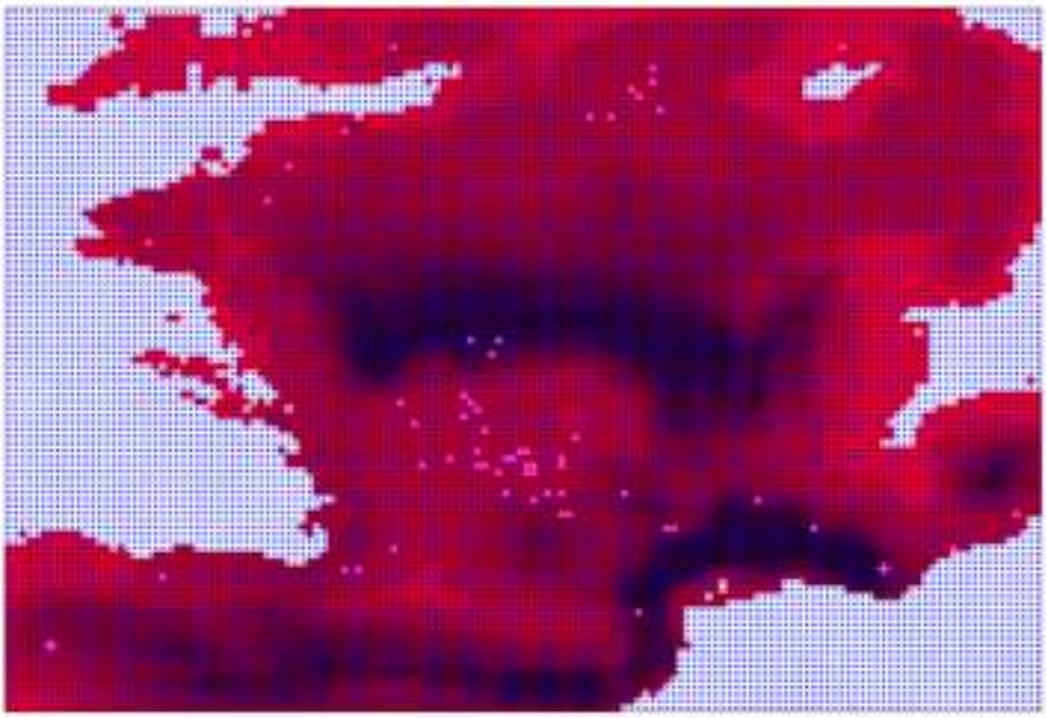
HomininSpace

Version 1.5

Overview, Design concepts and Details (ODD) of Model and implemented Simulation system



Universiteit Leiden

Fulco Scherjon

Faculty of Archaeology

f.scherjon@arch.leidenuniv.nl

Contents

Figures

Tables

Picture on the front depicts a population density map through time for the research area, created with the HomininSpace simulation system.

1. MODEL DESCRIPTION FOLLOWING THE OVERVIEW, DESIGN CONCEPTS AND DETAILS (ODD) PROTOCOL

This document provides the design of the model and simulation system for the research project: "HomininSpace 1.5, modelling and simulating hominins in a realistic world" (HS) according to the Overview, Design concepts and Details protocol (ODD). The ODD was introduced in an effort to standardize the publication and documentation of Agent Based Models (ABMs) and to allow duplication of simulation runs and replication of simulation results (Antón *et al.* 2014; Grimm *et al.* 2006; Müller *et al.* 2014). The open access repository for agent-based models at www.openabm.org adopted the ODD protocol as a requirement for their Model Review process[1]. The aim is to document the HomininSpace model in such a manner that it will pass this review process, fulfills the quality requirements and ultimately allowing incorporation of HS into the repository. This chapter documents the ODD protocol followed for HS. The protocol prescribes several elements including purpose of the model, process overview, design concepts, scheduling, state variables and input and output descriptions (see Table 1). Format and description of the model will facilitate re-creation of the research that is done within the HS simulation environment (Müller *et al.* 2014).

Table 1 - Overview of the elements within the ODD protocol. These elements have to be used in exactly the order specified here when implementing the protocol (Grimm *et al.* 2006).

Category	*Elements of the ODD protocol*
Overview	1. Purpose
	2. Entities, state variables, and scales
	3. Process overview and scheduling
Design concepts	4. Design concepts
	• Basic principles
	• Emergence
	• Adaptation
	• Objectives
	• Learning
	• Prediction
	• Sensing
	• Interaction
	• Stochasticity
	• Collectives
	• Observation
Details	5. Initialization
	6. Input data
	7. Submodels

[1] https://www.openabm.org/page/model-review-overview, accessed 27 October 2014.

In the following paragraphs each of these elements is described. To facilitate interpretation all elements are preceded by the relevant part of the original explanation according to Grimm *et al*. (2014), Supplement 2[2]. Extensive discussion of these elements can be found in Railsback (2001) and Grimm *et al*. (2005, Chapter 5). Additionally some elements of the ODD + Decision protocol (ODD+D) (Amiel *et al*. 2011) have been included in this document since the HomininSpace model includes a representation of human decision making, for which the ODD+D extension was designed. However, the guiding structure of the ODD protocol is kept intact, and the ODD+D elements are only referred to in the explanations and for clarity reasons there presented in bold face.

1.1 (1) Purpose

What is the purpose of the model? ***For whom is the model designed? How has the model been implemented? Is the model accessible, and if so where?***

The purpose of the model underlying the HS simulation system is to simulate hominin population dynamics in a realistic landscape under varying environmental conditions representative of those encountered in the past for the simulation period. Simulation results are used to explore how different parameters influence dispersal characteristics of the population. Individual groups are steered through the landscape following available energy in the form of edible ungulates. The model validity is assessed by comparing simulated population distributions through time versus archaeological presence and absence data for given points in space and time (dated archaeological sites). A better match indicates a 'better' model, one that is more likely to represent hominin behavior in the past. The results are used to evaluate the effect of different parameters and to find optimal parameter configurations. The implemented tool is an extension of the repertoire for researchers interested in past complex societies and is intended to be used in an explorative manner. The model is implemented in Repast Simphony 2.2 and code is written in the Java module. The model is available upon request from Fulco.scherjon@arch.leidenuniv.nl.

1.2 (2) Entities, state variables, and scales

What kinds of entities are in the model? By what state variables, or attributes, are these entities characterized? What are the temporal and spatial resolutions and extends of the model?

The ABM for HS comprises the following entities:

1. The landscape: the grid based environment in which the other agents reside is a collection of objects containing habitat information. A grid cell is a square area of 10x10 kilometers.
2. The agents:

[2] Available from the Railsback and Grimm 2014 book, online downloads: http://www.railsback-grimm-abm-book.com/Chapter03/GrimmEtAl2010_App2_ODD-template.docx, accessed 28 October 2014.

a. Moving through the landscape grid are groups of hominin (HomininGroups). Groups need resources which they obtain from the landscape. Groups can procreate and die.
b. The presence of hominin groups is detected by Checkpoints in Space and Time (CSTs). They are located in the grid representing archaeological sites and cannot move.
c. Refugia that can create new groups are implemented as hominin Factories (HFs) with a fixed location on the grid.

