 2400].

The quality of the soil of the cells is initialized by adding a number drawn from the distribution n(0, σ_{shv}) where σ_{shv} is the spatial harvest variance.

3.3 Input Data

A number of files are imported with input data. Input data provides information which cells are water sources for which periods (rivers, wells, etc.). The following data files are imported.

map.txt

- defines for each cell what kind of landcover zone it belongs to: (General Valley Floor, North Valley Floor, Midvalley Floor, Arable Uplands, Uplands Nonarable, Kinbiko Canyon or Dunes)

adjustedPDSI.txt

- defines for each year and each category of landcover what the value of the adjusted Palmer Drought Severity is. This index provides measurements of moisture conditions for agricultural activities.

Table 1: Classifications of Palmer Drought Severity Index

Value	Classifications of Palmer Drought Severity Index
4.0 or more	extremely wet
3.0 to 3.99	Very wet
2.0 to 2.99	Moderately wet
1.0 to 1.99	Slightly wet
0.5 to 0.99	Incipient wet spell
0.49 to -0.49	Near normal
-0.5 to -0.99	Incipient dry spell
-1.0 to -1.99	Mild drought
-2.0 to -2.99	Moderate drought
-3.0 to -3.99	Severe drought
-4.0 or less	Extreme drought

environment.txt

- the description for each cell which zone it relates to and when the cell is a water source.

settlement.txt

- contains estimates for each excavated settlement the time period of occupation and population numbers during this period

water.txt

- defines locations of water points and period in which they contain water Furthermore, we have the following parameter values:

Table 2: Main parameters of the Artificial Anasazi model and their default values

Variable	value
Simulation period	800 AD to 1350 AD
Nutritional need per household	800 kg per year
Number of individuals per household	5 persons
Maximum length of corn storage	2 years
Harvest adjustment	1.00
Annual variance in harvest	0.1
Spatial variance in harvest	0.1
Minimum household age for fission	16 years
Maximum household age for fission	30 years
Household age for death (maximum age household)	30 years
Annual probability of fission p_f	0.125
Corn stock given to new household f_{cs}	0.33
Maximum distance between residence and farm	1600m

3.4 Submodels

We will discuss some steps of the model in more detail. The baseyield BY is defined by the yield y, quality of the soils in the cell q and the harvest adjustment level H_a,

$BY = y * q * H_a$

Where yield y is defined for each zone and each PDSI index. For example, the zones North Valley, Mid Valley and Kinbiko Canyon have the same rules for yield: IF PDSI >= 3.0 THEN y = 1153

```
IF 1 <= PDSI < 3.0 THEN y = 988
IF -1 < PDSI < 1 THEN y = 821
IF -3 < PDSI <= -1 THEN y = 719
IF PDSI <= -3 THEN y = 617
```

The harvest of a household H_0 is equal to the baseyield of the location BY adjusted by some annual variation of the harvest using a normal distribution and a standard deviation σ_{ahv} $H_0 = BY * (1 + n(0,\sigma_{ahv}))$

Before determining how much households consume, we need to define how much food is available. Every year, the stock of two years ago not used is disregarded. Subsequently the stock of 1 year ago S_{-1} will become the stock of 2 years ago S_{-2}, and the harvest of the current year will be input for S_0:

```
S-2 = S-1
$S_{-1} = S_0$
$S_0=H_0$
```

A household needs to derive 800 kg a year.
We start with a nutrition need remaining NNR = 800

```
If $S_{-2}$ >= NNR THEN
    $S_{-2} = S_{-2} – NNR$
    NNR=0
ELSE
    $NNR=NNR–S_{-2}$
    $S_{-2}=0$
    If $S_{-1}$ >= NNR THEN
        $S_{-1} = S_{-1} – NNR$
        NNR=0
    ELSE
        $NNR=NNR–S_{-1}$
        $S_{-1}=0$
        If $S_0$ >= NNR THEN
            $S_0=S_0–NNR$
            NNR=0
        ELSE
            $NNR=NNR–S_0$
            $S_0=0$
If NNR > 0 the agent is removed from the system
```

In order to determine whether the agent should stay or go it estimates the harvest for next year on the same location. This estimation is the amount of stored corn left plus the expected harvest (equal to current harvest level):

$E[H] = S_0 + S_{-1} + H_0$

Searching for a new farm is performed by identifying all unoccupied cells which produce more than the minimum nutrition requirement (800 kg) which are within 1 mile from a water source. If there are multiple suitable locations, chose the one closest to the current location

Searching for a settlement is performed by executing the following conditions:

i. The settlement location must be unfarmed (although it may be inhabited, i.e., multihousehold sites permitted).

ii. The settlement must be within 1 mile of the new agricultural plot selected.

iii. The settlement must be in a less productive zone than the new agricultural land selected.