New groups can be created by procreation from other groups, by HFs and at initial locations at startup. Newly created groups are not able to interact with other groups for five turns after creation. There are three types of hominins: (1) the population that was already present at the start of a simulation, (2) those that were created by the Iberian HF and (3) those coming from the Italian HF. Offspring has the same type as the parent group.

Actions for the HomininGroup objects are initiated from the main loop. The HomininGroups scan the area around them for the best patch to move to, move to this patch and adapt their foraging range according to their needs and the environment. There is a clear tendency for grid cells to become more attractive in the underlying model when going south and west. This is due to the temperature distribution which favors locations more to the west and naturally south. To prevent all groups ending up in Iberia (South-west corner of the map), comparing two grid cells takes 10% error margin in account. That is, for an area to be more attractive it needs to be at least 10% better than other areas. When all groups have moved, the groups in random order subtract energy in the form of food from cells in the foraging range. Upon extraction the groups register their presence with checkpoints in the grid. They continue foraging until their subsistence needs are met or there is no more energy left. As such, not all energy from all the grid cells is necessarily consumed in each turn and sometimes not enough energy is available (especially when foraging ranges are overlapping).

1.2.1 HomininFactory class (implementation of refugia)

This agent produces new groups of hominins. Factories are positioned in the initialization phase only and cannot move. Each factory is called once every turn to check for production and can produce one new group. The Iberian peninsula is one of the areas within the HomininSpace model from where new populations can move into North-western Europe. The second source of possible population influx is the Italian peninsula, with hominins moving along the Mediterranean coastline passing the Alpine mountain ridges. Factories will check local conditions before production and will make sure that:

- there are no hungry groups within scanning range (default 30 cells);
- there are empty cells where the new group can be located;

When these conditions are met a new group is created at a random empty grid cell in the area. New groups immediately start moving and consuming, possibly in competition with the parent group or other groups in the area. Refugia can accept groups when the option ‘Open borders’ is selected (see section 1.6.5). With this option active, groups can pass the borders of the simulation area and are removed from the simulation and considered to have been absorbed into refugia.

1.2.2 Checkpoint class (sites)

For HomininSpace a chrono-archaeological database has been built to store Checkpoints in Space and Time (CSTs) (Scherjon 2012). These Checkpoints represent archaeological sites. Each CST has a list of dating results (intervals) that indicate when hominins were at that locality. A hominin group is assumed to be present when energy is extracted from the grid cell with the CST. Each defined interval is inclusive on the edges: from *a*-*b* means *a*, everything between *a* and *b*, and *b*. Visits are administrated per hominin group type.

Next to the regular checkpoints three other types have been defined: monitoring checkpoints that have no presence list but store visit results anyway, starting locations, sites that the user can list as input for the simulation where hominin groups are already present when the simulation starts, and climate monitoring points which store reconstructed precipitation and temperature values. Those checkpoints that are included in the simulations are mapped in Google Earth (GE) and visualized in Figure 1.

Figure 1 - Checkpoints in HomininSpace. In green regular checkpoints, in yellow monitoring points and in blue/red the starting locations.

1.2.3 System objects

The classes described here are used for administrative purposes, and are not part of the model±

- `Constants` provides the other classes with constant values for specific variables and with functionality to print debug information.
- `HSUtils` implements functionality to read input files and to administer dead groups.
- `HomininSpaceContextBuilder` class is the framework of the system, including initialization, main loop and finalization phases. Upon initialization the grid is constructed, all configuration files are read, and factories and checkpoints are created. Just before the first time step the *start*() routine is called, initializing all habitat cells and creating any initially present hominins. At the end of the simulation the *end*() routine is called, cleaning memory and finalizing output. This routine also collects statistics from the checkpoints and logs the results of the simulation.

1.3 (3) Process overview and scheduling

Who (i.e., what entity) does what, and in what order? When are state variables updated? How is time modeled, as discrete steps or as a continuum over which both continuous processes and discrete events can occur? Except for very simple schedules, one should use pseudo-code to describe the schedule in every detail, so that the model can be re-implemented from this code. Ideally, the pseudo-code corresponds fully to the actual code used in the program implementing the ABM.

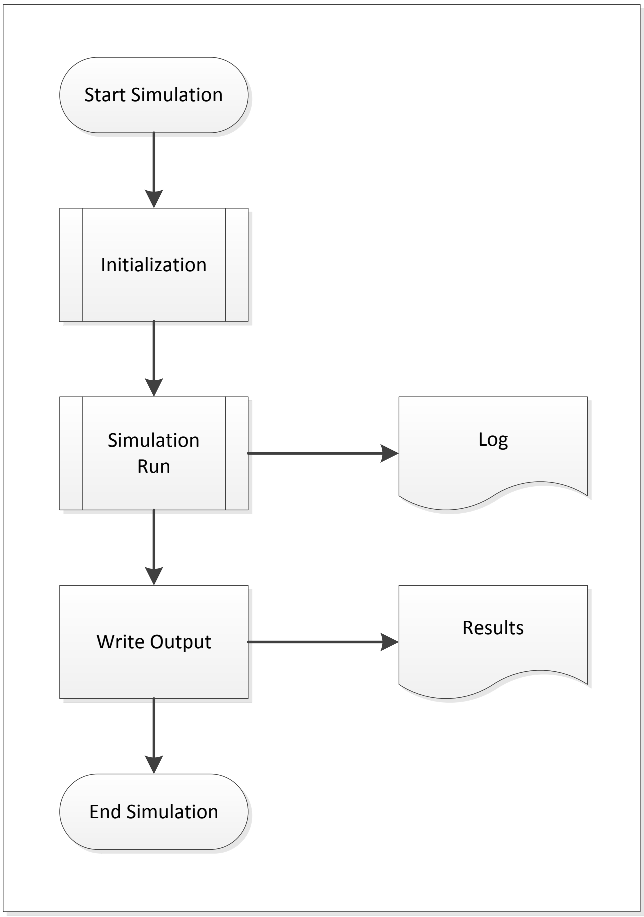


Figure 2 - Flowchart: overview of the simulation process.

An overview of the simulation process is given in Figure 2. At the start of a simulation the system is initialized by reading the configuration and data files (see . The configuration is a simulation number, a random seed and a set of parameter values. Data files include grid data, climate data for modern day and the past environments, and archaeological data. The simulation grid is created and filled with the environment information. Factories are created. Data for the reconstructed global sea level and the yearly mean temperature are read from file and stored in memory.

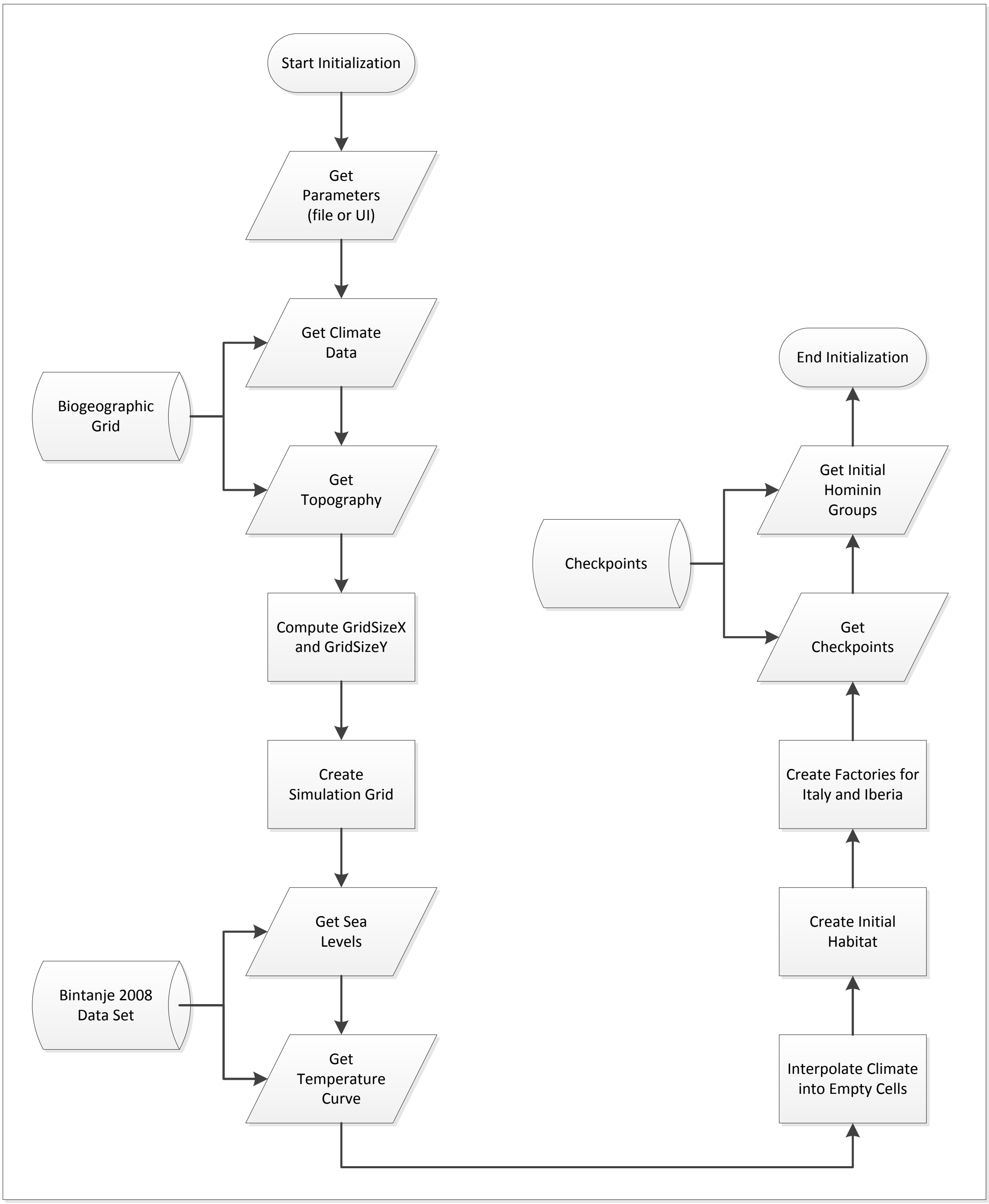


Figure 3 - Flowchart: Initialization routine

After initialization of the system the main loop is started. The flow chart depicting the main routine is in Figure 4. For each time step, the environment is updated with new climatic data. Next each hominin group that can move will scan the environment delimited by two times the foraging range. The location which is most attractive to this group becomes the destination for this group. Upon moving to this new location the foraging range for this group is recalculated, and presence is attested and resources claimed. The groups will then consume some of the claimed resources and the hominins in the group will either produce offspring or they will die, depending on the availability of resources. Logging information is written at the end of each timestep. When the simulation finishes, additional administrative data is written to output files and internal buffers (for charts and time series).

The main loop sequentially executes the following actions:

1. Increment timer, get new global sea level and global mean temperature;
2. Update all habitat cells with these new values and recalculate available energy;
3. Move all hominin groups and update foraging ranges;
4. For all groups, consume required energy, call the grow routine, check for merge possibilities, perform reproduction and then check for death conditions;
5. Finally check production conditions for each factory and if favorable, produce one new group.