If multiple sites satisfy these above criteria the location closest to the water resources is selected.

If no site meets these criteria, then first one looks at locations that meet condition i and

If still no site meets the criteria only sites who meet condition i are selected. Finally, if still no location is found, the agent leaves the system (which should not happen since the agent had already a settlement it came from).

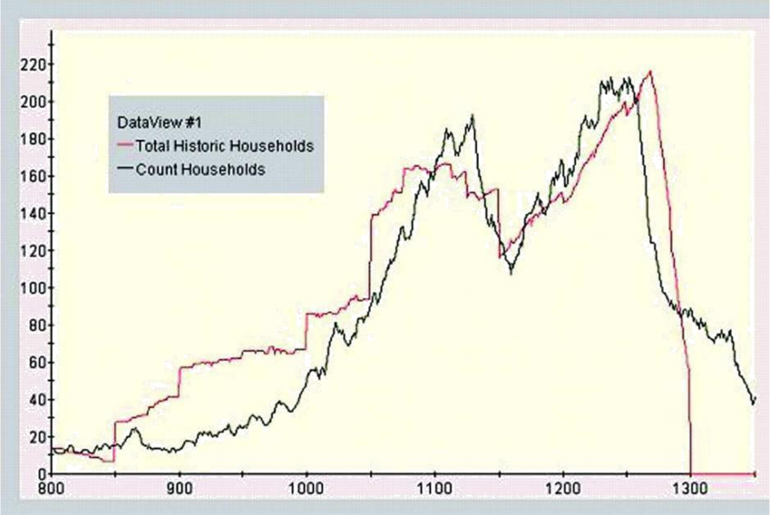


Figure 2a

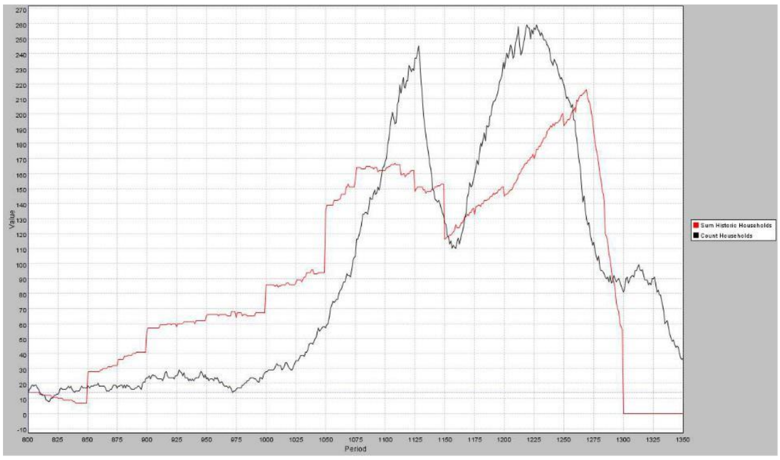


Figure 2b

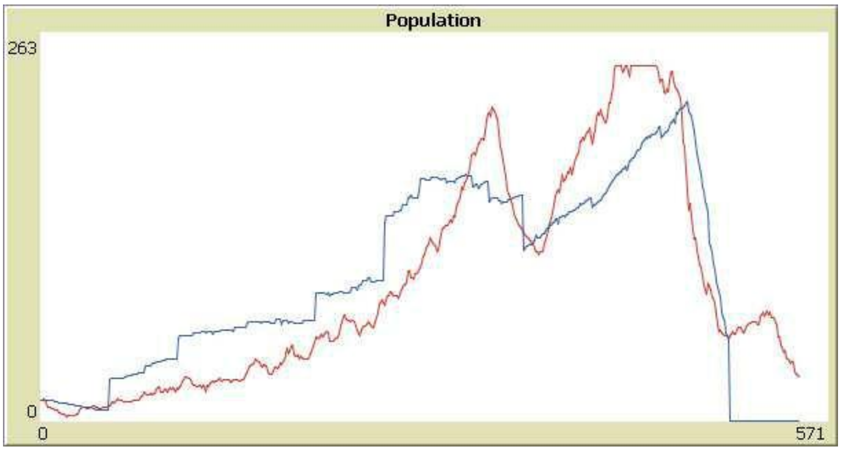


Figure 2c

References

Axtell R L, Epstein J M, Dean J S, Gumerman G J, Swedlund A C, Harburger J, Chakravarty S, Hammond R, Parker J, and Parker M. (2002) Population Growth and Collapse in a Multi-Agent Model of the Kayenta Anasazi in Long House Valley. Proceedings of the National Academy of Sciences of the United States of America 99(3): 7275-7279.