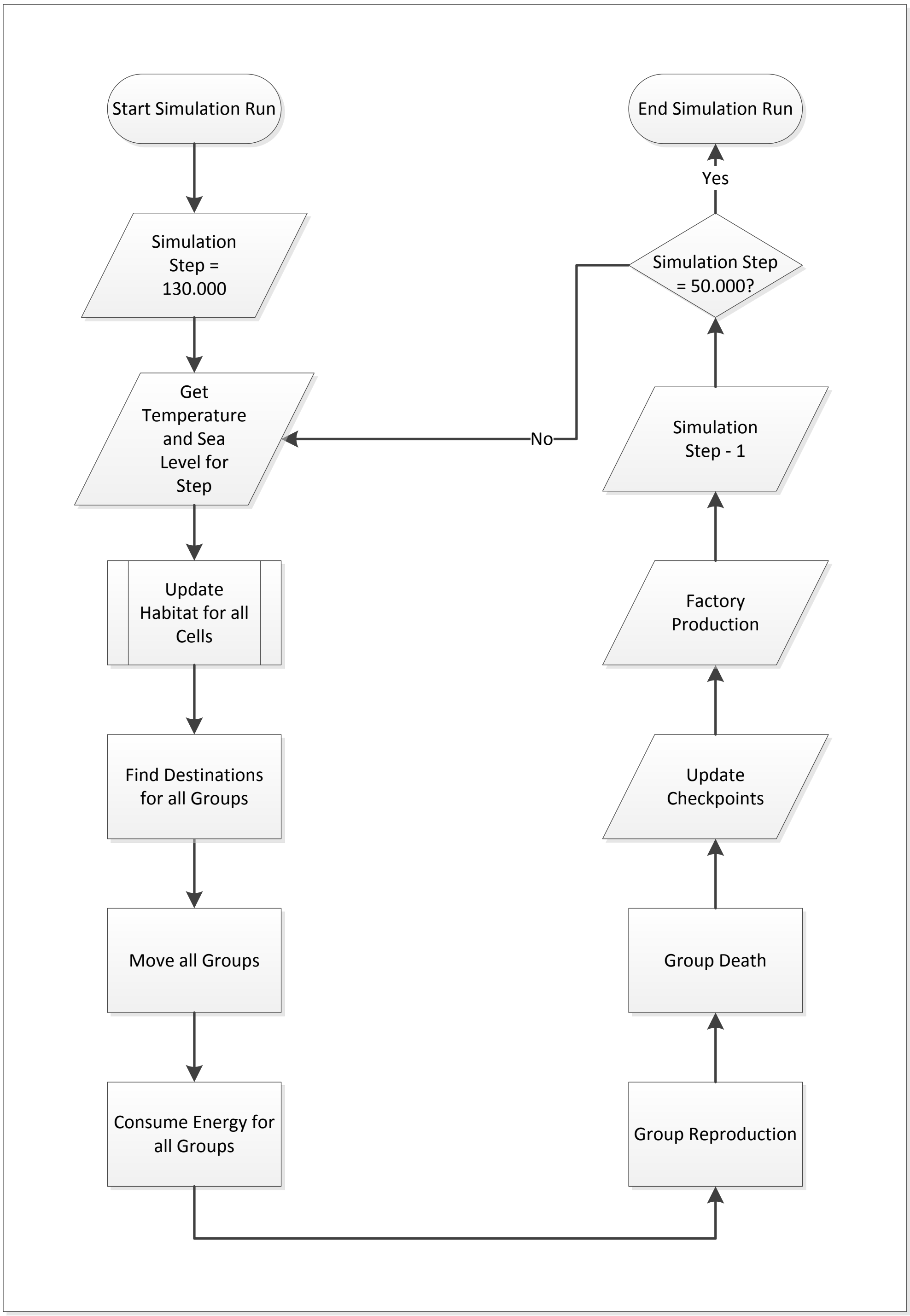


Figure 4 - Flowchart: Simulation loop.

1.3.1 Simulated time and period

Each simulation time step is defined to be one year. Each simulation has a starting time expressed in years before present (BP) and a number of time steps (years) to execute. All simulations start at 130 BP and end at 50 kya. That means that every simulation run for 81.000 time steps.

1.4 (4) Design concepts

1.4.1 Basic principles

*Which **assumptions**, general concepts, theories, hypotheses, or modeling approaches are underlying the model's design? Explain the relationship between these basic principles, the complexity expanded in this model, and the purpose of the study. How were they taken into account? **What is the Theoretical and Empirical Background***

The underlying assumptions and theoretical background to the core concepts are discussed in the respective submodels in section 1.7.

1.4.2 Random values

One of the aims for HomininSpace is to minimize the influence of random events and to create a deterministic simulation system. One place in the program where the influence of randomness is still present is in the movement and consumption order for the hominin groups. Each time step the complete list of groups is randomly ordered and then traversed sequentially with each group moving to its preferred area. Other locations where random values play a role is where a random number between -1 and 1 is added to the energy level of each HabitatCell to prevent a biased choice between equal destinations due to preferential system list ordering (this prevents groups always selecting for instance the left path when presented equal route choices), and when locations for new groups are randomly chosen. The random seed (a nine or ten digit number) driving the random generator is included as an identifying parameter for simulations.

1.4.3 Emergence

What key results or outputs of the model are modeled as emerging from the adaptive traits, or behaviors, of individuals? In other words, what model results are expected to vary in complex and perhaps unpredictable ways when particular characteristics of individuals or their environment change? Are there other results that are more tightly imposed by model rules and hence less dependent on what individuals do, and hence 'built in' rather than emergent results?

The key model result that is expected to change is the match with the archaeology (the checkpoints). Due to changes in the environment groups will re-distribute across the landscape, possibly matching the patterns of presence and absence. Responses to environmental change are governed by the parameter settings that define the reactions of individual agents. Population distribution and matching patterns then emerge through time.

1.4.4 Adaptation

What adaptive traits do the individuals have? What rules do they have for making decisions or changing behavior in response to changes in themselves or their environment? Do these traits explicitly seek to increase some measure of individual success regarding its objectives (e.g., "move to the cell providing fastest growth rate", where growth is assumed to be an indicator of success; see the next concept)? Or do they instead simply cause individuals to reproduce observed behaviors (e.g., "go uphill 70% of the time") that are implicitly *assumed to indirectly convey success or fitness?* ***Are the agents heterogeneous? If yes, which state variables and/or processes differ between the agents?***

At the beginning of each year the foraging ranges for all groups are computed. These are based on the subsistence needs of the members of the group, the available resources in the area, and the presence and subsistence needs of other groups. To calculate the foraging range, first the number of individuals in each age segment is multiplied by the energy needs for members of that segment (given as model parameters, see section 1.6.5). Then the available energy (the number of ungulates) from the home range (foraging range zero) is obtained. If this is less than the resources needed by the group the foraging range is increased by one grid cell. The amount of energy in each cell within the foraging range is added to the available energy for the foraging area. For a foraging range of one this is a maximum of eight cells that can provide energy for the group, additional to the home range. If this is not enough to satisfy the needs of the group the foraging range is again increased and this process is repeated until the energy requirements are theoretically met. In the consumption phase resource competition might impose resource insufficiencies within closely located groups. All grid cells within the foraging range receive a claim by this group. All grid cells that have received more than one claim can provide a fraction of the energy per group, based on the number of claims. The foraging range is calculated for all groups in random order.

Subsequently all groups consume the energy they need, in random order. If there is not enough energy (for instance because another group has consumed part of the claimed resources) a group is considered hungry. Hungry groups have penalties on birth rates and suffer from increased death rates. Groups are allowed to move through other foraging ranges in search for food. The most likely situation in which a group A is suddenly deprived of resources is when a mobile group B moves into the foraging range *after* the computation of the foraging range for A.

1.4.5 Objectives

If adaptive traits explicitly act to increase some measure of the individual's success at meeting some objective, what exactly is that objective and how is it measured? When individuals make decisions by ranking alternatives, what criteria do they use?

The main objective of each group is to obtain at least that amount of kilocalories that is needed for sustained existence and growth. When groups can move they will move to the most productive area within their scanning range, which is two times the foraging range. When groups are large enough they will procreate, constructing new groups randomly located in their foraging range.

1.4.6 Learning

Many individuals or agents (but also organizations and institutions) change their adaptive traits over time as a consequence of their experience? If so, how? ***Is collective learning implemented in the model?***

Individual agents do <u>not</u> learn from their experiences. Traits are not changed in individuals but between simulation runs.

1.4.7 Prediction

Prediction is fundamental to successful decision-making; if an agent's adaptive traits or learning procedures are based on estimating future consequences of decisions, how do agents predict the future conditions (either environmental or internal) they will experience?

Moving groups will go to those areas within scanning range that offer them the highest carrying capacity.

1.4.8 Sensing

What internal and environmental state variables are individuals assumed to sense and consider in their decisions? What state variables of which other individuals and entities can an individual perceive; for example, signals that another individual may intentionally or unintentionally send?

Groups will not move onto water grid cells. Groups can read the carrying capacity of each grid cell within scanning range (two times the foraging range) and the number of claims on that cell. Groups cannot move onto a grid cell that contains another group. Small groups (< 3 members) can send a merge request to other groups in their scanning range. Groups can register at checkpoints. Factories sense groups in the scanning range and can ask if a group is hungry.

1.4.9 Interaction

What kinds of interactions among agents are assumed? Are there direct interactions in which individuals encounter and affect others, or are interactions indirect, e.g., via competition for a mediating resource? If the interactions involve communication, how are such communications represented?

Hominin groups generally do not communicate with each other. Interactions are generally through sharing resources from neighboring grid cells. Direct communication between hominin groups happens only when a merge request is transmitted from one group to another group. There is no communication between Hominin Factories and other agents, but HFs can check if HGs are hungry. HGs register their presence with CSTs if they forage in grid cells where a Checkpoint is located.

1.4.10 Stochasticity

What processes are modeled by assuming they are random or partly random? Is stochasticity used, for example, to reproduce variability in processes for which it is unimportant to model the actual causes of the variability? Is it used to cause model events or behaviors to occur with a specified frequency?

Reproduction and mortality is implemented as stochastic processes. Since the basic agent is a group of hominins, reproduction occurs within that group, not between two individual hominins. Half of the group is assumed to be female, and each female can reproduce once every three years (this is a parameter value). Mortality rates are applied against every age class at the end of each year.

1.4.11 Collectives

Do the individuals form or belong to aggregations that affect, and are affected by, the individuals? Such collectives can be an important intermediate level of organization in an ABM; examples include social groups, fish schools and bird flocks, and human networks and organizations. How are collectives represented?

At the moment no collectives are present in HomininSpaces. Groups can be of different cultural type, but that does not affect the individual group in any way. It is only used for administrative purposes.

1.4.12 Observation

What data are collected from the ABM for testing, understanding, and analyzing it, and how and when are they collected? Are all output data freely used, or are only certain data sampled and used?

Next to the optional output data formats as described in the next paragraph, each simulation produces a results file with statistical information, used to determine the match with the archaeological data. This file contains the following information: the input parameters and defining constants for this model run, simulation totals (#groups, #hominins, #visits, #matching visits, #homerange matching visits, #homerange visits, #matching visits with confidence), totals per checkpoint including monitoring checkpoints, and produced new groups by each refugium and as offspring.

1.4.13 Output data: optional visualization maps, charts and time series

Output of a simulation run consists of a log file with debug information, a results file with statistics, and simulation data collected in maps, time series and charts. It is possible to take snapshots of the visual grid or create a movie from the simulation.

Visualization maps are used to show information on the simulation grid itself. The following can be displayed: topography (height levels and a varying sea level), reconstructed environment at any point in time (two variants: energy level and habitat reconstruction), visit density (count of all visits in the simulation this far), and death density (count of all groups that died this far).

Time series are graphs that display statistical information per timestep while running the simulation. The following time series are implemented: Global Temperature, Global Sea Level, X and Y Position statistics, Hominin Count, Hominin Type Count, Group Count, Hungry Group Percentage, Average Foraging Range, and Average Group Size.

Charts are graphical representations of some data sets in the form of bars. They present data for only one point in time, updated at each time step during a simulation. Two charts are available: Group Ages and Group Sizes.

1.5 (5) Initialization

What is the initial state of the model world, i.e., at time t = 0 of a simulation run? In detail, how many entities of what type are there initially, and what are the exact values of their state variables (or how were they set stochastically)? Is initialization always the same, or is it allowed to vary among simulations? Are the initial values chosen arbitrarily or based on data? References to those data should be provided.

All simulations start at the same point in time for comparative purposes. Parameters that define the model to be simulated are read, and initial groups are created from the list of starting points.

1.6 (6) Input data

Does the model use input from external sources such as data files or other models to represent processes that change over time?

At initialization the following datasets are read from file: geographical data, climate data for the simulation area for current and LGM conditions, mean temperature and sea levels for the last two million year, and archaeological data (CSTs).

1.6.1 Geography

The area used in the modelling system of HomininSpace is delimited in the North by latitude 51.5, in the South by 41.3, to the West the limit is longitude -6.3 and to the East the edge is defined at longitude 8.5 (see Figure 5). A grid cell is 0.1 x 0.1 degrees (roughly 10 x 10 kilometers) and the total area is composed of 14.948 grid cells. Topographical and climate parameters are distributed onto these grid cells. Altitude information and 19 bio values are derived from data from datasets provided by http://www.worldclim.org (accessed 16 April 2013). HomininSpace uses the 2.5 arc minutes distribution. For the sea grid cells bedrock data from the ETOPO1 dataset is used, from the National Geophysical Data Center (NOAA) (http://maps.ngdc.noaa.gov/viewers/wcs-client/, accessed 13 April 2013). Data is provided in one input file with the climate data (see paragraph 1.6.3).

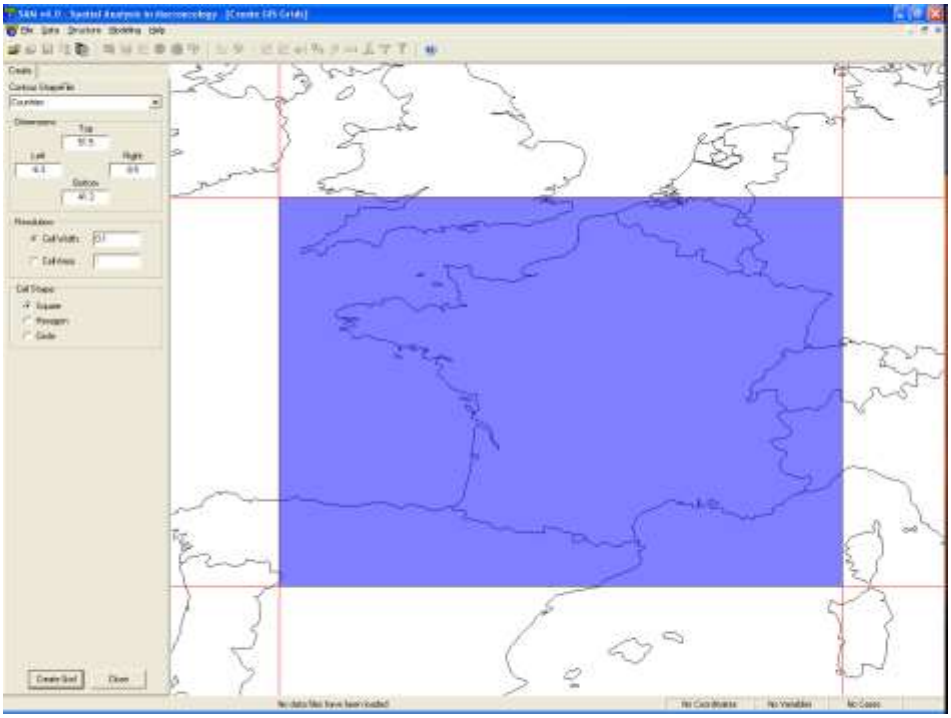


Figure 5 - The simulation grid with 14.948 grid cells.

1.6.2 Global mean temperature and reconstructed global sea level

Temperature and sea level are provided in the input file, taken from the published research by Bintanja and van de Wal (2008). The filename is "*Bintanja2008.txt*" and it is stored in the HomininSpace data directory. A description of the file is given in Table 2. From this file timestamp, surface temperature (interpreted as global mean temperature) and global sea level are used. When there is no value for any given year, the nearest value is returned. The values for the global mean temperature are used as an index for the interpolation between modern day climate data and climate data from the LGM.

Table 2 - Structure of the global temperature and sea level input file.