Dean J S, Gumerman G J, Epstein J M, Axtell R L, Swedlund A C, Parker M T, and McCarroll S. (2000) Understanding Anasazi Culture Change Through Agent Based Modeling in Dynamics in Human and Primate Societies: Agent Based Modeling of Social and Spatial Processes, T. Kohler and G. Gumerman (eds.), Santa Fe Institute. New York & London: Oxford University Press.

~~Grimm V, Berger U, Bastiansen F, Eliassen S, Ginot V, Giske J, Goss-Custard J, Grand T, Heinz S, Huse G, Huth A, Jepsen J U, Jørgensen C, Mooij W M, Müller B, Pe'er G, Piou C, Railsback S F, Robbins A M, Robbins M, Rossmanith E, Rüger N, Strand E, Souissi S, Stillman R A, Vabø R, Visser U, DeAngelis D L. (2006) A standard protocol for describing individual-based and agent-based models. Ecological Modelling 198:115-126.~~

Gumerman G J, Swedlund A C, Dean J S, and Epstein J M (2003) The Evolution of Social Behavior in the Prehistoric American Southwest. Artificial Life 9: 435-444.

Müller, B., F. Bohn, G. Dressler, J. Groeneveld, C. Klassert, R. Martin, M. Schlüter, J. Schulze, H.
Weise, N. Schwarz, 2013b. Describing human decisions in agent-based models - ODD + D, an
extension of the ODD protocol. *Environmental Modelling & Software* 48, 37-48.

Appendix 4: ODD + D of Joyce’s (2019) NetLogo implementation of ROMFARMS

Created by Stefan Weijgertse, 2020

The model description follows the ODD + D protocol for describing socio-ecological agent-based models (Müller *et al.* 2013). All elements of the framework are included, even when an element does not apply to this specific model. The described model is the ROMFARMS model as reported in Joyce (2019). This ODD + D is originally employed for a comparison with the ODD + D (Weijgertse 2020, adapted after Janssen 2009) of the Artificial Anasazi model from Janssen (2009). The writer of this ODD + D has not developed the ROMFARMS model, but has utilized literature (Joyce 2019) and the ROMFARMS source code to create this ODD + D.

Overview

1.1 Purpose

The purpose of the model is to generate different quantified spatio-temporal scenarios of agricultural subsistence strategies in the Lower Rhine delta, where the Dutch *limes* (northern border of the Roman Empire) was located, from 15 BC to 270 AD. It models the interactions between farming, animal husbandry, fuel use and timber collection. The model is part of the “Finding the Limits of the Limes” project (Verhagen 2011).

The model is developed to produce detailed results for the baseline economy of the Lower Rhine delta and is a heuristic toolkit for those interested in the investigation of Roman agriculture in the past, as well as agriculture in general.

1.2 Entities, state variables and scales

The model includes three different kinds of agent: The settlement, the person and the animal. At each timestep, which represents a year, a settlement engages in arable farming, animal husbandry, fuel collection and (only every 20 timesteps) timber collection. The following description names the attributes of the different agents. Please note that an attribute can both refer to a single or overarching group of agent variables.

The main attributes of a settlement are:

- Sex
- Age
- Spouse
- House
- Travelled distance
- Number of children
- Mortality rate
- Fertility rate
- Carrying volume and maximum capacity
- Fuel and timber processing time

A settlement sprouts six persons: two 16-year adults and 4 sub-adults per household. Each person has the following attributes:

- Sex
- Age
- Spouse
- House
- Travelled distance
- Number of children
- Mortality rate
- Fertility rate

- Carrying volume and maximum capacity
- Fuel and timber processing time

Each settlement produces, when animals are enabled, a herd of animals (sheep, cattle and horse). The attributes of an animal are:

- Age
- Sex
- Species
- Fertility rate
- Survivorship
- Owner
- Milk producing?

Each patch represents one hectare (100x100 meter) and has the following attributes:

- Raster value
- Owner
- Landscape type (Levee, flood-basin or undefined)
- Biomass
- Restoration countdown after being worked

The total number of patches forms the landscape. It is randomly generated based on the area parameter in the interface screen. In Joyce's research GIS Datasets from the Lower Rhine delta were used, but the model derived from modellingcommons.org does not come equipped with this GIS data.

Despite its reliance on palaeoenvironmental data, this data is included as a closed system. No specific exogenous factors are present in this model.