Field	***Description***
1	Time (kyr BP)
2	Modeled marine oxygen isotope value, relative to present (o/oo)
3	Ice sheet contribution to the marine isotope signal, relative to present (o/oo)
4	Deep-ocean temperature contribution to the marine isotope signal, relative to present (o/oo)
5	Atmospheric surface air temperature relative to present (degC)
6	Deep-ocean temperature relative to present (degC)
7	Eurasian ice volume relative to present (m sea level equivalent)
8	North American ice volume relative to present (m sea level equivalent)
9	Global sea level relative to present (m)

1.6.3 Modern day climate data and past climate data

The input file that contains the climate data sets is called "*HomininSpace Climate Grid.txt*" and is prepared with the SAM tool. In each grid cell the datapoints for each parameter are collected, and made available as aggregrate values. For most variables HomininSpace uses the mean value, only for bedrock the maximum value is used. It contains one line for comments and 14948 data lines. Each line contains the values for one grid cell. The structure of the file is given in Table 3, with an explanation of each field.

Table 3 - Structure of the climate data file. Note: the column Description is not part of the input file.

Field	***Fieldname***	***Type***	***Description***
1	Longitude (X Centroid)	Float	Logitude
2	Latitude (Y Centroid)	Float	Latitude
3	n_alt	Integer	Number of altitude data points
4	Mean_alt	Float	Mean of the altitude points
5	Mean_bio1	Float	Annual mean temperature, current
6	Mean_bio5	Float	Max temp of warmest month, current
7	Mean_bio6	Float	Min temp of coldest month, current
8	Mean_bio12	Float	Annual precipitation, current
9	Mean_bio13	Float	Precipitation of wettest month, current
10	Mean_bio14	Float	Precipitation of driest month, current
11	Mean_wc_2_5m_CCSM_21k_bio_14	Float	Precipitation of driest month, LGM
12	Mean_wc_2_5m_CCSM_21k_bio_1	Float	Annual mean temperature, LGM
13	Mean_wc_2_5m_CCSM_21k_bio_5	Float	Max temp of warmest month, LGM
14	Mean_wc_2_5m_CCSM_21k_bio_6	Float	Min temp of coldest month, LGM
15	Mean_wc_2_5m_CCSM_21k_bio_12	Float	Annual precipitation, LGM
16	Mean_wc_2_5m_CCSM_21k_bio_13	Float	Precipitation of wettest month, LGM
17	Max_etopo1_bedrock	Integer	Bedrock height for current water cells

1.6.4 Checkpoint file (archaeological sites)

Archaeological input data for the simulation are initial starting locations and the Checkpoints in Space and Time (CSTs). Both are stored in one file named: "*Neandertal sites - north west Europe.txt*". This file is read after the grid is initialized. More than one checkpoint can be in the same topographical location. Each checkpoint will keep a separate presence record.

The checkpoint file starts with one line with comments, followed by any number of checkpoints, one per line. Fields are separated by tabs. The layout of this input file is given in Table 4. There are four types of checkpoints indicated by the value of the first field in each line. A value of 1 indicates a regular checkpoint, a value of 2 is a monitoring checkpoint, a value of 3 is a starting location for local groups that are present when the simulation starts, and a 4 indicates a climate monitoring point. For types 2, 3, and 4 only latitude and longitude (and an optional name) has to be provided, for regular checkpoints all other fields must have a value as well. A checkpoint can have more than one interval, and each interval is followed by a confidence indicator (used for output statistics).

Table 4 - Structure of the checkpoint input file. Note: the last column is for explanation only.

Field	*Name*	*Value type*	*Description / remarks*
1	Type checkpoint	Integer	1 = regular checkpoint 2 = monitoring checkpoint 3 = starting location 4 = climate monitoring point
2	Locality (site name)	String	Any characters (but no tabs)
3	Latitude	Floating point	If (latitude, longitude) falls outside the grid, it is ignored.
4	Longitude	Floating point	
5	Age (kya)	Floating point	Interval Example: Artenac 110 +9 -7
6	Sigma plus (ka)	Integer	
7	Sigma minus (ka)	Integer	
8	Confidence level	Integer	1 = unreliable 2 = modest confidence 3 = high confidence in measurement
			Repeat intervals

1.6.5 Parameters and constants

The user can parameterize the underlying model through changing simulation parameters in the configuration file and by changing certain constants in the source code. From the user interface the simulation number (an index into the configuration file) and the random seed can be set. It is also possible to run simulations in batch mode where a subset of simulations is executed.

Parameters for the simulation are stored in a file within the HomininSpace model directory and is named: `HomininSpace Scenario Settings.txt`. The file contains a header with the parameter names and then per line a different set of values. The parameters are detailed in Table 5, with defaults from a standard scenario. The names in source are: `ScenarioNumber`, `RandomSeed`, `BirthRate`, `DeathRate_PreFertileCohort`, `DeathRate_FertileCohort`, `DeathRate_PostFertileCohort`, `Subsistence_PreFertileCohort`, `Subsistence_FertileCohort`, `Subsistence_PostFertileCohort`, `Years_Before_Group_Maturity`, `GroupSize_BeforeMerge`, `GroupSizeFertile_BeforeMerge`, `GroupSize_BeforeSplit`, `Temperature_Tolerance`, `CohortSize_PreFertile`, `CohortSize_Fertile`, `Calories_Per_Kg_Meat`, `Max_ForagingRange`, `Parent_Scenario_1`, `Parent_Scenario_2`. The random seed indexes the random number generator

Table 5 - The parameters and default values for the modelled hominins in HomininSpace.

#	**Parameter name**	**Description**	**Default**
	Demographics		value
1	Birthrate	The number of females that conceive	33

		this year (as a percentage of the total)	
 2 **3** **4**	Deathrate per cohort - pre fertile - fertile - post fertile	The percentage of individuals that do not survive to the next year (percentages per cohort)	 4 3 8
 5 **6**	Cohort size - pre fertile - fertile	Size of the pre-fertile and fertile cohorts (age boundaries).	 13 30
	Energetics		
7	Calories per kg meat	The number of calories that is extracted from a kilogram of meat.	3000
 8 **9** **10**	Subsistence per cohort - pre fertile - fertile - post fertile	The required number of calories for any individual from each cohort. If less is available than the total amount required, the group becomes hungry.	 3000 4000 3500
11	Temperature tolerance	Minimum temperature in degrees Celsius that the group can survive.	-18
	Group dynamics		
12	Years before group interaction	The first *n* years a group cannot interact (merge, settle) with other groups.	5

The constants that can be used to modify the model in the source code with their default values are:

```
INITIAL_SIZE_PREFERTILE          = 8.0;
INITIAL_SIZE_FERTILE             = 11.0;
INITIAL_SIZE_POSTFERTILE         = 6.0;
SEA_LEVEL_HEIGHT                 = 0.0;
GROUP_SIZE_SMALL                 = 10;
FERTILE_SIZE_SMALL               = 3;
YEARS_BEFORE_GROUP_MATURITY      = 5;
```

1.7 (7) Submodels

What, in detail, are the submodels that represent the processes listed in 'Process overview and scheduling'? What are the model parameters, their dimensions, and reference values? How were submodels designed or chosen, and how were they parameterized and then tested?

1.7.1 The environment

HomininSpace uses the two most important climatic parameters, temperature and precipitation, as driving forces for the reconstructed climate in the simulations. The aim of the calculations is an amount of energy available in the form of edible secondary biomass (large ungulates) and this is stored in each grid cell and is used to calculate the carrying capacity per cell (CC). Reconstruction of the available secondary biomass has been implemented in two variants: (1) the continuous energy model, and (2) the habitat reconstruction model. Both begin with computing PP and end with calculating SB.

1. Primary Productivity (PP) is based on mean yearly temperature (T) and precipitation (P) using formulas designed by Lieth (1973). For each grid cell in each turn (year) both values are computed and the smaller of the two is used.

$$\mathbf{PP = \frac{3000}{(1+e^{1.315-0.119*T})} \ in \ \frac{\frac{g}{m^2}}{year}} \qquad \textbf{(Equation 1)}$$

$$\mathbf{PP = \ 3000 * (1 - e^{-0.000664*P}) \ \ in \ \frac{g}{m^2}/year} \quad \textbf{(Equation 2)}$$

2. In the *continuous energy* variant the Secondary Biomass (SB) is computed based on empirical data. This data was collected by McNaughton *et al.* (1989) and presents values for SB that are found in a wide variety of habitats for measured PP levels. A linear regression formula has been derived and is implemented in HomininSpace:

$$\log SB = 1.52 * (\log PP) - 4.79 \ \ in \ kJm^{-2} \quad \textbf{(Equation 3)}$$

3. In the *habitat reconstruction* model the type of climate is determined using temperature ranges based on data collected by Binford (2001). For each cell Effective Temperature (ET) is computed using the mean temperature for the warmest month (MWM) and the mean temperature for the coldest month (MCM) (Binford 2001,59):

$$ET = \frac{(18 * MWM) - (10 * MCM)}{(MWM - MCM + 8)} \ \ in \ °C \qquad \textbf{(Equation 4)}$$

   The ET value specifies which kind of climate (polar, boreal, cool temperate, etc.) is realized. Boundary ET values for different types of climate are given in the table below.

Table 6 - Types of climate and ET boundaries, derived from Table 4.04, page 70 (Binford 2001).

#	Type climate	Min ET (°C)	Max ET (°C)
1	Polar	----	9.99
2	Boreal	10.00	12.49
3	Cool temperate	12.50	14.55
4	Warm temperate	14.56	16.61
5	Subtropical	16.62	18.15
6	Tropical	18.16	22.57
7	Equatorial	22.58	----

   Then the type of vegetation is considered per environment. For several types of environment more than one type of vegetation can be present, depending on the level of precipitation. Primary Biomass (PB) can be calculated for the chosen type of vegetation using the PP mass computed in step 1, with formulas based on research by Kelly (1983) including the mean annual precipitation. The vegetation community that is the transition between deciduous forest and more water-stressed environments is the tall grass prairie-forest steppe, with annual rainfall 717 +/- 257 mm (Binford 2001,Table 4.08, 98). For cooler temperatures Binford identifies mid-latitude short grass prairie (97). Rainfall for this vegetation type is 431 +/- 178 mm, with an upper value of 609mm. In the absence of similar information for other vegetation transitions, this latter value is rather arbitrarily chosen as the default value for precipitation levels for arid, cold environments. Although both the climate types qualify as arid, following Kelly (1983, 284) the difference between the biome types is defined as 900mm rainfall per year. Coe *et al.* (1976) use 700mm rainfall per year as an indication for semi-arid areas.

The amount of Secondary Biomass (SB) that can be sustained by the computed Primary Biomass (PB) can be calculated and is based on empirical key values from Kelly (1983), the reconstructed vegetation type and the calculated PB. From PP for a given biome it is possible to calculate (PB), using the empirical derived ratios as given by Kelly (1983, 284) (c.f. Keeley 1988,379), see Table 7. PB is the total amount of standing plant material present in an area at any time. The PB can be used to calculate the SB with the given ratios.

Table 7 - Calculating PB and SB from PP. Values derived from (Kelly 1983,284, Table 3). Type climate from Table 6 matched against biome type.

Biome#	Type climate (Binford 2001)	Type biome (Kelly 1983)	a = PP / PB	b = SB / PB
1	Polar	Tundra	0.2333	0.0006667
2	Boreal	Boreal forest	0.0400	0.0002500
3	Cool temperate	Temperate Evergreen forest	0.0371	0.0002857
4		Temperate Grassland	0.3800	0.0043750
5	Warm temperate	Temperate Deciduous Forest	0.0400	0.0005333
6		Woodland/Scrubland	0.1200	0.0008333
7	Subtropical	Tropical Savanna	0.2250	0.0037500
8	Tropical	Tropical Seasonal Forest	0.0457	0.0003429
9	Equatorial	Tropical Rain Forest	0.0489	0.0004222

To calculate the PB and SB values for each biome the formulas in the last two columns of Table 7 are combined. This translates into Equation 5 for calculating SB (PP is per cell and in gram, hence the division by 1000 to get to kilograms):

$$SB = \frac{PP * b}{a} / 1000 \quad \text{in kg/cell/year} \qquad \text{(Equation 5)}$$

Temperature and precipitation values used in these steps are calculated per grid cell using linear interpolation between the modern day temperature and precipitation distributions and the modeled distributions from the Last Glacial Maximum (LGM) using the global mean temperature for linear interpolation.

Carrying capacity and the number of ungulates

The model in HomininSpace assumes that 50% of the herbivore mass in any area are individuals from species that can be hunted and consumed. It is further estimated that from every kilogram of secondary ungulate biomass only 60% is consumable by hominins. The remaining 40% are bones, skin and assorted inedible items including hooves and horns. The amount of resources steer the number of hominin groups. If there are insufficient resources groups become hungry, and hungry groups have higher mortality rates and lower birth rates (see below).

1.7.2 The hominins

The basic agent or unit in HomininSpace is the hominin group. Groups of hominins can be created and can also be removed from the simulation (by leaving the simulation area, through extinction or through merging with another group). Values for parameters mentioned in this section can vary between simulations. Values are read from the Settings input file.

Demography

The availability of sufficient amounts of energy allows procreation of hominins in HomininSpace. This demographic process is implemented as a statistical process. This means that the system does not track individuals, but accounts for group dynamics with births and deaths recorded in floating point numbers for the whole group. This process mimics population statistics, but is implemented at the group level.

Three age categories are identified for the hominins in HomininSpace: pre-fertile, fertile and post-fertile with each their own mortality rate. For each category 50% of the hominins are assumed to be female. The demographic model in HomininSpace takes the pre-fertile period to last until 15 years of age and the fertile period to end at 40 years of age which signifies the start of the surprisingly long post-fertile stage. The HomininSpace model assumes that all fertile women bear one child every three years. This is implemented with 33% of all female adults giving birth to a newborn. The demographic subsystem is summarized in Table 8. A newly created Neandertal group (Initial group of produced by a Hominin Factory) theoretically would consist of around 25 people, with 15 being of pre-fertile age (Sørensen 2009).