1.3 Process overview and scheduling

At each timestep and after initialization, the following sequence of events occurs:
1. New births, deaths and marriages within each settlement population;
2. Based on these population dynamics the creation of new, or abandonment of old, settlements;
3. Present settlements calculate grain needs, needed land to cultivate these yields, which land can actually be cultivated, total grain yield and the amount of labour needed to do so;
4. The settlements harvest grain;
5. New births and deaths of animals, while also calculating slaughter rates, needed land, required labour and yields for animal husbandry;
6. The settlements engage in animal husbandry and require resources from the animals;
7. Calculate fuel requirements and required labour to collect fuel;
8. Collect fuel;
(Every 20 timesteps, perform step 7 and 8 for Timber collection as well)
9. Update all age variables with one year, (partial) regrowth of patch variables;
10. Repeat step 1

Design concepts

2.1 Theoretical and Empirical Background

The main concept behind ROMFARMS is that all settlements pursue a mixed farming strategy: farming, animal husbandry and fuel/timber collection. This idea is based on a combination of literary sources (this applies to farming and animal husbandry) and expert judgment (this applies to fuel and timber collection). This mixed farming strategy creates a complex and holistic agricultural subsistence system and allows for an evaluation of the impact that different agricultural subsistence strategies would have had on the land. On the level of the submodels, use is made of palaeoenvironmental and archaeozoological data, ethnographical parallels, (ancient) literary sources, economic data and (as few as possible) expert judgment. Data aggregation levels vary between the level of the settlement, household and individual for all submodels.

The submodel for population dynamics utilizes calorific/nutritional needs based on data from the United Nations Farming and Agriculture Organization. Grain caloric contents and abundance in the arable farming submodel is based on archaeobotanical evidence. For animal husbandry, slaughter rates are derived from system dynamics models and yields from archaeological experimental research.

Agents are assumed to make economically rational decisions by calculating current presence of resources, needs and available land. This is chosen because the ROMFARMS model aims to approximate a null scenario for different agricultural subsistence strategies, whose complexity can be expanded in the future.

2.2 Individual Decision-making

The subject of decision-making is the settlement. The objects of the decision-making are the other two classes of agents and the patches. One could argue that certain population factors, like the transfer of an adult to another settlement for marriage, is a decision on the individual level but this is still steered by the settlement. Settlements measure success by being able to provide for all the basic requirements of their inhabitants. Settlements make economically rational decisions and are driven by mathematical programming. Agents adapt their farming, animal husbandry and fuel/timber strategies based on new calculations after changes in the landscape (last steps of the process overview), the current composition of their inhabitants (first steps of the process overview) and the agricultural strategy they pursue. Social norms and cultural values are not incorporated in this model. Space plays a significant role in the sense that it determines farming, animal husbandry and fuel/timber collection possibilities for the new timestep. Settlements act based on information available at the moment of decision-making/calculation and do not take past events and future predictions into account. Settlements have thus no knowledge on possible outcomes of future events, but this uncertainty does not affect their decisions.

2.3 Learning

Both individual learning and collective learning are not present in this model.

2.4 Individual Sensing

Settlements engage in sensing when determining their arable land and pasture/meadow land required. Settlements also engage in sensing when locating the presence of wood and timber sources. The agents consequently send labour forces to work land, to manage the animals and to collect resources. After being send, persons can sense the presence of arable land or fuel/timber sources and work the land or collect resources. This sensing process cannot be erroneous and information obtainment is explicitly expressed as equations. The spatial sensing scale of the settlements is in theory unlimited but involves those patches that are closest to the currently worked area. Settlements with an agricultural extensification strategy would also more quickly decide to work new patches, whereas settlements with an intensification strategy would more quickly decide to work and manure the land more efficiently.

2.5 Individual prediction

Agents do not employ collected data to predict future conditions but are able to build a buffer when yields for a following year are relatively low.

2.6 Interaction

The model primarily focuses on the interaction of the submodels in a holistic framework of agricultural subsistence strategies. Agents only engage in interaction in the case of a deficit of grain yields, in which case they will borrow from neighbouring settlements to compensate for this deficit.

2.7 Collectives

Persons belong to a household and a settlement, whereas animals belong only to a settlement. Despite the household being a collective of two adult agents and four subadult agents, it is an assumed class that is not explicitly modelled. The total number of animals owned by a settlement comprises the herd, a class that is also not explicitly modelled.

2.8 Heterogeneity

All agent classes are heterogeneous. Settlements differ from each other in terms of number of households, inhabitants, livestock, resource yields and storage and worked land for different agricultural purposes. Persons differ in terms of sex, age, house, spouse, fertility rates, mortality rates and number of children. Animals differ in terms of age, sex, species and owner. Agents are however not heterogeneous in their decision-making and will all follow the same set of procedures.

2.9 Stochasticity

Stochasticity occurs in different forms. It is primarily related to population dynamics of persons and animals: Fertility, mortality and marriage (the latter only for persons) are events based on random probabilities. There is a possibility for the creation of random landscapes and resource contents and regrowth is (to a certain extent) random. Upon initialization, settlements are randomly distributed throughout the simulated environment.

2.10 Observation

The key data that ROMFARMS produces relates to the composition of settlement demographics, livestock constitution, resource quantities, resource consumption, constitution and efficiency of labour forces and exploited land.

Emerging patterns and key results of the ROMFARMS model are the relevance of surplus production each year to both feed population and use for sowing the next year. Furthermore, differences in land-use via arable intensification and extensification were visible. Settlements pursuing an intensification strategy often tend to have sufficient labour to work all the land but are often severely limited in exploiting their land to the maximum capacity. Settlements pursuing extensification are often limited in labour, due to the fact that larger areas are worked. Regardless of the exploited area, there was often enough land to be used for all settlements.

Details

3.1 Implementation Details

The model has been implemented in NetLogo (version 6.0.2) and is accessible via http://modelingcommons.org/browse/one_model/5687#model_tabs_browse_info

3.2 Initialisation

The state of the model at t = 0 is a model where no parameters are filled in. As such, the model is not functional when opening the NetLogo file. The downloaded model comes equipped with a .jpg image illustrating a setup for a random generated environment, but this does not represent a prescribed initial setup of the model.

3.3 Input Data

The model does not require any external input data, but only when generating a random environment. Joyce has utilized geographical GIS data in his study, but the model does not come equipped with this spatial data. The model does furthermore not explicitly use processes directly derived from other models.

3.4 Submodels

The model consists of five different submodels:

- Population dynamics: The population dynamics submodel focuses on the births, deaths, marriages and transfer of inhabitants. Children are generated when a random number is smaller than a married female's fertility rate. A person dies when a randomly generated number is smaller than their mortality rate. A person marries if it is an unmarried adult and there is another unmarried adult in the own or another settlement, one of the spouses moves to the settlement of the other based on the max number of possible households of a settlement. In the case both settlements have reached their household maximum, both spouses create a new settlement or are removed from the simulation when maximum settlement density has been reached. When a settlement has no adults, it is removed and remaining subadults are moved to the nearest settlement. The simulation ends when there are no settlements for the subadults to be moved to.
- Arable farming: The settlement calculates the required arable land and the available maximum area of arable land. Based on availability, either the maximum area (when the minimum is higher than what is available) or the minimum area will be cultivated. Based on extensive strategy extra sowing seed is used to cultivate more land and based on intensive strategy the cultivated land is manured. The yield is calculated and if a settlement has a deficit, it will borrow grain from a neighbouring settlement. Variables will then be updated.
- Animal husbandry: This submodel focuses on the reproduction, natural mortality and slaughter of animals. Adult female animals will produce offspring when a random number is higher than the fertility rate, sheep have an added random probability of producing more than one offspring relative to cattle and horses. Newborn animals have a neonatal mortality rate and adult animals have a natural mortality rate, if a random number is higher than this rate the animal will die. After slaughter rates are set, all male neonates are slaughtered and based on the slaughter rates other animals are being slaughtered if a random number is lower than the slaughter rate.
- Fuel collection: This submodel focuses on the determination of the workforce, the finding of fuel and the collection of fuel. The settlement calculates the required fuel. If the max load of all adult male inhabitants is higher than the required fuel, only they will be employed. If the required fuel is higher, female adults are also employed and if the requirements are still higher, elders and adolescents are being employed as well. The persons then move throughout the landscape and select the cells with wood and consequently collect the maximum load of wood. If fuel has been collected it is consumed and a surplus or deficit is noted.
- (Every 20 timesteps) timber collection: Identical to the fuel collection submodel.

References

Joyce, J., 2019. *Farming along the limes: Using agent-based modelling to investigate possibilities for subsistence and surplus-based agricultural production in the Lower Rhine delta between 12BCE and 270CE.* Vrije Universiteit Amsterdam Research Portal.

Verhagen, J.W.H.P., 2011. Finding the Limits of the Limes project: https://www.nwo.nl/onderzoek-en-resultaten/onderzoeksprojecten/i/71/8871.html

Appendix 5: ODD + D of the docked model

Made by Stefan Weijgertse, 2020

The model description follows the ODD + D protocol for describing socio-ecological agent-based models (Müller *et al.* 2013). All elements of the framework are included and thus also lists when an element does not apply to this specific model. The described model is an adapted version of the ROMFARMS model by Joyce (2019) that utilizes the demographic system from the replicated Artificial Anasazi model by Janssen (2009).

Overview

1.1 Purpose

The purpose of this model is to identify whether the implementation of a different version of the similar ontological components between the Artificial Anasazi model (Janssen 2009) and the ROMFARMS model (Joyce 2019) leads to similar output results. As a result, this model is intended to study the relationship between both models. The conceptual household structures were determined as ontologically similar between the two models via a qualitative comparison. In this model the household structure from the Artificial Anasazi has therefore been implemented in the ROMFARMS environment.