Table 8 - Demographic parameters that define the three population segments.

	Age period	*Period length*	*Mortality rate (% per year)*	*Births (% per year)*
Pre-fertile	0-15	15	4	Infertile
Fertile	15-40	25	2	33
Post-fertile	40-		8	Infertile

Population growth continues until carrying capacity levels of the environment are reached. Then, for each year where population size exceeds carrying capacity growth rates are halved and death toll is doubled. Death and birth rates are applied at the end of every year. Population segments are assumed to be uniform (with for instance 20,000 women between 15 and 40 this means that 800 of them are of the age 20-21, 800 are 21-22 etc.). In each timestep for each age range the number of people in that age range divided by the length of that range move up to the next segment (with deaths subtracted). Newborns, to be added to the first category, are calculated using women of the adult age range.

The model in HomininSpace uses the following defaults for energy requirements, which are overall mean values, constant through time and space and irrespective of gender:

4000 kcal for the fertile section
3500 kcal for the post-fertile segment
3000 kcal for the pre-fertile part of the population

Meat provides 3000 kcal/kg.

Demography: new groups, group extinction and joining groups

Groups can procreate by splitting new groups from the mother group. When that happens a new viable group is created consisting of enough fertile adults of both sexes to ensure healthy growth potential. Groups look for opportunities to split up when the total group size exceeds 50 and become progressively unstable.

Groups can also join other groups when they become very small (< 3 individuals). When another group agrees to such a merger proposal the populations are added together and the joining group, now an empty administrative unit, is removed from the simulation. A group can refuse a join request because there are too few resources in the area, or when the group size already exceeds 50. If a group contains three or less reproductive adults it is considered not viable and if it does not join another group that year (see above) it is removed from the simulation.

Social interaction: home range and foraging range

Each grid cell in HomininSpace can contain a maximum of one group, at all times. When a group is present this cell is the *home range* of the group and will be defended against other groups at all costs. The model in HomininSpace prevents other groups from occupying an already claimed home range or even using it in foraging activities. One grid cell, and thus the home range, is 10x10km or 100km^2. Such a limited area hardly ever produces enough resources to sustain the subsistence needs of a group. Members of the group will have to constantly leave the home range and forage outside the protected area. They will collect resources only within the *foraging range* calculated for that group. A maximum foraging range is enforced, and this maximum is read from the scenario file.

Mobility: dispersal strategies and refugia

Groups move though the landscape and available energy steers their direction. In the model of HomininSpace there are two types of hominin mobility strategies: static mobility and dynamic mobility. Hominins that follow the static mobility strategy will move through the landscape until they find an area that suits their subsistence needs. This is where they will stay even when resources become depleted. There is one exception, and that is when the group size is getting below the viability threshold. When the group is about to go extinct they have the possibility to join another group if such a group is in the area and if that group will take them. Before they find an area to stay they behave exactly like dynamic groups. Dynamic mobility is characterized by constant movement, always directed towards the most favorable habitat in the area. As groups tend to consume the resources in the area they stay in groups will move almost every year towards other areas that have relatively more resources.

Movement for both strategies is identical. A group that is about to move will scan the environment for the best *habitat patches* in an area with a radius twice the foraging range. Habitat patches are defined as the sum of the available energy for each cell combined with its surrounding adjacent neighbors per cell (to avoid *death traps*: a grid cell with high productivity surrounded by cells with very low productivity like small islands and promontories).

When the destination cell with the highest productivity for this group has been found, the group moves onto that cell, the new foraging range for this location is computed and for each cell in the new foraging range the presence of the group is recorded. Thus other groups will know that they have to share the resources of those cells with this group.

References

Amiel, Joshua J., Reid Tingley, and Richard Shine. 2011. Smart moves: effects of relative brain size on establishment success of invasive amphibians and reptiles. *PLOS One* 6 (4):e18277.

Antón, Susan C., Richard Potts, and Leslie C. Aiello. 2014. Evolution of early Homo: An integrated biological perspective. *Science* 345 (6192):1236828.

Binford, Lewis R. 2001. *Constructing Frames of Reference: An Analytical Method for Archaeological Theory Building Using Hunter-Gatherer and Environmental Data Sets*: University of California Press, Berkeley and Los Angeles, California.

Bintanja, R., and R. S. W. van de Wal. 2008. North American ice-sheet dynamics and the onset of 100,000-year glacial cycles. *Nature* 454 (7206):869-872.

Coe, MJ, DH Cumming, and J Phillipson. 1976. Biomass and production of large African herbivores in relation to rainfall and primary production. *Oecologia* 22 (4):341-354.

Grimm, V., E. Revilla, U. Berger, F. Jeltsch, W. M. Mooij, S. F. Railsback, H. H. Thulke, J. Weiner, T. Wiegand, and D. L. DeAngelis. 2005. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *Science* 310 (5750):987-91.

Grimm, Volker, Uta Berger, Finn Bastiansen, Sigrunn Eliassen, Vincent Ginot, Jarl Giske, John Goss-Custard, Tamara Grand, Simone K Heinz, and Geir Huse. 2006. A standard protocol for describing individual-based and agent-based models. *Ecological modelling* 198 (1):115-126.

Keeley, Lawrence H. 1988. Hunter-gatherer economic complexity and “population pressure”: A cross-cultural analysis. *Journal of Anthropological Archaeology* 7 (4):373-411.

Kelly, R.L. 1983. Hunter-gatherer mobility strategies. *Journal of Anthropological Research* 39:277-306.

Lieth, Helmut. 1973. Primary production: Terrestrial ecosystems. *Human Ecology* 1 (4):303-332.

McNaughton, S. J., M. Oesterheld, D. A. Frank, and K. J. Williams. 1989. Ecosystem-level patterns of primary productivity and herbivory in terrestrial habitats. *Nature* 341 (6238):142-144.

Müller, Birgit, Stefano Balbi, Carsten M. Buchmann, Luís de Sousa, Gunnar Dressler, Jürgen Groeneveld, Christian J. Klassert, Quang Bao Le, James D. A. Millington, Henning Nolzen, Dawn C. Parker, J. Gary Polhill, Maja Schlüter, Jule Schulze, Nina Schwarz, Zhanli Sun, Patrick Taillandier, and Hanna Weise. 2014. Standardised and transparent model descriptions for agent-based models: Current status and prospects. *Environmental Modelling & Software* 55 (0):156-163.

Patterson, Samantha K., Aaron A. Sandel, Jordan A. Miller, and John C. Mitani. 2014. Data quality and the comparative method: the case of primate group size. *International Journal of Primatology*:1-14.

Railsback, Steven F. 2001. Concepts from complex adaptive systems as a framework for individual-based modelling. *Ecological Modelling* 139 (1):47-62.

Scherjon, F. 2012. SteppingIn – Modern humans moving into Europe – Implementation. In *CAA 2012 International Conference,* edited by D. Whitelock, W. Warburton, G. Wills and L. Gilbert. University of Southampton.

Sørensen, Bent. 2009. Energy use by Eem Neanderthals. *Journal of Archaeological Science* 36 (10):2201-2205.