ROMFARMS is a heuristic tool that has originally been developed to generate different land-use scenarios based on the execution of a diversity of arable strategies. It specifically applies to the Lower Rhine Delta in the Netherlands between 15 BC and 270 AD.

1.2 Entities, state variables and scales

The model includes two different agents: The household and the animal. Each timestep represents a year wherein farming, animal husbandry, fuel collection and timber collection activities are being performed. The following two lists discuss the main attributes of the two agent categories.

The main attributes of a household are:

- Age and predefined age of death;
- Fertility-rate
- Number of people (inhabitants);
- Number of animals (livestock);
- An area of arable land;
- Grain storage;
- Labour costs for ploughing, sowing, manuring and harvesting;
- A workforce;
- Calorie consumption;
- Farming and animal husbandry areas;
- A fuel store;
- Construction activities

Each household owns a herd of animals (cattle, sheep and horse). The attributes of this animal agent are:

- Age
- Sex
- Species
- Fertility rate
- Survivorship
- Owner
- Milk producing?

Each patch represents one hectare (100x100 meter) and has the following attributes:

- Raster value
- Owner
- Landscape type (Levee, flood-basin or undefined)
- Biomass
- Restoration countdown after being worked

The total number of patches represents the landscape. This landscape is randomly generated based on the area parameters that can be found in the interface screen.

1.3 Process overview and scheduling

At each timestep and after the initialization, the following sequence of events occurs:
1. Households age with one year, due to their age they might die or reproduce;
2. Households calculate grain requirements; how much arable land is required to foresee in these requirements and cultivate the land;
3. Grain surpluses or deficits are reported, settlements compensate for each other for their deficits;
4. Animal population dynamics, households engage in animal husbandry and collect resources from the animals;
5. Calculation of fuel requirements and collection of biomass to foresee in these requirements;
6. Depending on the reconstruction frequency, households engage in the collection of construction materials;
7. Update patch and agent variables;
8. Repeat the first step.

Design concepts

2.1 Theoretical and Empirical Background

The following section is taken from the ODD+D of ROMFARMS by Weijgertse (2020).
"The main concept behind ROMFARMS is that all settlements pursue a mixed farming strategy: farming, animal husbandry and fuel/timber collection. This idea is based on a combination of literary sources (this applies to farming and animal husbandry) and expert judgment (this applies to fuel and timber collection). This mixed farming strategy creates a complex and holistic agricultural subsistence system and allows for an evaluation of the impact that different agricultural subsistence strategies would have had on the land. On the level of the submodels, use is made of palaeoenvironmental and archaeozoological data, ethnographical parallels, (ancient) literary sources, economic data and (as few as possible) expert judgment. Data aggregation levels vary between the level of the settlement, household and individual for all submodels.

The submodel for population dynamics utilizes calorific/nutritional needs based on data from the United Nations Farming and Agriculture Organization. Grain caloric contents and abundance in the arable farming submodel is based on archaeobotanical evidence. For animal husbandry, slaughter rates are derived from system dynamics models and yields from archaeological experimental research.

Agents are assumed to make economically rational decisions by calculating current presence of resources, needs and available land. This is chosen because the ROMFARMS model aims to approximate a null scenario for different agricultural subsistence strategies, whose complexity can be expanded in the future. "

Despite the fact that no alterations have been made to the origin of all the data, the purpose of this model is to get a better insight in the relationship. Therefore, the population dynamics model has been altered in relation to the original ROMFARMS model. This is done based on a structured qualitative comparison of the Artificial Anasazi and ROMFARMS model.

2.2 Individual Decision-making

The decision-making in this model is regulated via the household, the animal agents are objects in the decision-making process of the household and cannot portray behaviour on their own. Despite being a theory-building tool, meaning that the agent-based model is not built to achieve success, success in this model can generally perceived as the households being able to produce and collect sufficient resources to not achieve resource deficits. Success is also achieved when the total number of households has not disappeared before the final timestep of the simulation. The following relevant description is taken and altered from the ROMFARMS ODD+D:
"~~Settlements~~ Households make economically rational decisions and are driven by mathematical programming. Agents adapt their farming, animal husbandry and fuel/timber strategies based on new calculations after changes in the landscape ~~(last steps of the process overview)~~ , the current composition of their inhabitants (first steps of the process overview) and the agricultural strategy they pursue. Social norms and cultural values are not incorporated in this model. Space plays a significant role in the sense that it determines farming, animal husbandry and fuel/timber collection possibilities for the new timestep. ~~Settlements~~ Households act based on information available at the moment of decision-making/calculation and do not take past events and future predictions into account. Settlements have thus no knowledge on possible outcomes of future events, but this uncertainty does not affect their decisions."

2.3 Learning

Both individual learning and collective learning are not present in this model.

2.4 Individual Sensing

Households only engage in sensing when determining their arable lands and meadows required for arable farming and animal husbandry, or when sensing for fuel sources when collecting fuel. Settlements have an unlimited sensing range and cannot be erroneous in their sensing.

2.5 Individual Prediction

Agents do not engage in prediction based on previous events. Based on parameter settings, agents can however create a buffer of sowing seed for the next year.

2.6 Interaction

No direct interaction between agents takes place. Households will only compensate for the grain deficits of other households when they have a surplus themselves, but this is regulated via mathematical formulae and does have social consequences for the agents.

2.7 Collectives

No collectives are explicitly models, but it is assumed that the total number of animals belonging to one household comprise the total herd of that household.

2.8 Heterogeneity

Both agent classes are heterogeneous. Households differ from each other in terms of age, livestock and yields. Animals differ in terms of age, sex, species and owner. Agents are however homogeneous in their decision-making.

2.9 Stochasticity

Stochasticity is involved in relation to the setup of the environment in the form of a random distribution of landscape patches and households. Stochasticity is furthermore involved in the reproduction capabilities of settlements and the reproduction capabilities and mortalities of animals, and the order in which agents act per timestep.

2.10 Observation

This implemented model is capable of producing the same output measurements as the original ROMFARMs model: composition of settlement demographics, livestock constitution, resource quantities, resource consumption, constitution and efficiency of labour forces and exploited land.

These output results have been compared for equivalence with the same output results from the ROMFARMS model. It was concluded that the implemented model has parameter combinations that provide demographic developments that are statistically indistinguishable from the original ROMFARMS model. Despite being statistically indistinguishable, the output results of the other submodels are not statistically equivalent as well. This is primarily caused by the fact that household populations are static, whereas they are more fluid and diverse in the original model.

Details

3.1 Implementation details

This model has been implemented in NetLogo (version 6.1.0). It is at this moment not accessible online, but the main goal is to upload this model to the Leiden University repository.

3.2 Initialization

The initial state of the model has the variables of the demography submodel set to those values that produced a statistically indistinguishable population development compared to ROMFARMS. Values of other submodels are naturally set to middle values.

3.3 Input Data

This model does not require any external input data.

3.4 Submodels

The model consists of five different submodels:

- Population dynamics: This submodel regulates the development of the deaths and reproduction of settlements. Settlements die when their age is similar to the predefined age of death. Reproduction is regulated via a stochastic mechanism and a fertility value; a settlement has the opportunity to reproduce a maximum of 4 times since the number of assumed subadults per household is 4.
- Arable farming: *"This submodel regulates the cultivation of arable land. The household calculates the required arable land and the available maximum area of arable land. Based on availability, either the maximum area (when the minimum is higher than what is available) or the minimum area will be cultivated. Based on extensive strategy extra sowing seed is used to cultivate more land and based on intensive strategy the cultivated land is manured. The yield is calculated and if a settlement has a deficit, it will borrow grain from a neighbouring settlement. Variables will then be updated."* (ROMFARMS ODD+D)
- Animal husbandry: *"This submodel focuses on the reproduction, natural mortality and slaughter of animals. Adult female animals will produce offspring when a random number is higher than the fertility rate, sheep have an added random probability of producing more than one offspring relative to cattle and horses. Newborn animals have a neonatal mortality rate and adult animals have a natural mortality rate, if a random number is higher than this rate the animal will die. After slaughter rates are set, all male neonates are slaughtered and*

based on the slaughter rates other animals are being slaughtered if a random number is lower than the slaughter rate." (ROMFARMS ODD+D)

- Fuel collection: The household calculates the required amount of biomass that needs to be used as fuel, calculates the best workforce to establish this and then collects all of the biomass itself.
- Timber collection: Identical to the fuel collection submodel, with the side note that it is only executed once the age of the household is the same or a multiplication of the reconstruction frequency.

References

Janssen, M., 2009. Understanding Artificial Anasazi. *Journal of Artificial Societies and Social Simulation* 12(4)13, 1-17.

Joyce, J., 2019. *Farming along the limes: Using agent-based modelling to investigate possibilities for subsistence and surplus-based agricultural production in the Lower Rhine delta between 12BCE and 270CE.* Vrije Universiteit Amsterdam Research Portal.

Weijgertse, S., 2020. *ODD+D of Joyce's (2019) NetLogo implementation of ROMFARMS.* Leiden (unpublished agent-based model description).

Appendix 6: Work log for the implementation phase and data analysis phase

Total estimation of minimum development time for the docked model (Section 1+2 and related non-categorical steps): ± 27:40

Total estimation of minimum time for data analysis (section 3+4 and related non-categorical steps): ± 32:15

1. Design and preparation of ROMFARMS for the implementation

1. Initial exploration of the source code to get insight in its structuring, using the NetLogo dictionary for guidance (3:00)
2. Re-structuring of the source code (00:10)
3. Replacing input boxes with parameter sliders (00:10)
4. Implementing mechanisms that allow all submodels besides demography and farming to be turned off (00:15)

Sub-total: 3:35

2. Code modifications

2.1 Globals and variables

1. Creation of household breed and transferring the agent variables of settlements and person breeds to this breed (00:03)
2. Removal of person and household breeds (00:02)
3. Determining and transferring global variables from Artificial Anasazi model (00:10)
4. Creation of new household variables to ensure functionality with all submodels (00:05)

Sub-total: 00:20

2.2 Setup procedure

1. In-depth exploration of the source code to assess where changes must be made (01:00)

1. Implement stop-mechanism when total area parameters exceed available patches (00:05)
2. Remove source code for loading the unavailable GIS landscape files (00:05)
3. Alteration of global variables to suit the intended household structure (00:10)
4. Import and calibrate Artificial Anasazi global variable values (00:10)
5. Removal of 2-, 3- and 5-household settlement initialization and adapt the 1-household settlement initialization and configuration to suit the intended household structure (00:20)

Sub-total: 1:50

2.3 Demography submodel

1. In-depth exploration of the source code to assess where changes must be made (1:30)
2. Creation of sliders with Artificial Anasazi global variables for population dynamics (00:05)
3. Implement changes to *update-settlements-1*, *deaths* and *marriages* procedures (1:30)
4. Removal of *update-households-2*, *births*, *set-fertility*, *set-mortality* procedures and excessive code in other procedures (00:15)
5. Debugging and testing submodel functionality (1:00)

Sub-total: 4:20

2.4 Arable Farming submodel

1. In-depth exploration of the source code to assess internal functionality and locations where changes must be made (1:00)
2. Adaptation of source code to suit household structure (00:15)
3. Debugging and testing submodel functionality (00:30)

Sub-total: 1:45

2.5 Animal Husbandry submodel

1. In-depth exploration of the source code to assess internal functionality and locations where changes must be made (00:30)
2. Adaptation of source code to suit the household structure (00:15)
3. Debugging and testing submodel functionality (00:10)

Sub-total: 00:55

2.6 Fuel and Timber submodels

1. In-depth exploration of the source code to assess internal functionality and locations where changes must be made (2:00)
2. Implementation of reconstruction multiplier system to allow aging of households (00:30)
3. Adaptation of source code to suit the household structure (1:15)
4. Removal of *calc-workforce-fuel* and *calc-workforce-construction* procedures and excessive source code (00:10)
5. Debugging and testing submodel functionality (1:15)

Sub-total: 5:10

***2.7* Update variables and patches**

1. In-depth exploration of the source code to assess internal functionality and locations where changes must be made (00:30)
2. Implement mechanisms that variables are only being update if the corresponding submodel is activated (00:10)
3. Adaptation of the source code to suit changes made to the submodels (00:20)
4. Debugging and testing of functionality (00:30)

Sub-total: 1:30

3. Running the model

1. Running sensitivity analysis experiments
    - A. Demography submodel (00:20)
    - B. Arable farming submodel (00:10)
    - C. Fuel and timber submodels (00:15)
2. Running experiments for model alignment
    - A. Demography submodel: ROMFARMS (2x), docked model (1x) (00:30)
    - B. Arable farming submodel: ROMFARMS (1x) docked model (1x) (00:15)
    - C. Animal husbandry submodel: ROMFARMS (1x) docked model (1x) (00:15)
    - D. Fuel/timber submodels: ROMFARMS (1x) docked model (1x) (00:15)
3. Processing the CSV data of all experiments for suitable importation to RStudio (1:30)

Sub-total: 3:30

4. Data analysis in RStudio

1. Sensitivity analysis:
    - A. Demography submodel (4:30)
    - B. Arable farming submodel (3:30)
    - C. Fuel and timber submodels (4:00)
2. Model alignment analysis
    - A. Demography submodel: ROMFARMS (2x), docked model (1x) (5:00)
    - B. Arable farming submodel: ROMFARMS (1x) docked model (1x) (2:30)
    - C. Animal husbandry submodel: ROMFARMS (1x) docked model (1x) (1:30)
    - D. Fuel/timber submodels: ROMFARMS (1x) docked model (1x) (2:45)

Sub-total: 23:45

5. Non-categorical steps

1. Total consultation of the NetLogo Dictionary and Programming Guide (5:00)
2. Consultation of R and StackOverflow websites for data analysis (5:00)
3. Consultation of other online sources (2:00)
4. Consultation of literary sources related to the case studies (00:45)
5. Meetings with thesis supervisors focused on the implementation phase (1:30)